

Université du Québec
Institut National de la Recherche Scientifique
Centre INRS-Institut Armand-Frappier

**Caractérisation, distribution et régulation du nouveau fimbria Efo chez
*Escherichia coli***

Par
Romain Coeurt

Mémoire présenté pour l'obtention du grade de
Maître sciences (Ms.)
en Virologie et Immunologie

Jury d'évaluation

Président du jury et
examineur interne

François Shareck
INRS-Institut Armand-Frappier

Examineur externe

George Szatmari
Département de microbiologie-immunologie
Université de Montréal

Directeur de recherche

Charles M. Dozois
INRS-Institut Armand-Frappier

RÉSUMÉ

La bactérie *Escherichia coli* est présente naturellement chez plusieurs animaux dont elle fait partie intégrante de la flore normale. Cependant, plusieurs souches peuvent causer des maladies infectieuses chez différents hôtes. Elles provoquent notamment des colibacilloses chez le poulet ; des diarrhées chez le porc ; des méningites, des infections urinaires et des diarrhées chez l'être humain. Ceci est dû à la plasticité et à l'hétérogénéité du génome de cette bactérie. Elle acquiert avec une certaine facilité des facteurs de virulence qui lui permettent d'avoir des hôtes et un tropisme cellulaire diversifiés.

Les fimbriae sont un de ces facteurs de virulence. Ils jouent un rôle critique pour l'établissement d'une majorité des infections causées par *E. coli*. Plusieurs gènes, spécifiques à la souche χ 7122 (O78:K80:H9) et qui ne sont pas présents chez la souche non pathogène K-12 χ 289, exprimant des fimbriae putatifs, ont été isolés *in vivo* lors d'infections de poulets. Donc, le but de cette étude est de caractériser un de ces nouveaux fimbriae présumés, de déterminer sa distribution parmi les multitudes de souches que compte l'espèce *E. coli* et enfin d'établir sa régulation.

Un cluster de quatre gènes présomptifs, codant pour quatre protéines ayant des homologies avec plusieurs opérons fimbriaires, fut identifié. Ces gènes ont été nommés *efoABCD*, pour « *exu-linked fimbrial operon* ». En effet, un criblage par PCR et une analyse *in silico* de différentes souches ont révélé que les gènes *efoABCD* sont localisés systématiquement entre les gènes *exuRT*, qui sont eux impliqués dans le transport des sucres hexuronates. Une analyse *in silico* a aussi permis de déterminer des rôles hypothétiques pour chaque protéine composant le fimbria Efo : EfoA (238 aa) pour la chaperonne ; EfoB (166 aa) pour la sous-unité majeure ; EfoC (900 aa) pour l'usher et EfoD (362 aa) pour la sous-unité mineure. Une étude comparative des séquences protéiques des « ushers » (placiers) de différents fimbriae et avec EfoC, montre une forte association du fimbria Efo avec les fimbriae Tcf de *Salmonella enterica* et Cbl de *Burkholderia cepacia*. Ces deux fimbriae font partie de la famille alternative de la voie du « chaperonne-usher ». De plus, la distribution phylogénétique de *efo* se fait en majorité chez le groupe B1, avec quelques sous-groupes chez les groupes A et D. Par ailleurs, une analyse des bases de données génomiques démontre la présence de trois types de séquences *efo* distincts.

Grâce à une fusion transcriptionnelle du promoteur hypothétique du cluster *efoABCD* et du gène *lacZ*, plusieurs formes de régulations ont été mises en évidence, notamment une régulation par la température, les sources de carbone et plus particulièrement par la protéine ArgR, de façon indirecte. Enfin, les gènes *efo* ont été clonés chez la souche afimbriaire ORN172. Des structures filamenteuses ont été observées par microscopie électronique et une diminution de la nage des bactéries, lorsqu'elles expriment les fimbriae Efo, est observée.

Ces résultats démontrent que le cluster *efoABCD* code pour un nouveau fimbria fonctionnel et qu'il est présent chez différents types de souches d'*E. coli*.

Romain Coeurt
Étudiant

Charles M. Dozois
Directeur de recherche

ABSTRACT

Escherichia coli is naturally present in many animals and is an integral part of the normal intestinal microbiota. However, several strains can cause infectious diseases in different hosts. They cause various diseases including colibacillosis in chickens, porcine diarrhoea, and meningitis, urinary tract infections and diarrhoea in humans. This is due to the heterogeneity and plasticity of the genome of this bacterium. It can readily acquire many virulence factors that promote adherence and survival in different hosts and diverse cell tropisms.

Fimbriae are one type of virulence factor. They play a critical role in adherence to host cells and tissues and the establishment of a majority of infections caused by *E. coli*. Several genes specific to avian pathogenic *E. coli* strain χ 7122 (O78:K80:H9) expressing putative fimbriae that are not present in the nonpathogenic strain *E. coli* K-12 χ 289 were transcribed *in vivo* during infection of chickens. Therefore, the aim of this study was to characterize a novel putative fimbriae, determine its distribution among a variety of strains, and investigate its expression under different conditions.

A cluster of four putative genes coding for four proteins with homology with several fimbrial proteins was identified. These genes were named *efoABCD* for exu-linked fimbrial operon, due to the association of this gene with the *exuRT* genes. Indeed, a PCR screening and *in silico* analysis of different strains showed that *efoABCD* genes are located systematically between *exuRT* genes, which are themselves involved in the transport of hexuronate sugars. *In silico* analysis also identified putative roles for each protein component of Efo fimbriae : EfoA (238 aa) for the chaperone ; EfoB (166 aa) for the major subunit ; EfoC (900 aa) for the *usher* and EfoD (362 aa) for a minor subunit. A comparative study of protein sequences of different fimbrial ushers with EfoC, shows a strong association of Efo fimbriae with Tcf fimbriae of *Salmonella enterica* and Cbl fimbriae of *Burkholderia cepacia*. Both fimbriae belong to the family of the alternative pathway of “chaperone-usher”. In addition, the phylogenetic distribution of *efo* was mostly associated to group B1, with some strains belonging to groups A and D. In addition, an analysis of the genomic databases revealed that Efo fimbriae comprise three distinct branches present among a variety of strains.

Through a transcriptional fusion of the putative promoter of the *efoABCD* fimbrial cluster and the *lacZ* gene, regulation of expression was investigated, including regulation by temperature, carbon sources and more particularly by the ArgR protein, indirectly. Finally, *efoABCD* genes were cloned in an afimbriates strain ORN172. Filamentous structures were observed by electron microscopy and a decrease in swimming motility when strains produced Efo fimbriae was also observed.

These results demonstrate that the *efoABCD* fimbrial genes code for a new functional fimbriae and that these gene clusters, encode three distinct fimbrial variants distributed among different *E. coli* strains.

REMERCIEMENTS

Tout d'abord, j'aimerais remercier mon directeur de recherche le Dr Charles Dozois, qui m'a offert un stage et une place dans son laboratoire pour effectuer ma maîtrise, ce qui m'a permis de confirmer mon désir de faire de la recherche en microbiologie. Merci pour ton soutien tout au long de mon séjour, de ta rigueur scientifique, de m'avoir permis de participer à plusieurs congrès scientifiques et surtout de m'avoir appris l'importance des contrôles expérimentaux !

Merci à tous les membres du laboratoire présents et passés pour votre aide et surtout d'avoir fait de mon passage dans le laboratoire quelque chose de particulier. J'aimerais ajouter une mention spéciale pour Sébastien Crépin qui a été bien plus qu'un simple collègue de laboratoire pour moi, à Sébastien Houle pour son aide et sa patience légendaire et à Gaëlle Porcheron pour son implication directe dans le projet.

Merci à ma famille, et surtout à mes parents, de m'avoir poussé à partir au Canada. Merci mille fois à tous pour votre soutien moral et financier (et alimentaire !). C'est une expérience que je n'oublierai jamais. Et j'espère que ce mémoire va vous permettre de comprendre un peu plus ce que j'ai fait pendant tout ce temps où j'étais loin de vous.

Enfin, merci aux organismes subventionnaires pour leur apport financier. La Fondation Universitaire Armand-Frappier de l'INRS pour ma bourse d'études, le Conseil de Recherches en Sciences Naturelles et en Génie du Canada (CRSNG) et la Chaire de recherche du Canada sur les maladies infectieuses bactériennes pour leurs subventions à la recherche.

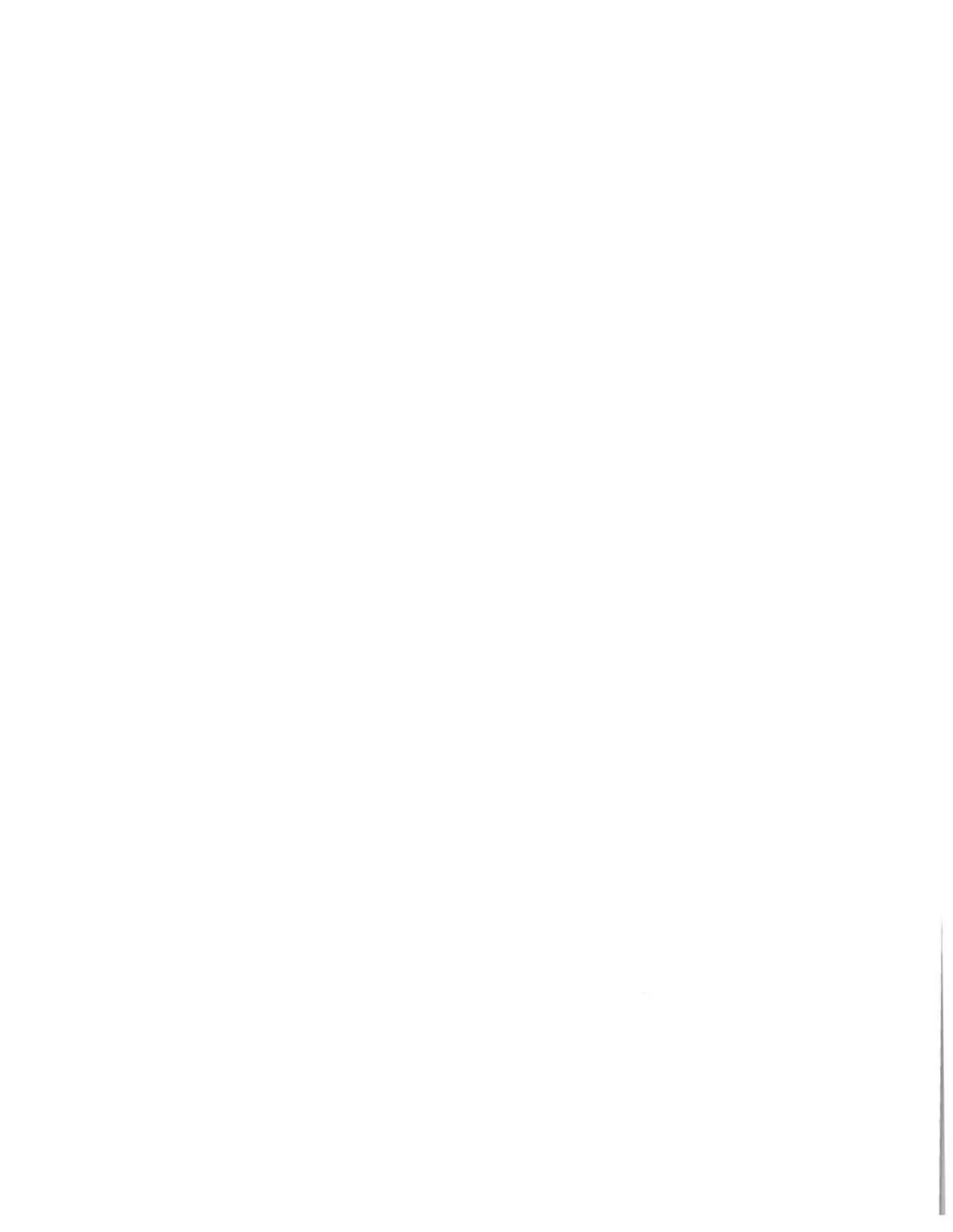


TABLE DES MATIÈRES

Identification du jury	i
Résumé	iii
Abstract	v
Remerciements	vii
Table des matières	ix
Liste des tableaux	xiii
Liste des figures	xv
Liste des sigles et abréviations	xvii
INTRODUCTION	1
REVUE DE LITTÉRATURE	3
1 DIVERSITÉ ET CLASSIFICATION D'ESCHERICHIA COLI	3
1.1 Schéma clonal	4
1.1.1 La collection ECOR	5
1.2 Classification sérologique	5
1.3 Classification par pathotypes	6
1.3.1 Les InPEC	7
1.3.2 Les ExPEC	11
1.4 Classification par groupes phylogénétiques	17
1.4.1 La méthode Clermont	19
1.4.2 La technique du MLEE	19
1.4.3 La méthode du MLST	20
2 LES FACTEURS DE VIRULENCE	20
2.1 Acquisition des facteurs de virulence	21
2.1.1 Les transferts horizontaux	21
2.1.2 Les ilots de pathogénicité	23
2.1.3 Les séquences d'insertion (IS)	23
2.2 Les facteurs de virulence associés aux UPEC	24
2.2.1 Les systèmes d'acquisition du fer	24
2.2.2 Les flagelles.....	28
2.2.3 Les toxines.....	28
2.2.4 L'adhésine afimbriaire lha	31
2.3 Les fimbriae chez les UPEC	31
2.3.1 Caractéristiques générales	31
2.3.2 Classification.....	33
2.3.3 Les fimbriae Type 1	34
2.3.4 Les fimbriae P	37

2.3.5	Régulation de leur expression	38
2.3.6	Comparaison des familles fimbriaires de la voie du « chaperonne-usher » et de la voie alternative	41
2.3.7	Les autres types de fimbriae importants pour les ITU	41
2.3.8	Conclusion sur les fimbriae	42
3	L'ARGININE ET LE RÉGULATEUR ARG R	42
3.1	Les différentes sources d'arginine	43
3.1.1	Le transport de l'arginine	43
3.1.2	La synthèse de novo de l'arginine et sa dégradation	45
3.2	La protéine ArgR	48
3.2.1	Structure et site de reconnaissance de ArgR	48
3.2.2	Fonctions de ArgR	49
3.3	Aperçu de la protéine ArgR chez les autres espèces bactériennes	50
	HYPOTHÈSES ET OBJECTIFS	53
	MATÉRIEL ET MÉTHODES	55
1	LISTE DES SOUCHES ET PLASMIDES UTILISÉS AINSI QUE LES MILIEUX ET LES CONDITIONS DE CROISSANCE	55
2	LISTE DES AMORCES UTILISÉES	57
3	MANIPULATIONS GÉNÉTIQUES ET DE L'ADN	60
4	CONSTRUCTION DES MUTANTS POUR ECOR72	60
5	CLONAGE ET SÉQUENÇAGE DES GÈNES DU CLUSTER <i>EFO</i>	62
6	DÉTERMINATION DE LA FORMATION D'UN OPÉRON PAR LES GÈNES <i>EFOABCD</i>	63
7	PRÉPARATION DE L'ANTISÉRUM SPÉCIFIQUE À EFOB	63
8	VISUALISATION PAR MET ET IMMUNOBUVARDAGE DE TYPE WESTERN	64
8.1	Visualisation par microscopie électronique à transmission (MET)	64
8.2	Extraction fimbriaire	64
8.3	Immunobuvardage de type Western	65
9	PRÉSENCE DE <i>EFOC</i> PARMIS LES SOUCHES D' <i>ESCHERICHIA COLI</i>	65
10	ANALYSES <i>IN SILICO</i>	66
10.1	Analyse par recherche de similitudes	66
10.2	Production des arbres phylogénétiques	66
10.3	Analyse du promoteur de <i>efo</i>	66
11	EXPÉRIENCES DE B-GALACTOSIDASE	66
12	EXPÉRIENCES DE RETARD SUR GELS DE POLYACRYLAMIDE	67
12.1	Expression et purification d'une protéine recombinante ArgR	67
12.2	Expérience de retard sur gel	68
13	TESTS PHÉNOTYPIQUES	69
13.1	Mesure de la production de biofilm	69

13.2	<i>Mesure de la nage des bactéries</i>	69
13.3	<i>Modèle d'infection du tractus urinaire murin</i>	69
14	TESTS STATISTIQUES.....	70
	RÉSULTATS	71
1	IDENTIFICATION DU CLUSTER DE GÈNES <i>EFOABCD</i>	71
2	VISUALISATION PAR MET ET PAR IMMUNOBUVARDAGE DE TYPE WESTERN	74
3	DISTRIBUTION DE <i>EFOC</i> PARMIS LA COLLECTION ECOR, DES SOUCHES PORCINES (PPEC) ET DES SOUCHES AVIAIRES (APEC).....	77
4	RELATION PHYLOGÉNÉTIQUE DES DIFFÉRENTES SÉQUENCES POUR LES GÈNES <i>EFOABCD</i> ET LEUR PROTÉINE RESPECTIVE	81
5	RELATION PHYLOGÉNÉTIQUE AVEC LES AUTRES FIMBRIAE DE LA VOIE DU « <i>CHAPERONNE-USHER</i> »	82
6	ANALYSE <i>IN SILICO</i> DU PROMOTEUR PRÉSUMÉ DU CLUSTER <i>EFOABCD</i>	84
7	RÉGULATION PAR DIFFÉRENTS FACTEURS ENVIRONNEMENTAUX DE L'EXPRESSION DU CLUSTER <i>EFO</i>	86
8	EXPÉRIENCES D'INTERACTIONS DE LA PROTÉINE ARGR SUR LE PROMOTEUR DE <i>EFOABCD</i>	90
9	TESTS PHÉNOTYPIQUES.....	95
9.1	<i>Production de biofilm et le fimbria Efo</i>	95
9.2	<i>Nage des bactéries et surexpression de efoABCD</i>	98
9.3	<i>Rôle des fimbriae Efo dans la colonisation du tractus urinaire</i>	98
	DISCUSSION	101
	CONCLUSION	109
	RÉFÉRENCES	111
	ANNEXES	139
	ANNEXE I.....	141
	ANNEXE II	142
	ANNEXE III	143
	ANNEXE IV	147
	ANNEXE V	148
	ANNEXE VI.....	150

LISTE DES TABLEAUX

TABLEAU 1 : LISTE DES SOUCHES ET DES PLASMIDES UTILISÉS DANS CETTE ÉTUDE.....	56
TABLEAU 2 : LISTE DES AMORCES UTILISÉES DANS CETTE ÉTUDE.....	60
TABLEAU 3 : DISTRIBUTION DES SOUCHES POSITIVES POUR <i>EFO</i> PARMI LA COLLECTION ECOR, DES SOUCHES PPEC ET APEC.....	78
TABLEAU 4 : DISTRIBUTION DES SOUCHES POSITIVES POUR <i>EFO</i> PARMI LES DIFFÉRENTS SÉROGROUPES DES APEC	79
TABLEAU 5 : RÉSULTAT DE L'ANALYSE <i>IN SILICO</i> DE L'ORF <i>EFOC</i> ₅₇₁₂₂ PAR LE LOGICIEL EN LIGNE IS FINDER (372) .	141
TABLEAU 6 : LISTE DES POOLS DE SOUCHES UTILISÉS ET LE RÉSULTAT DE LA PCR MULTIPLEX AVEC LES AMORCES ECS4-F/R	142
TABLEAU 7 : LISTE DES SOUCHES CRIBLÉES POUR LA PRÉSENCE DE <i>EFO</i>	143

LISTE DES FIGURES

FIGURE 1 : FACTEURS DE VIRULENCE ACQUIS PAR LES SOUCHES PATHOGÈNES.....	4
FIGURE 2 : SCHÉMA PATHOGÉNIQUE DES <i>ESCHERICHIA COLI</i> DE LA FAMILLE DES INPEC	10
FIGURE 3 : REPRÉSENTATION DES ÉTAPES D'UNE INFECTION DU TRACTUS URINAIRE	15
FIGURE 4 : OBSERVATION D'UNE CBI	16
FIGURE 5 : DISTRIBUTION PHYLOGÉNÉTIQUE DE LA COLLECTION DE SOUCHES ECOR	18
FIGURE 6 : ARBRE DE DÉCISION POUR LA MÉTHODE CLERMONT DE DÉTERMINATION DU GROUPE PHYLOGÉNÉTIQUE ..	19
FIGURE 7 : STRUCTURE CHIMIQUE DES SIDÉROPHORES PRODUITS PAR LES <i>E. COLI</i> PATHOGÈNES	25
FIGURE 8 : MODÈLE DE L'ASSEMBLAGE DU FIMBRIA TYPE I PAR LA VOIE DU « <i>CHAPERONNE-USHER</i> ».....	33
FIGURE 9 : SCHÉMA DU CLUSTER DE GÈNES <i>FIM</i>	35
FIGURE 10 : SCHÉMA DE LA RÉGULATION DE LA VARIATION DE PHASE DU FIMBRIA TYPE I PAR LES RECOMBINASES ..	39
FIGURE 11 : SCHÉMA SIMPLIFIÉ DE LA RÉGULATION DE LA VARIATION DE PHASE DU FIMBRIA P	40
FIGURE 12 : SCHÉMA DES TROIS TRANSPORTEURS DE L'ARGININE CHEZ <i>ESCHERICHIA COLI</i>	44
FIGURE 13 : BIOSYNTHÈSE DE L'ARGININE ET SON CATABOLISME PAR LA VOIE DES AST.....	47
FIGURE 14 : ORGANISATION GÉNÉTIQUE DE LA RÉGION ENTRE LES GÈNES <i>EXUR</i> ET <i>EXUT</i> CHEZ TROIS SOUCHES DIFFÉRENTES D' <i>ESCHERICHIA COLI</i>	54
FIGURE 15 : MIGRATION SUR GEL D'AGAROSE DES PCR, POUR LA DÉTERMINATION DE LA PRÉSENCE D'UN OPÉRON <i>EFOABCD</i>	73
FIGURE 16 : MIGRATION SUR GEL D'AGAROSE DES PCR, POUR LA DÉTERMINATION DE LA PRÉSENCE DES GÈNES <i>EFOAB</i> SUR UN MÊME TRANSCRIT	73
FIGURE 17 : PHOTOS DE MICROSCOPIE ÉLECTRONIQUE À TRANSMISSION DE LA SOUCHE QT1306.....	75
FIGURE 18 : PHOTOS DE MICROSCOPIE ÉLECTRONIQUE À TRANSMISSION DES SOUCHES QT3044 ET QT3045.....	75
FIGURE 19 : PHOTOS DE MICROSCOPIE ÉLECTRONIQUE À TRANSMISSION DE LA SOUCHE QT1307.....	76
FIGURE 20 : RÉVÉLATION DE L'IMMUNOBUVARDAGE DE TYPE WESTERN AVEC L'ANTICORPS ANTI-EFOB	77
FIGURE 21 : ARBRE PHYLOGÉNÉTIQUE DE LA COLLECTION ECOR À PARTIR DES SÉQUENCES MLST.....	80
FIGURE 22 : ARBRE PHYLOGÉNÉTIQUE DU GÈNE <i>EFOC</i> DES SOUCHES OBTENUES PAR BLASTN	82
FIGURE 23 : ARBRE PHYLOGÉNÉTIQUE DES PROTÉINES <i>USHER</i> DES FIMBRIAE DE LA VOIE « <i>CHAPERONNE-USHER</i> ».....	83
FIGURE 24 : ARBRE PHYLOGÉNÉTIQUE DES FIMBRIAE-A (ZOOM DE LA FIGURE 23).....	84
FIGURE 25 : ANALYSE DES PROMOTEURS PRÉSUMÉS DE L'OPÉRON <i>EFOABCD</i>	85
FIGURE 26 : EXPRESSION DE LA B-GALACTOSIDASE SOUS CONTRÔLE DU <i>PEFOABCD</i> ET SELON DIFFÉRENTES CONDITIONS DE CULTURE.....	87
FIGURE 27 : EXPRESSION DE L'ACTIVITÉ B-GALACTOSIDASE SELON DIFFÉRENTES SOURCES DE CARBONE	88
FIGURE 28 : ÉTUDE DE L'INFLUENCE DE LA PROTÉINE ARG _R SUR L'EXPRESSION DE L'ACTIVITÉ B-GALACTOSIDASE...90	90
FIGURE 29 : VÉRIFICATION DE LA PURIFICATION DE FAÇON SOLUBLE DE LA PROTÉINE ARG _R	92
FIGURE 30 : EXPÉRIENCE DE RETARD SUR GEL AVEC LA PROTÉINE ARG _R ET LES PROMOTEURS DE <i>EFO</i> ET DE <i>ARG_R</i>94	94

FIGURE 31 : EXPÉRIENCE DE RETARD SUR GEL AVEC ARGR, LE PROMOTEUR <i>EFO</i> ET DE L'ADN COMPÉTITEUR NON SPÉCIFIQUE.....	95
FIGURE 32 : RÔLE DU FIMBRIA <i>EFO</i> DANS LA FORMATION DE BIOFILM.....	97
FIGURE 33 : MESURE DE LA NAGE DES BACTÉRIES EN FONCTION DE LA CONCENTRATION EN AGAR ET DE L'EXPRESSION DES FIMBRIAE <i>EFO</i>	98
FIGURE 34 : CO-INFECTIONS DU TRACTUS URINAIRE MURIN AVEC <i>ECOR72</i> ET SES MUTANTS POUR LES FIMBRIAE TYPE I ET <i>EFO</i>	100
FIGURE 35 : PRÉDICTION DES PEPTIDES SIGNAUX POUR LES PROTÉINES <i>EFOABCD</i>	147
FIGURE 36 : COURBE DE CROISSANCE DE <i>ECOR72</i> ET SES MUTANTS POUR LES FIMBRIAE TYPE I ET <i>EFO</i>	150

LISTE DES SIGLES ET ABRÉVIATIONS

Aa : Acide aminé
ABC : « ATP-Binding cassette »
ADN : Acide désoxyribonucléique
ADNc : ADN complémentaire
ADP : Adénosine diphosphate
A/E : « Attachante et effaçante »
AIEC : *Escherichia coli* adhérent et invasif
AMPC : Adénosine monophosphate cyclique
ARN : Acide ribonucléique
ARNm : ARN messenger
ARNt : ARN de transfert
AST : Arginine succinyltransferase
A/T/C/G : Adénine / Thymine / Cytosine / Guanine
ATP : Adénosine-5'-triphosphate
AO : Arginine et ornithine
APEC : *Escherichia coli* pathogène aviaire
BAS : Bactériurie asymptomatique
BHE : Barrière hématoencéphalique
CBI : Communauté bactérienne intracellulaire
CNF1 : Facteur nécrosant cytotoxique 1
DAEC : *Escherichia coli* à adhérence diffuse
DAP : Acide meso-diaminopimélique
EAEC : *Escherichia coli* entéroaggrégative
EHEC : *Escherichia coli* entérohémorragique
EIEC : *Escherichia coli* entéroinvasive
EPEC : *Escherichia coli* entéropathogène
ETEC : *Escherichia coli* entérotoxigénique
ExPEC : *Escherichia coli* pathogène extra-intestinal
FV : Facteurs de virulence
HUS : Syndrome hémorragique urémique
InPEC : *Escherichia coli* pathogène intestinal
IPA : Ilots de pathogénicité
IS : « Insertion sequence »
ITU : Infection du tractus urinaire
KLH : « Keyhole limpet hemocyanin »
LAO : Lysine arginine et ornithine
LEE : Locus d'effacement des entérocytes
LPF : « Long polar fimbriae »

LPS : Lipopolysaccharide
LT : Enterotoxine thermolabile
ME : Membrane externe
MET : Microscopie électronique à transmission
MGE/DGE/TGE : Entérobactine mono/di/tri-glycosylée
MLEE : « Multilocus enzyme electrophoresis »
MLST : « Multilocus sequence typing »
MI : Membrane interne
NMEC : *Escherichia coli* de la méningite néonatale
ORF : « Open reading frame » Fenêtre de lecture ouverte
PCR : « Polymerase chain reaction »
PPEC : *Escherichia coli* pathogène porcine
SCOTS : « Selective capture of transcribed sequences »
SNC : Système nerveux central
SNP : « Single-nucleotide polymorphism »
SPATE : Sérine protéase autotransporteur des *Enterobacteriaceae*
SRI : Séquence répétée inversée
ST : Enterotoxine thermostable
STEAEC : *Escherichia coli* entéroaggrégative-hémorragique
STEC : *Escherichia coli* avec shiga toxines
UPEC : *Escherichia coli* pathogène urinaire
WT : « wild-type » soit sauvage

UNITÉ DE MESURE

x g : Nombre de g (gravitationnel)
Kb : Kilobase
Kd : Constante de dissociation
kDa : Kilodalton
nm : Nanomètre
Mb : Mégabase
M : Molaire
mM : Millimolaire
pb : Paires de bases
rpm : Rotation par minute
UFC : Unité formant colonie
μL : Microlitre
μM : Micromolaire
μm : Micromètre

INTRODUCTION

La découverte de la bactérie *Escherichia coli* fut, pour la première fois, initiée par Theodor Escherich, en effectuant l'isolement de bacilles Gram négatif dans des selles de nouveau-nés, en 1885 (1). Elle fait partie intégrante de la flore intestinale des mammifères. Mais là où certaines souches sont bénéfiques pour l'hôte, d'autres provoqueront des maladies infectieuses. Ces infections peuvent être différenciées, selon leur localisation, en deux grandes familles : les infections intestinales provoquées par les InPEC (pour *Escherichia coli* intestinal) et les infections extra-intestinales occasionnées par les ExPEC (pour *Escherichia coli* extra-intestinal).

Dans chaque famille, les souches sont subdivisées selon différents pathotypes. Par exemple, chez les ExPEC, l'appellation du pathotype dépendra de l'hôte infecté, comme les APEC pour les infections chez le poulet ou les PPEC pour le porc. Ainsi que de la localisation de l'infection : NMEC pour les méningites et UPEC pour les infections du tractus urinaire.

Ces différences d'hôtes ou de tropismes cellulaires peuvent simplement s'expliquer par la spécialisation de ces souches, grâce à l'acquisition de facteurs de virulence spécifiques. En effet, le génome de la bactérie *E. coli* est particulièrement plastique. Les différences de taille des génomes entre une souche non pathogène et pathogène sont remarquables (2). Par exemple, MG1655, qui est la première souche non pathogène à avoir été séquencée (3), possède un génome de 4,64 Mb, alors que des souches pathogènes, comme IAI39 (UPEC) et O157:H7 (EHEC), possèdent respectivement un génome de 5,13 Mb et 5,62 Mb. Ce qui représente une différence de près de 1000 gènes entre les souches pathogènes et non pathogènes. Ceci est dû à l'intégration de larges séquences d'ADN, appelées îlots de pathogénicité, présentes sur des transposons, des plasmides ou encore des bactériophages. De plus, d'autres mécanismes tels que des transferts horizontaux entre espèces, des recombinaisons intrachromosomiques et des duplications de gènes peuvent être à l'origine d'une multiplication et de la dérivation des facteurs de virulence.

Dans le but d'identifier de nouveaux facteurs de virulence, la technique SCOTS (« Selective Capture Of Transcribed Sequences »), qui est basée sur le principe de comparaison d'expression *in vivo* de transcrits entre, dans cette étude, la souche pathogène χ 7122 (APEC) et la souche K-12 non pathogène χ 289, a été utilisée lors d'infections de poulets. Plusieurs

séquences, codant potentiellement pour des fimbriae, ont été mises alors en évidence (4). Les fimbriae sont un de ces facteurs de virulence essentiels pour l'initialisation et le développement des maladies infectieuses. Ils sont notamment critiques lors des infections du tractus urinaire (5, 6). Or, un des clusters identifiés lors de cette étude chez χ 7122, mais interrompu par une séquence d'insertion (IS), a été retrouvé intact (non interrompu) chez la souche UPEC ECOR72 (appartenant à la collection ECOR (7)).

L'objectif de ce projet a été donc de faire la caractérisation complète de ce cluster de gènes, dénommé *efo* pour « exu-linked fimbrial operon ». Un premier aspect phylogénétique a été abordé afin de déterminer la prévalence de la présence de ce cluster parmi différents pathotypes, ainsi qu'une étude comparative des fimbriae déjà connus chez les différentes espèces bactériennes avec le fimbria putatif Efo. Le deuxième aspect portait sur son expression, à la suite d'un clonage sur plasmide à haut nombre de copies, ainsi que les différentes formes de régulation. Enfin, la détermination du phénotype composait le dernier aspect du projet de maîtrise, avec une approche basée sur la souche ECOR72 et les infections du tractus urinaire.

REVUE DE LITTÉRATURE

1 **Diversité et classification d'*Escherichia coli***

Escherichia coli fait partie de la grande famille des *Enterobacteriaceae*, qui comprend à savoir les genres *Shigella*, *Salmonella* et *Klebsiella*. C'est un bacille à Gram négatif avec des flagelles péritriches et c'est la bactérie anaérobie facultative prédominante de la flore intestinale humaine (8). En effet, l'être humain est colonisé dès les premières heures après sa naissance. Demeurant dans l'intestin, les souches sont, pour la plupart, non pathogènes et en symbiose avec l'hôte (9). Mais, certaines souches peuvent coloniser d'autres organes et provoquer des infections et septicémies.

Les souches pathogènes sont différenciées en deux catégories distinctes : les InPEC qui occasionnent des infections intestinales et les ExPEC qui s'établiront dans de multiples organes et hôtes. Dans chacune de ces catégories, il existe différents pathotypes qui correspondent à un type d'infection selon le mode de colonisation, l'organe atteint ou encore l'hôte cible (9). Outre un noyau de gènes conservés essentiels à la vie de la bactérie, les souches pathogènes ont acquis diverses larges séquences d'ADN, qui contiennent notamment des facteurs de virulence. Ces séquences leur confèrent un pouvoir pathogène et donc une capacité à envahir des milieux non propices à la colonisation (Fig. 1) (10).

Ceci démontre une grande diversité parmi l'espèce *Escherichia coli* et nécessite, par conséquent, des systèmes de classements. Ils dépendront soit du pathotype, soit du polymorphisme antigénique ou soit de la variation allélique.

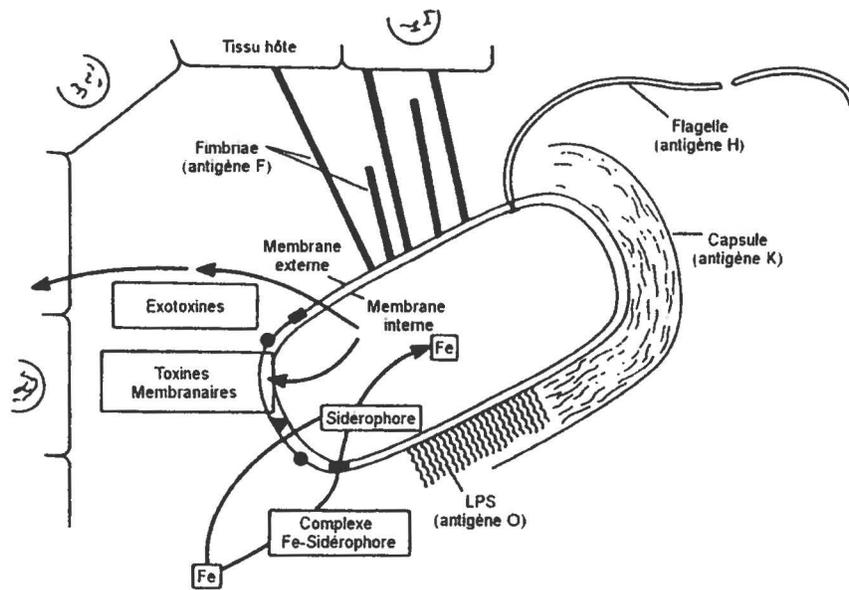


Figure 1 : Facteurs de virulence acquis par les souches pathogènes

Les facteurs de virulence sont le plus souvent associés avec le matériel génétique acquis par les souches pathogènes. Les toxines, LPS, toxines membranaires, sidérophores (Fe pour Fer) et les fimbriae sont des acteurs clés dans différentes pathologies. Modifiée de J R Johnson (1991).

1.1 Schéma clonal

Lors de comparaisons des génomes entre des souches commensales (comme K-12) et des souches pathogènes, il s'est révélé que, par exemple, près de 75 % des gènes exprimés chez la souche MG1655 (K-12) sont retrouvés chez CFT073 (UPEC) (9). Ils forment un noyau de gènes conservés entre les différentes souches d'*Escherichia coli*, ce qui donne comme information que même si cette espèce présente une grande diversité génétique, toutes les souches seraient liées génétiquement.

Des analyses, fondées sur l'hétérogénéité génétique et protéique, ont permis de confirmer cette hypothèse. En effet, ces techniques, décrites plus bas, ont montré que les allèles étudiés ne sont pas disséminés parmi la population bactérienne de façon aléatoire. Mais que des isolats prélevés en des lieux, des temps et des hôtes différents présenteront des profils alléliques identiques et donc auront une liaison phylogénétique forte. Étant donné qu'une des seules explications pour cette observation est une origine commune, la population bactérienne d'*Escherichia coli* aurait une structure clonale (12-15). De plus, des infections spécifiques peuvent être causées par des groupes de clones liés : les facteurs de virulence, outre le fait qu'ils

proviennent d'éléments d'ADN mobiles (voir chapitre « 2.1 : Acquisition des facteurs de virulence »), suivent aussi ce même schéma clonal (13, 16).

1.1.1 La collection ECOR

Dans le but d'avoir une large représentation de la diversité d'*Escherichia coli*, Ochman *et al.* (1984) ont effectué des prélèvements sur des fèces d'humains et d'animaux sains ou durant des infections du tractus urinaire, en différents lieux géographiques (7). Depuis, plusieurs analyses génétiques et phénotypiques ont été effectuées. Ces résultats ont démontré et démontrent encore que cette collection est une bonne représentation de la diversité génétique de cette espèce (12).

1.2 Classification sérologique

La classification sérologique est l'une des premières classifications établies par Kauffmann (1947). Elle se base sur trois antigènes présents à la surface des bactéries: les antigènes O pour une courte chaîne polysaccharidique du LPS (lipopolysaccharide), les antigènes H pour des facteurs sérologiques présents sur la flagelline (protéine des flagelles) et les antigènes K pour des facteurs polysaccharidiques acides de la capsule (17). À ce jour, plus de 175 antigènes O, 56 antigènes H et 80 antigènes K furent découverts (18).

L'antigène O, avec l'antigène H, définit le sérotype de la souche testée. Il est composé d'un polymère de glycanes, qui fait lui-même partie du constituant majeur de la membrane externe : le lipopolysaccharide. Le LPS est formé de trois structures distinctes : le lipide A qui est l'ancrage du LPS à la membrane externe, le « noyau » qui est un oligosaccharide phosphorylé non répétitif et enfin l'antigène O qui est la région variable du LPS. Les différences notées entre les multiples antigènes O proviennent de leur composition en sucres, des liaisons entre les sucres et leurs substitutions par des résidus non glucidiques (19). La détermination du sérotype se fait par agglutination des bactéries, grâce à un antisérum de lapin, préalablement immunisé avec les différents antigènes O connus. Mais, cette technique présente plusieurs limitations venant du fait que de plus en plus d'antigènes O sont découverts, ce qui demande une plus grande production d'antisérum. De plus, si la bactérie possède une capsule (antigène K), cet antigène sera masqué. Par ailleurs, une multitude de réactions croisées entre les antigènes O d'*E. coli*, mais aussi avec

les antigènes présents chez les genres *Shigella*, *Salmonella* et *Klebsiella*, ont été rapportées (17, 20, 21).

L'antigène H est un élément des flagelles et plus précisément de la flagelline qui les compose. Les flagelles assurent la mobilité des bactéries. Ils sont formés de trois différentes structures : le corps basal sert à l'ancrage du flagelle à la cellule, le crochet pour le transfert du mouvement et enfin le filament flagellaire qui est constitué de nombreux monomères de la flagelline. Comme pour l'antigène O, cet antigène est identifié par agglutination des cellules en utilisant des antisérums (22). De même, plusieurs réactions croisées peuvent se produire entre les différents antigènes H. Les *E. coli*, possédant certains sérogroupes (par exemple O157:H7), peuvent être associés avec certains types d'infections. Cependant, ce n'est pas le sérotype qui est directement à l'origine de la virulence des souches, ils font juste office de marqueurs sérologiques (20, 23).

L'antigène K est le polysaccharide qui compose la capsule des bactéries. Les différences de monomères, de liaisons entre eux et le remplacement des résidus glucose par d'autres résidus, constituent la diversité des antigènes K. Mais, étant donné que la capsule n'est pas un élément vital chez *Escherichia coli*, plusieurs souches peuvent ne pas posséder d'antigène K. Ils sont classés selon quatre groupes qui dépendent de l'organisation génétique et de la régulation des gènes responsables de la production de la capsule (24).

Un dernier antigène, l'antigène F, a été proposé par Orskov *et al.* (1983) qui se base sur les différents fimbriae (25). Ils sont responsables de l'adhérence des bactéries aux cellules de l'hôte. On notera que plusieurs groupes sérologiques peuvent être attribués à un seul fimbriae. C'est le cas notamment du fimbria P qui possède 12 groupes sérologiques qui lui sont propres (26). Cet antigène permet une meilleure dissociation entre les sérogroupes définis par les antigènes O et H.

1.3 Classification par pathotypes

Une autre classification a vu le jour en même temps que la mise en évidence de la responsabilité d'*Escherichia coli* dans différentes infections. En effet, ce classement se base sur les hôtes, le tropisme cellulaire et le mode d'infection pour déterminer les pathotypes. Il existe deux grandes familles de pathotypes, la première qui regroupe toutes les souches qui provoquent des infections intestinales : les InPEC. Et la seconde regroupant celles qui provoquent des

infections extra-intestinales : les ExPEC. Par contre, étant basé sur le type de pathologie, les souches non pathogènes ne sont pas comprises dans ce classement (9). Seuls ici les pathotypes principaux seront décrits.

1.3.1 Les InPEC

Les pathotypes, faisant partie de cette famille, causent tous des diarrhées. Mais des stratégies différentes de colonisation, des facteurs de virulence et des symptômes caractéristiques feront que ces souches seront classées différemment (Fig. 2).

Les souches entéropathogènes (EPEC) sont les premières à avoir été décrites (27). Cette infection se caractérise par une diarrhée aqueuse persistante, pouvant aller jusqu'à 120 jours d'hospitalisation (20, 28). Elle touche principalement les enfants en bas âge, mais aussi les adultes, dans les pays développés à cause d'une contamination fécale-orale (20, 28). De plus, elle est potentiellement mortelle dans les pays en voie de développement. Une caractéristique histopathologique est particulièrement associée avec les EPEC : c'est la lésion « attachante et effaçante » (A/E). Ces bactéries vont se lier intimement avec les cellules épithéliales intestinales, réassembler le cytosquelette de ces cellules pour former des structures ressemblant à des piédestaux et par conséquent, détruire les microvillosités. Ceci est possible grâce au « locus d'effacement des entérocytes » (LEE), qui n'est, bien sûr, pas présent chez les souches d'*E. coli* commensales (9).

Les EHEC, pour les *Escherichia coli* entérohémorragiques, ont été associés, dans les pays développés, avec « la maladie du hamburger ». En effet, après avoir mangé des hamburgers crus dans les chaînes de restauration rapide, les personnes infectées développaient des diarrhées hémorragiques aqueuses (dysenteries), associées avec des syndromes hémolytiques urémiques (HUS) (20, 29). Le principal réservoir de ce pathotype est le tube digestif des bovins, d'où la relation entre la consommation des viandes bovines et les dysenteries. Le sérotype O157:H7 est celui qui est le plus fortement lié avec cette infection. Ce qui montre une certaine relation entre sérotype et présence des facteurs de virulence (se référer au chapitre « 1.1.1 : Schéma clonal ») (20). Les EHEC, comme pour les EPEC, possèdent le LEE. Mais un facteur de virulence clef est identifié en plus chez les EHEC : la Shiga toxine. Elle est notamment responsable du HUS. Toutefois, il faut faire attention de ne pas confondre le pathotype EHEC

avec le pathotype STEC. Les souches, faisant partie de ce dernier, ne possèdent pas le LEE mais expriment la Shiga toxine (9).

Le pathotype ETEC est retrouvé préférentiellement chez les enfants de moins de cinq ans dans les pays en voie de développement. De plus, les ETEC sont caractéristiques de la diarrhée dite « du voyageur », lors de séjours dans ces pays. Cette infection provient d'une transmission fécale-orale (30). Elle se traduit par une diarrhée aqueuse, avec un temps d'incubation très court de 14 à 50 heures (20). C'est le petit intestin qui est colonisé. Les ETEC produisent au moins une entérotoxine thermolabile (LT) ou thermostable (ST). Ces toxines sont responsables d'une abondante sécrétion intestinale (9).

Les souches entéroaggrégatives (EAEC) provoquent des diarrhées persistantes aqueuses, notamment chez les enfants dans les pays développés et en voie de développement (31). Elles entraînent une augmentation de sécrétion du mucus par la libération d'entérotoxines (différentes des toxines LT et ST) et de cytotoxines (9). Cela conduit à l'apparition d'un biofilm. C'est ce qui les différencie des EPEC, qui eux forment des microcolonies à la surface des cellules de l'hôte (20).

Les *Escherichia coli* entéroinvasives (EIEC) peuvent, comme leur nom l'indique, envahir les cellules de l'hôte (28). Elles sont responsables de diarrhées aqueuses et mucoïdes et occasionnellement des dysenteries qui sont non distinguables de celles provoquées par les ETEC (9). Ces souches, comme la plupart des pathotypes de cette famille, se transmettent par la nourriture ou l'eau contaminée. C'est plus une infection que l'on retrouve dans les pays en voie de développement, mais des cas d'éclairs peuvent aussi apparaître dans les pays développés, comme le cas d'une épidémie associée à un restaurant au Texas (20). Ce sont des souches qui sont très proches biochimiquement, génétiquement et pathologiquement des *Shigella* spp. Leur mode d'infection se découpe en trois étapes : la première étape constitue l'invasion des cellules épithéliales de l'hôte, la seconde est la multiplication des bactéries dans ces cellules et enfin grâce au réarrangement du réseau de microfilaments d'actine, les bactéries se propagent aux cellules adjacentes.

Les DAEC, pour *Escherichia coli* à adhérence diffuse, ont été longtemps classés parmi les EPEC. Cependant, leur stratégie d'infection est quelque peu différente. Grâce à des effecteurs cytopathogènes fimbriaires, la bactérie provoque un développement d'excroissance des

microvillosités qui vont l'entourer (9), ce qui donne un phénotype d'adhérence diffuse aux cellules de l'hôte, différent des EPEC (32). La pathologie produite est une diarrhée aqueuse avec absence de sang et de leucocytes. Elle touche plus particulièrement les enfants qui ont un âge entre un et cinq ans (20).

Découvertes pour la première fois en 1999 (33), les bactéries du pathotype AIEC sont associées avec la maladie de Crohn (comme étant une conséquence et non une cause (34)) et les diarrhées canines (35, 36). Les AIEC peuvent véritablement envahir les cellules des plaques de Peyer qui sont l'un des constituants du tissu lymphoïde de l'intestin, grâce notamment au LPF et aux fimbriae Type 1 (34, 36). Le mécanisme d'invasion est très similaire aux EIEC avec une réplication intracellulaire (36), de plus, et c'est ce qui caractérise ce pathotype, les AIEC peuvent envahir les macrophages et survivre. Aucune nécrose, apoptose ou modification des macrophages n'est observée, contrairement aux autres pathotypes. En outre, les macrophages restent toujours actifs (36, 37).

Un nouveau pathotype a fait son apparition pendant l'épidémie de 2011 en Europe. La souche responsable, de sérotype O104:H4, est un hybride entre les EAEC et les EHEC, soit un *E. coli* entéroaggrégative-hémorragique (STEAEC). Elle provoque des diarrhées hémorragiques et des HUS. Ce pathotype est un exemple de la grande plasticité du génome d'*E. coli* qui lui permet, dans ce cas, d'acquérir par transfert horizontal des facteurs de virulence des EAEC (38, 39).

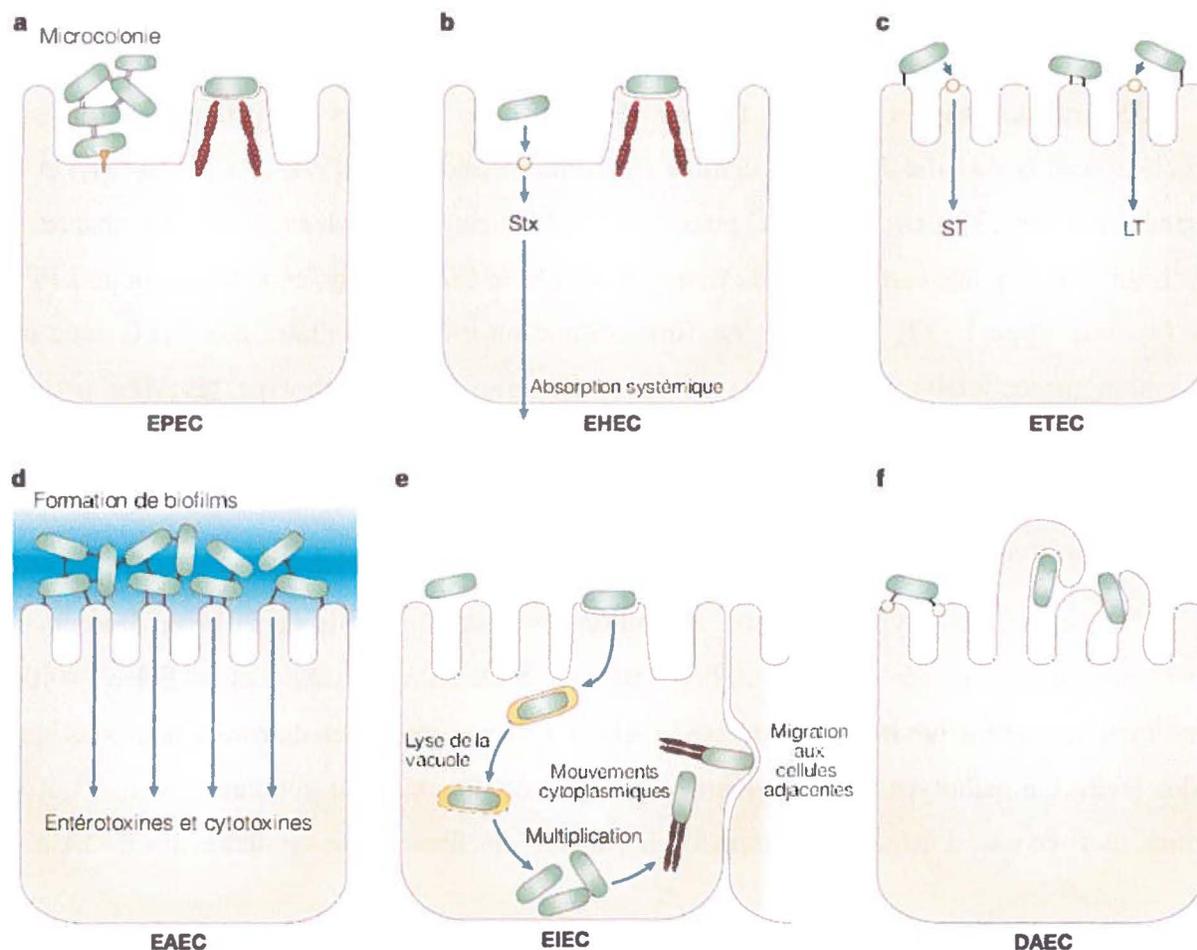


Figure 2 : Schéma pathogénique des *Escherichia coli* de la famille des InPEC

Modifiée de Kaper *et al.* (2004). Les six pathotypes de la famille des InPEC interagissent différemment avec les cellules de l'hôte. Ces interactions sont schématisées ici. **a)** Les EPEC adhèrent aux cellules en formant des microcolonies et en détruisant les microvillosités. Ils effectuent un réarrangement du cytosquelette pour former un piédestal. **b)** Les EHEC forment aussi un piédestal et détruisent les microvillosités. De plus, ils produisent la Shiga toxine (Stx). **c)** Les ETEC produisent deux entérotoxines (ST et LT). **d)** Les EAEC forment des biofilms et sécrètent des entérotoxines et cytotoxines. **e)** Les EIEC envahissent la cellule hôte grâce à une vacuole, la détruisent, se multiplient et enfin peuvent migrer aux cellules adjacentes. **f)** Les DAEC induisent la production d'une excroissance qui les enveloppe (9).

1.3.2 Les ExPEC

Seulement aux États-Unis, plus de neuf millions de personnes par an contractent une infection extra-intestinale, due à la bactérie *Escherichia coli* (40). Ces souches, qui font partie de la grande famille des pathogènes extra-intestinaux (ExPEC) (41), présentent une particularité : elles sont souvent retrouvées dans la flore commensale intestinale humaine sans causer de maladie (contrairement aux InPEC), par contre, dès qu'elles colonisent un autre organe ou un nouvel hôte, elles provoquent des infections avec une morbidité et une mortalité assez élevées. Les souches, appartenant à ces pathotypes, possèdent des facteurs de virulence et composent des groupes clonaux qui leur sont propres. Ainsi, elles peuvent être différenciées des souches strictement commensales et des souches pathogènes intestinales (40-42). Plusieurs pathotypes sont associés avec cette famille. Chacun provoquant une infection caractéristique qui dépendra, notamment, du tissu et de l'hôte impliqués. Les pathotypes NMEC, APEC et UPEC représentent la majeure partie des ExPEC et seront les seuls décrits ci-dessous (43).

Les NMEC

Les NMEC, pour *Escherichia coli* de la méningite néonatale, sont responsables de l'une des maladies les plus mortelles et handicapantes, particulièrement pendant la période néonatale. Le taux de mortalité se situe entre 15 et 40 %, tandis que les bébés qui survivent sont, dans 50 % des cas, atteints de séquelles neurologiques, comme des convulsions, des paralysies ou encore des hydrocéphalies (44, 45).

Chez les nouveau-nés, les souches de ce pathotype proviennent de la flore commensale de la mère ou de l'environnement (46). Elles colonisent le tractus intestinal puis, par translocation à travers la paroi, elles envahissent le système circulatoire sanguin (46-48). Il a été constaté une corrélation entre le niveau de la bactériémie (minimum de 10^3 UFC/mL) et le développement d'une méningite. Mais cela n'est pas suffisant, car des facteurs de virulence clés sont nécessaires pour l'invasion du système nerveux central (SNC) (9, 46, 47). Pour atteindre le SNC, les bactéries ont besoin de traverser la barrière hématoencéphalique (BHE). Pour cela, elles effectuent une translocation par une vacuole fermée, sans multiplication cellulaire, ni disruption de la BHE. L'interaction entre les bactéries et les cellules de la BHE induit la production de pseudo-microvillosités, ce qui facilite la translocation (49). Les facteurs spécifiques, pour cette opération, sont principalement : des fimbriae et des facteurs afimbriaires, pour l'adhérence aux

cellules de la BHE ; la capsule et des protéines de la membrane externe, pour la résistance au sérum et le transport dans la cellule hôte (9). Plusieurs études épidémiologiques ont montré que seulement certains sérogroupes O, associés dans 80 % des cas avec la capsule K1, sont responsables des méningites néonatales (48, 50).

Les APEC

Les *Escherichia coli* peuvent infecter une multitude d'hôtes animaux, comme les veaux, les agneaux, les porcs... (51) Mais, ce sont les *E. coli* pathogènes aviaires (APEC) qui possèdent la plus grandes incidences sur l'économie agricole. En effet, ces souches occasionneront 20 % de mortalité chez les poulets infectés, diminueront leur croissance et donc réduire leur valeur marchande et enfin augmenter le taux de refus aux abattoirs (52, 53).

Les animaux s'infectent en respirant des particules fécales contaminées avec les APEC (présentes naturellement dans la flore intestinale de la volaille), qui sont contenues dans la poussière (52, 53). Les infections sont assez variées, elles incluent les péritonites, les omphalites, le syndrome de la tête enflée, la cellulite et l'aérosacculite qui peut déboucher sur une septicémie. Fréquemment, pour que les poulets développent ce genre d'infections, ils sont déjà infectés par *Mycoplasma gallisepticum* ou par un virus. De même, les stress environnementaux peuvent les rendre plus disposés ou peuvent accentuer la maladie (52).

Toute une panoplie de facteurs de virulence est exprimée par ces souches APEC : les fimbriae, notamment, joueront un rôle prépondérant dans la mise en place de l'infection (52, 54), comme les sidérophores (55, 56) et les toxines (52). Les sérotypages des isolats d'animaux malades ont révélé une prévalence des sérogroupes O1, O2 et O78 (15 à 61 % selon les études) (53, 57). Or, le séro groupe O78 se retrouve couramment chez les ETEC infectant l'être humain, avec des clones très proches phylogénétiquement des APEC. Évidemment, ils possèdent d'autres facteurs de virulence. Ceci donne comme indication que les animaux seraient un réservoir potentiel de souches ETEC (57). À cela s'ajoute que plusieurs études ont montré un risque de zoonose direct, impliquant des souches APEC et PPEC (souches porcines) et pouvant provoquer des infections du tractus urinaire ou encore du système respiratoire (58-61).

Les UPEC

Chez les humains, le tractus urinaire est le site d'infection le plus commun des ExPEC. Les souches, infectant le système urinaire, sont qualifiées comme étant des *E. coli* uropathogènes

(UPEC) (40). Au cours de leur vie et pour des raisons anatomiques, près de 40 % des femmes sont susceptibles de contracter une infection du tractus urinaire (ITU) contre 12 % chez les hommes. Cela représente, seulement aux États-Unis, plus de onze millions de femmes infectées par an et plus de 2,5 milliards de dollars de dépenses pour cette seule infection (6, 62), et *Escherichia coli* est à lui seul responsable de près de 80 % de ces ITU (63).

Par des analyses sérologiques et phylogénétiques, il a été constaté que ces souches font partie intégrante de la flore intestinale humaine, sans pour autant provoquer de maladie. En effet, le côlon est le premier réservoir de souches UPEC. Ceci est notamment dû à la distance minimale entre l'anus et la région périurétrale, qui est d'autant plus faible chez la femme. De plus, le vagin, préalablement colonisé par ces mêmes souches, peut devenir un réservoir potentiel (64-66). À cela s'ajoute le fait que l'urètre est plus grand chez l'homme que chez la femme, donc le délai est plus long pour atteindre la vessie, ce qui favorise l'élimination des bactéries par le flux d'urine. Outre le fait que des différences morphologiques influencent la prédisposition aux ITU, plusieurs facteurs de risque ont été établis : par exemple, l'utilisation de différents moyens de contraceptions, tels que le condom, le diaphragme ou les spermicides, peut augmenter cette susceptibilité. De même, les activités sexuelles et la prise de certains antibiotiques sont un de ces facteurs (66-68).

Dans les six mois après une première infection, une femme a trois fois plus de risque de développer une seconde infection urinaire par rapport à une femme n'ayant jamais eu d'infection urinaire (69). En effet, même si la première infection est traitée, la souche impliquée est toujours présente dans la flore intestinale ou vaginale de la personne. De plus, ce taux de récurrence est inégal selon : l'âge de la personne (les femmes de plus de 55 ans sont plus sensibles), l'âge de la contraction de la première ITU, l'activité sexuelle, le nombre de partenaires différents, l'utilisation de moyens de contraception... (65, 70)

Trois types d'ITU sont distingués : les ITU symptomatiques non compliquées, les ITU symptomatiques compliquées et les bactériuries asymptomatiques (BAS). Une ITU est dite compliquée lorsque la personne présente des dysfonctionnements physiologiques tels qu'une malformation du tractus urinaire, du diabète ou que la personne possède un cathéter urinaire (63, 68, 71, 72). Mais étant donné la plus forte prévalence de l'ITU non compliquée, c'est principalement celle-ci qui sera décrite dans les prochains paragraphes.

L'ITU non compliquée est une infection ascendante qui se décompose en plusieurs étapes (Fig. 3). Elle débute par une contamination de l'air périurétrale, par une souche UPEC présente dans la flore intestinale ou vaginale. Ces bactéries migreront vers la vessie, à travers l'urètre, et elles se fixeront à l'urothélium. Cette fixation est notamment médiée par les fimbriae Type 1. L'interaction entre les bactéries et les cellules épithéliales de l'hôte provoque l'exfoliation de ces cellules et leur apoptose. Ceci est considéré comme étant un mécanisme de défense inné de l'hôte (73). Malgré un environnement peu favorable à leur croissance, ces souches se multiplieront et envahiront la vessie jusqu'à provoquer une cystite. Cela se traduit par une inflammation de la vessie, de la présence de bactéries dans les urines (bactériurie) et des symptômes variés comme la possible présence de sang, la sensation d'urgence et la difficulté de produire une miction. Si la cystite est mal ou non soignée, les bactéries, grâce aux flagelles, se déplaceront vers les reins afin de les coloniser (74). L'un des facteurs de virulence principaux est le fimbria P, qui permet la fixation à l'endothélium rénal. Cela entraînera le développement d'une pyélonéphrite, accompagnée de symptômes tels que de la fièvre, des douleurs aux flancs et des vomissements. De même, si elle n'est pas soignée, des dommages importants apparaîtront au niveau des reins et permettront le passage des bactéries dans la circulation sanguine. Cela termine l'ITU par une septicémie et une mort possible du malade (9, 18).

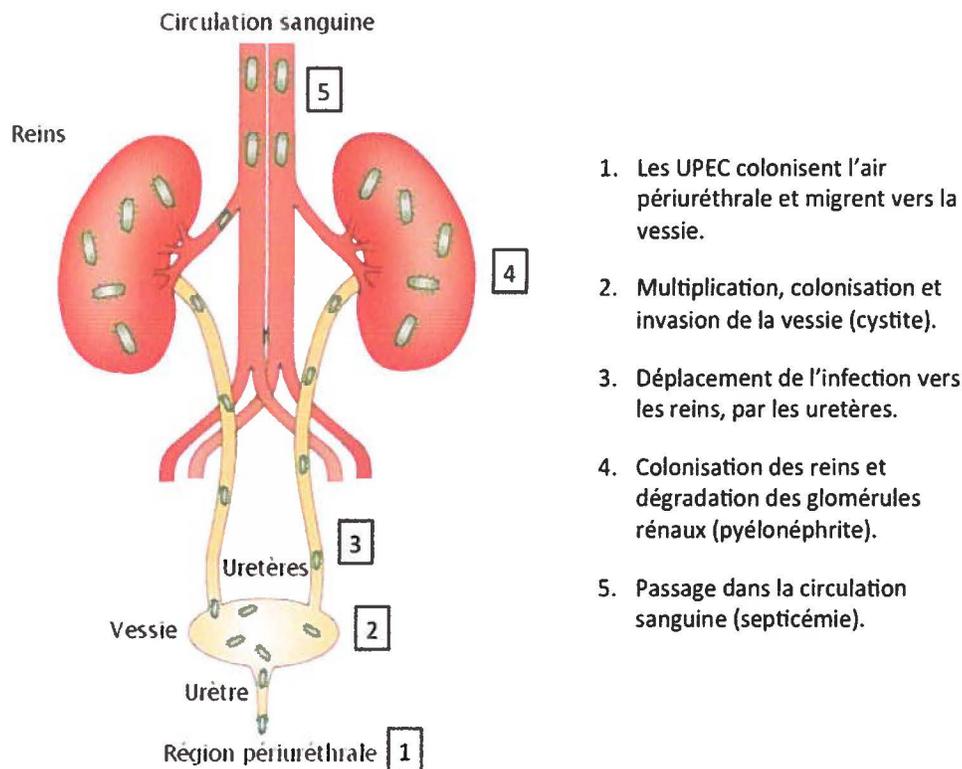


Figure 3 : Représentation des étapes d'une infection du tractus urinaire

Modifiée de Kaper *et al.* (2004). En vert sont représentées les bactéries évoluant à travers le système urinaire (9).

Il est maintenant reconnu que certaines souches d'*Escherichia coli* sont des pathogènes intracellulaires opportunistes (75). C'est-à-dire que, certaines UPEC, peuvent envahir les cellules épithéliales de la vessie, se multiplier et former des inclusions qui se nomment des CBI, pour Communautés Bactériennes Intracellulaires (Fig. 4) (73, 76-79). Une preuve assez importante de l'existence des CBI a été donnée par une étude, où les urines, de femmes atteintes d'une ITU, contenaient des filaments bactériens (caractéristiques des CBI), ainsi que des cellules colonisées (79). Cette internalisation est principalement médiée par le fimbria Type 1. Il induira un réarrangement du cytosquelette de la cellule hôte pour permettre la création d'une vacuole (80). Cette capacité constitue un réservoir quiescent de bactéries pathogènes, il les protège du système immunitaire de l'hôte ainsi que de la prise d'antibiotiques. Ceci expliquerait la forte proportion des infections récurrentes qui impliquent souvent la même souche (75-77, 80, 81).

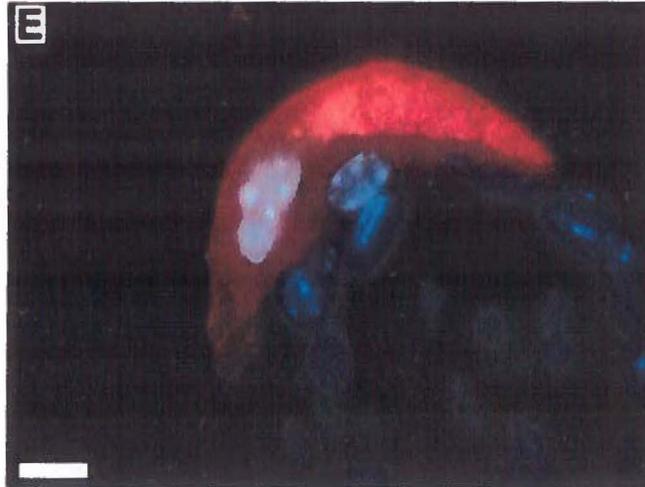


Figure 4 : Observation d'une CBI

Les noyaux des cellules de l'urothélium sont colorés en bleu. Les bactéries sont colorées en rouge. L'échelle représente 10 μ m. Tirée de M A Mulvey *et al.* (2001) (77).

Les bactériuries asymptomatiques (BAS), qui se caractérisent par une bactériurie supérieure à 10^5 UFC/mL, touchent diverses populations comprenant les femmes enceintes, les patients diabétiques et les personnes âgées et/ou avec cathéter urinaire (82, 83). Les souches responsables des BAS sont un cas intéressant d'évolution des bactéries pathogènes vers des bactéries commensales (84). En effet, ces bactéries présentent une diminution de l'expression des fimbriae, qui est compensée par une augmentation de leur croissance (85). Du point de vue génomique, cela se traduit par : un génome plus petit que les APEC et les UPEC provenant des ITU symptomatiques, et des points de mutations/réarrangements/délétions dans les gènes de virulence (84). Par exemple, des points de mutations sont présents dans le gène *papG* (l'adhésine) du fimbria P et dans le gène de transport *focD* du fimbria F1C. Une large délétion peut être aussi observée dans l'opéron *fim* de plus de 4,2 Kb (84, 86, 87). Ce sont des exemples de modifications directes des séquences des facteurs de virulence. Mais ces changements peuvent aussi altérer des protéines de régulation globales, comme la protéine RfaH qui augmente, à l'état native, l'expression des facteurs de virulence (88). Une des explications apportées pour comprendre cette évolution est l'action du système immunitaire de l'hôte qui influencera sur la production de souches moins virulentes, en supprimant les bactéries qui activent le plus les défenses de l'hôte. Cependant, la persistance dans le tractus urinaire ne peut s'accompagner que d'une augmentation de la capacité à croître dans le milieu urinaire (85, 89).

Ceci conduit à la création de souches provoquant des infections urinaires asymptomatiques, ayant comme ancêtres des souches d'ITU symptomatiques (88).

1.4 Classification par groupes phylogénétiques

Comme décrite précédemment, la population d'*Escherichia coli* suit un schéma clonal (voir chapitre « 1.1.1 : Schéma clonal »). Partant de ce principe, la communauté scientifique a commencé à vouloir classer les souches par leurs différences alléliques. Grâce à la collection ECOR, qui s'est révélée être une bonne représentation de la diversité de l'espèce, et des techniques de différenciation allélique, qui seront décrites dans les prochains paragraphes, quatre principaux groupes phylogénétiques (A, B1, B2 et D) ont pu être ainsi distingués (Fig. 5)(14). Cependant, il faut souligner que de nouvelles études, prenant en compte une plus large diversité de souches d'*Escherichia coli* et de *Shigella*, ont révélé des regroupements plus précis et ont apporté des nuances quant aux quatre groupes phylogénétiques. Les souches ne peuvent pas être seulement classées dans ces quatre groupes (considérés comme les groupes ancestraux), mais aussi dans des groupes hybrides AxB1 et ABD, apparus plus récemment (90, 91).

Diverses études ont mis en relation les groupes phylogénétiques avec la pathogénicité. Par exemple, les souches pathogènes, faisant partie de la famille des ExPEC, se retrouvent préférentiellement dans les groupes B2 et D. Alors que, des souches commensales, de type K-12, sont plus représentées dans les groupes A et B1 (14, 41, 92-100). Or cette relation de groupe phylogénétique/pathogénicité se reflète aussi avec la présence de certains facteurs de virulence parmi ces groupes (41, 93-96, 98, 101, 102). Mais certaines exceptions subsistent : plusieurs facteurs de virulence peuvent apparaître dans un groupe clonal précis, qui ne fait pas partie du groupe phylogénétique (et pathotype) normalement associé avec ces facteurs. Ceci est dû à un biais lors de la détermination des groupes phylogénétiques qui ne prend en compte que les gènes conservés. Donc, cela révèle un événement d'intégration d'un élément ADN (contenant les facteurs de virulence) qui s'est produit chez l'ancêtre commun de ce groupe clonal (94). Ces éléments et ces événements d'intégrations, en relation avec les facteurs de virulence, seront revus plus en détail dans le chapitre « 2.1 : Acquisition des facteurs de virulence ». Enfin, d'autres liaisons ont pu être mises en évidence entre les groupes phylogénétiques et les niches écologiques (103) (101), entre les groupes phylogénétiques et le type d'alimentation (102) et entre les groupes phylogénétiques et l'environnement (104).

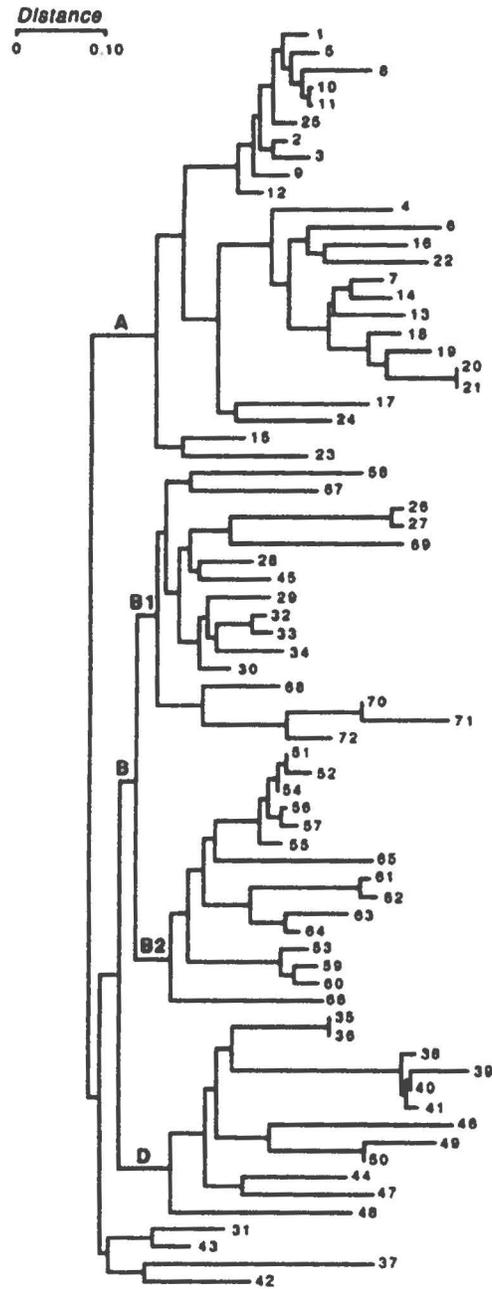


Figure 5 : Distribution phylogénétique de la collection de souches ECOR

Arbre phylogénétique construit avec l'algorithme « neighbor-joining » basé sur le polymorphisme de 38 enzymes.
Tirée de Herzer *et al.* (1990) (14).

1.4.1 La méthode Clermont

Cette méthode est simplement basée sur une PCR multiplex (105). Trois gènes sont amplifiés simultanément dans une même réaction PCR. La présence ou l'absence d'amplification pour l'un, ou plusieurs gènes, détermine le groupe phylogénétique (Fig. 6). Ces trois gènes sont : *chuA*, qui est requis pour le transport de l'hème chez O157:H7 ; *yjaA*, qui joue un rôle dans la réponse au stress acide et est induit par le peroxyde d'hydrogène (106) ; TSPE4.C2, qui correspond à un fragment d'ADN « anonyme » contenant potentiellement un gène. Outre de bons résultats obtenus avec cette technique, elle présente clairement une limitation majeure : elle n'est basée que sur la présence ou non d'une amplification. Donc un simple mésappariement entre les amorces et l'ADN génomique testé peut biaiser le résultat obtenu. C'est notamment ce qui a été observé au laboratoire pour les souches APEC.

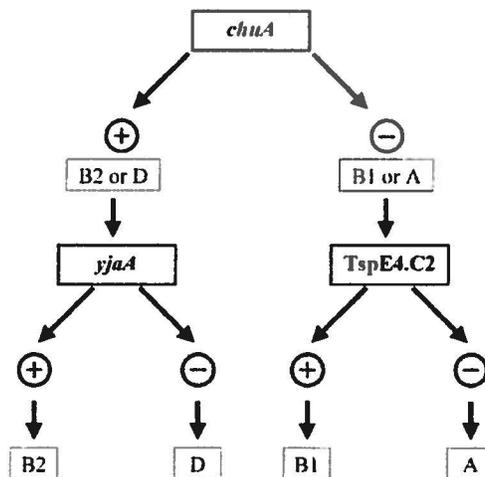


Figure 6 : Arbre de décision pour la méthode Clermont de détermination du groupe phylogénétique
Les signes \oplus représentent la présence d'une amplification pour le gène correspondant. Les signes \ominus leur absence.
Tirée de Clermont *et al.* (2000) (105).

1.4.2 La technique du MLEE

Le MLEE, pour « MultiLocus Enzyme Electrophoresis », est une technique de classification phylogénétique, fondée sur la différence des charges électriques nettes, dans un gel d'électrophorèse, entre les différentes formes (isoenzymes) d'enzymes métaboliques hydrosolubles et conservées (15). Cette différence de migration reflète directement leur composition en acides aminés. C'est une méthode qui est plus fiable que les classements par les

phénotypes (le classement sérologique et les pathotypes). Par contre, elle est sujette aux modifications post-traductionnelles, ce qui génère une source d'erreur (100). De plus, elle ne dissocie pas les différents allèles donnant les mêmes profils de migration et elle est difficilement reproductible entre les multiples laboratoires de recherche (107).

1.4.3 La méthode du MLST

Contrairement à la précédente technique, la méthode du MLST, pour « MultiLocus Sequence Typing », tient compte des mutations neutres de l'ADN. En effet, elle repose sur l'amplification par PCR et le séquençage de sept gènes conservés, dits de ménages. Donc, elle permet : de détecter plus de variations alléliques ; de comparer simplement les résultats entre les laboratoires ; de créer une base de données électronique ; de l'utiliser pour d'autres espèces bactériennes (107). Cependant, comme n'importe quelles techniques, elle possède plusieurs limitations. Une étude a démontré que les gènes utilisés sont insuffisants pour discriminer les souches du sérotype O157:H7 et par conséquent, le MLST ne peut être employé comme outil d'étude épidémiologique (108).

De nouveaux procédés ont vu le jour grâce à la multiplication des séquençages des génomes bactériens. Ils s'appuient notamment sur le polymorphisme d'un seul nucléotide (109-111). Cela reste encore des procédés marginaux, mais qui tendent à se développer avec les plateformes de génotypage SNP à haut débit.

2 Les facteurs de virulence

Comme indiquée précédemment, la différence entre une souche pathogène d'*E. coli*, d'une souche non pathogène est l'acquisition de facteurs de virulence (FV), permettant ainsi à la bactérie pathogène de mieux coloniser l'hôte et surtout d'envahir des organes normalement non propices au développement bactérien. Ces FV peuvent provenir de différentes sources comme des plasmides, des bactériophages, des éléments mobiles... Par exemple, le plasmide ColV contient des gènes considérés comme des facteurs de virulence (112-114).

Pour que soit considéré un gène comme étant un FV, il doit répondre au postulat moléculaire de Koch, édité par Falkow (1988) (115) :

- Le(s) gène(s) doit (doivent) être associé(s) à un trait de virulence.

- La mutation et donc l'inactivation du, ou des gènes, liés à ce trait, doit diminuer la virulence de la bactérie.
- Le remplacement de ce(s) gène(s) muté(s) par un (ou des) gène(s) normal(-aux) doit restaurer le phénotype de la souche sauvage.

2.1 Acquisition des facteurs de virulence

2.1.1 Les transferts horizontaux

Après le séquençage du premier génome bactérien de la souche d'*Escherichia coli* MG1655, la tâche des laboratoires de recherche fut de l'analyser, tant sur le plan des cadres de lecture ouverts et de déterminer leur fonction, que sur les séquences génomiques qui ne codent pas pour des protéines. Ce séquençage a permis de révéler que le chromosome bactérien est une mosaïque d'éléments divers tels que : des gènes de ménages (conservés parmi les genres et espèces bactériennes), des duplications de gènes, des séquences d'insertions (IS) et des éléments transposables provenant possiblement de transferts horizontaux, des gènes de bactériophages... (3, 116)

Or, l'observation du pourcentage de bases G et C, de différentes portions du génome, a mis en évidence que près de 18 % de l'ADN bactérien (\approx 850 Kb) provenait de transferts horizontaux interespèces, plus ou moins éloignés dans le temps. En effet, les gènes conservés ont un taux de G+C aux environs de 50 % et ce taux est spécifique pour chaque espèce. Mais lors de transferts horizontaux, les fragments d'ADN contiennent le taux de G+C du donneur, qui sera différent du taux des gènes natifs, puis au fil du temps et des mutations, ils acquerront le même taux de G+C. Cette particularité a comme origine une disparité dans le taux de mutation pour chaque base des codons. Elle est nommée « pression de mutation directionnelle » (116, 117). C'est un « marqueur » du transfert horizontal de gènes (2, 116-118). Cependant, l'acquisition de gènes est limitée : un large chromosome induit une diminution de la vitesse de croissance, due essentiellement à la vitesse cinétique des ribosomes. Ce qui se traduit par une taille moyenne de chromosome assez stable dans le temps chez *Escherichia coli* (2).

Les bactériophages sont des vecteurs importants de transferts de gènes entre bactéries. Étant donné que la plupart de ces virus ont besoin d'intégrer le chromosome de la bactérie pour se répliquer, ils peuvent transporter (en plus de leurs gènes vitaux) des fragments d'ADN bactériens. Tant et si bien que lorsqu'ils infecteront une autre bactérie, ils apporteront ces

fragments, que la bactérie pourra incorporer dans son génome. Cette intégration est susceptible d'empêcher le cycle lytique du bactériophage. Un cas typique de transfert de gène par bactériophage est l'apparition du pathotype EHEC, avec le sérotype O157:H7. En effet, cette souche proviendrait d'une EPEC O55:H7 qui, après l'infection par un bactériophage, a acquis les gènes codant pour la Shiga toxine (98, 119). Dans ce cas, une plus grande virulence est acquise par une souche déjà pathogène. Mais cela peut aussi se faire avec une bactérie commensale, non pathogène, de sorte qu'en une seule étape elle deviendrait pathogène (116).

De même, les plasmides sont un bon moyen de propagation de nouveaux gènes. Le plasmide ColV, par exemple, est un plasmide transférable. Il est responsable de la résistance au sérum, de la production de toxines, ainsi que d'une amélioration de l'adhérence aux cellules de l'hôte, chez les souches le possédant (112-114).

De plus, une étude, prenant en compte une large diversité de souches d'*Escherichia coli* et de *Shigella*, a pu mettre en évidence une relation entre la reproduction sexuée et la virulence des souches. Les bactéries commensales conservent préférentiellement une structure clonale, avec un faible taux de recombinaisons. Après un événement de transfert horizontal de gènes, la souche mutante produit une population mixte, avec des groupes clonaux et des groupes « sexuels ». En outre, le nombre de recombinaisons et de mutations est plus important chez les souches virulentes. Enfin, d'autres événements de transferts horizontaux permettent un accroissement de la virulence. Ces souches très virulentes présentent un fort taux de mutations/recombinaisons et une population sexuelle exclusivement. Ceci démontre une forte relation entre la virulence, les groupes sexuels et les mutations/recombinaisons (91).

Un autre moyen de « communication » a été découvert récemment. Ce sont des nanotubes composés de couches de la membrane bactérienne. Grâce à ces nanotubes, des bactéries, qui ne sont pas forcément de la même espèce, peuvent transférer des effecteurs cytoplasmiques, ainsi que du matériel génétique. Cela ouvre de nouvelles voies d'études sur les modes d'évolution bactérienne (120).

Ce transfert ne concerne pas seulement des gènes de virulence : des gènes du métabolisme sont aussi transmis. L'opéron *lac* est impliqué dans le catabolisme du lactose. Pourtant cet opéron provient d'un transfert horizontal de gène. Cela en est de même pour le gène *argF* engagé dans la synthèse de l'arginine (2). En définitive, ces transferts horizontaux aident à la

colonisation de nouvelles niches écologiques, en rendant les souches plus compétitives pour l'exploitation de l'environnement et ils sont conservés grâce à la sélection naturelle (23, 117, 118, 121).

2.1.2 *Les ilots de pathogénicité*

Bien souvent, les facteurs de virulence sont regroupés sur ce qu'on appelle des ilots de pathogénicité (IPA). Un exemple, déjà donné, est le LEE des pathotypes EPEC et EHEC. C'est un regroupement de tous les gènes impliqués dans le phénotype des lésions A/E. Les IPA sont des éléments transférables par les bactériophages ou les plasmides (121, 122).

Les ARNt sont fortement reliés avec les IPA : ils possèdent des séquences répétées qui sont reconnues par les bactériophages, afin d'intégrer leur ADN contenant un IPA (121). C'est pour cela que, à un emplacement identique dans le chromosome d'*Escherichia coli*, deux IPA différents peuvent être retrouvés. C'est le cas avec un IPA spécifique aux UPEC et le LEE des EPEC/EHEC. Mais le pathotype n'est pas totalement « fait » par l'IPA : les souches UPEC CFT073 et J96, possède deux IPA différents cependant elles font partie du même pathotype (122). L'acquisition des IPA peut se faire à de multiples temps. Ainsi de nouveaux clusters de souches apparaissent et les facteurs de virulence sont disséminés parmi la population bactérienne (93-95, 99, 122).

Comme mentionné dans un autre chapitre, ce n'est pas le sérotype qui donne la virulence de la souche. Mais il y est fortement associé : en plus des gènes de virulence, les IPA peuvent contenir des gènes qui seront responsables de la modification de l'antigène O ou K (98, 121, 122).

Comme pour l'acquisition, la perte de certains gènes est possiblement bénéfique. Par exemple, *Shigella flexneri* et *Escherichia coli* EIEC ont un large « trou » dans leur génome, appelé « trou noir ». Les gènes métaboliques supprimés réduisaient la pathogénicité des souches, car l'un des produits métaboliques de ces gènes diminuait l'efficacité de facteurs de virulence. Donc cette perte a eu pour conséquence une augmentation de leur virulence (122).

2.1.3 *Les séquences d'insertion (IS)*

Les IS sont des éléments mobiles de l'ADN. Ils se composent essentiellement d'un cadre de lecture codant pour une transposase, flanqué par de petites séquences répétées terminales, qui

seront reconnues par la transposase. Généralement, les IS ont une taille inférieure à 2,5 Kb. Ils sont regroupés dans différentes familles, selon le pourcentage d'identité de séquence de la transposase, la conservation du site catalytique, l'organisation et la conservation des séquences répétées terminales (123, 124). On compte aujourd'hui plus de 1500 IS différentes, classées dans une vingtaine de familles (123).

Les IS sont capables d'engendrer des mutations par des réarrangements de l'ADN comme des délétions, des insertions et des fusions de réplicons. Ces modifications s'effectuent préférentiellement dans des zones riches en bases A et T. Normalement, seulement l'élément IS est transloqué, mais un segment d'ADN, flanqué de deux IS, peut être lui aussi transloqué. Dans ce cas, ils jouent le rôle de transposon composé. Par ce principe, les IS sont souvent associées avec des gènes de résistance aux antibiotiques ou des gènes du métabolisme (2, 124). Ils sont moins responsables de la translocation des IPA, car plus la taille du fragment d'ADN est grande, plus le réarrangement d'ADN, effectué par la transposase, sera compliqué. Même s'ils sont moins utilisés pour ce transport, il n'en est pas moins qu'ils font partie intégrante des IPA (124). En effet, par l'intégration de petites séquences, ils construiront, petit à petit, l'IPA. C'est le cas des gènes de l'aérobactine sur le plasmide ColV (112).

Chez la souche MG1655, plus de 68 % des IS sont groupées avec des transferts horizontaux. Par contre, une certaine disparité apparaît : six sur les sept IS2, qui compte ce génome, sont reliés avec ces transferts alors qu'aucuns IS186 n'y sont joints (116).

Pour conclure, ces différents mécanismes montrent une évolution rapide des populations bactériennes par l'acquisition de grands fragments d'ADN, comprenant, bien souvent, des facteurs de virulence. Ces idées d'évolutions lentes (par mutations spontanées seulement) et d'espèces cloisonnées sont dorénavant désuètes. Les génomes bactériens ont une dynamique forte et une grande plasticité, qui font intervenir plusieurs éléments mobiles et protéines de « transport » de l'ADN.

2.2 Les facteurs de virulence associés aux UPEC

2.2.1 Les systèmes d'acquisition du fer

Le fer est un élément essentiel à la vie sur Terre. Seulement très peu d'organismes sont connus pour ne pas dépendre du fer. Étant le quatrième métal le plus répandu sur la Terre, il est

normal que la sélection naturelle ait favorisé son utilisation. Mais, dans l'environnement, le fer est le plus souvent retrouvé comme des hydroxydes insolubles (125). De plus, l'hôte, comme première barrière au développement des microorganismes, séquestre le fer afin de le rendre moins disponible. Ainsi, les concentrations de fer varient de l'ordre du 10^{-18} à 10^{-24} M (selon l'organe ou le système biologique). Cependant, *Escherichia coli* a besoin d'un minimum de 10^{-6} M pour croître (53, 126-128).

Donc, les bactéries ont développé des systèmes d'acquisition du fer environnant, qui se nomment sidérophores (Fig. 7). Ces petites molécules pourront « arracher » le fer contenu dans les protéines de séquestration de l'hôte (transferrine et lactoferrine principalement), car elles ont une très forte affinité pour le Fe^{3+} . Cela permettra aux bactéries de coloniser des niches biologiques pauvres en fer (127-129).

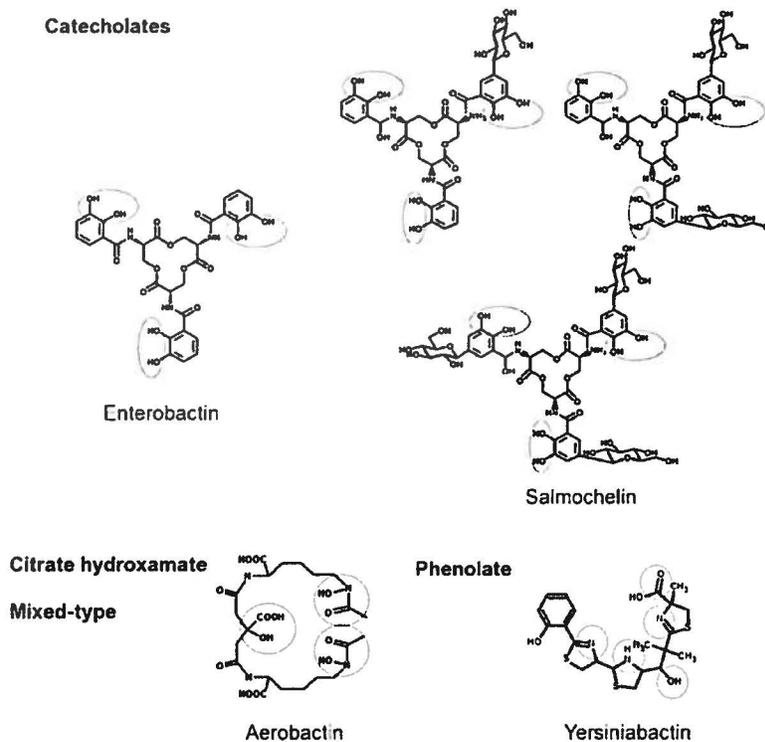


Figure 7 : Structure chimique des sidérophores produits par les *E. coli* pathogènes

Représentation des catecholates, avec l'entérobactine et les trois formes de salmocheline (MGE, DGE, TGE), ainsi que des sidérophores de type citrate hydroxamate (aérobactine) et de type phénolate (yersiniabactine). Les cercles représentent les régions de fixation du fer. Tirée de Garénaux *et al.* (2011).

L'entérobactine

L'entérobactine (enterocheline) fut découverte pour la première fois par O'Brien *et al.* (1970) (130). Elle fait partie de la famille des catecholates et est identifiée dans la quasi-totalité des souches d'*Escherichia coli* (Fig. 7) (131, 132). Cette petite molécule possède un fort pouvoir de fixation du fer, avec un $K_d \approx 10^{-49}$ M. Ce qui est largement supérieur aux protéines de séquestration du fer de l'hôte ($K_d \approx 10^{-20}$ M). Donc, l'entérobactine captera facilement le fer de ces protéines (127). Ce sidérophore est synthétisé par un cluster de six gènes : *entABCDEF* (126). Les sidérophores, étant la contre-mesure des protéines de séquestration, ont fait apparaître un autre type de défense par l'hôte : la lipocaline-2 (ou NGAL) peut emprisonner l'entérobactine et donc limiter l'apport en fer aux bactéries. Cette protéine agit comme un agent bactériostatique (133).

Les salmochelines

La relation de l'hôte humain avec un pathogène est toujours un équilibre entre les différents moyens mis en place par l'hôte pour contrer le pathogène, mais aussi par la bactérie pour empêcher ces mesures d'agir. C'est le cas avec les salmochelines (134). Le NGAL pouvant séquestrer l'entérobactine, un nouveau sidérophore est apparu avec une structure modifiée. Le corps de ce sidérophore est composé d'une molécule d'entérobactine et la protéine IroB rajoute un à trois éléments glucose (Fig. 7) (134-138). Ainsi, la salmocheline ne sera plus séquestrée par le NGAL, à cause d'un trop grand encombrement stérique (127).

Ce sidérophore peut être retrouvé sous des formes plus ou moins glycosylées : soit des entérobactines monoglycosylées (MGE), des entérobactines diglycosylées (DGE) ou des entérobactines triglycosylées (TGE). Ces trois formes sont appelées les salmochelines (137). Plusieurs études ont permis de démontrer que le locus, contenant les gènes *iroN* et *iroBCDE*, est responsable de la synthèse/transport/dégradation des salmochelines et est important pour la virulence de différents pathotypes comme les APEC et les UPEC (4, 56, 126, 139-142).

L'aérobactine

L'aérobactine est une petite molécule de la famille des citrates hydroxamates (Fig. 7). Son nom provient de *Aerobacter aerogenes*, où elle fut pour la première fois identifiée (112). Très tôt, elle a été associée à la virulence des bactéries *Escherichia coli*, notamment des entéroinvasives. En effet, les locus *iuc* (quatre gènes de synthèse) et *iut* (un gène de transport),

codant pour ce sidérophore, se situent sur un plasmide de virulence ColV (11, 114, 143). Mais une copie presque exacte de ces gènes peut se retrouver sur le chromosome, comme chez les NMEC (11, 112, 144, 145). C'est un exemple de transposition ancienne plasmide-génome d'un facteur de virulence.

Cette molécule est composée de la condensation de deux lysines et d'une molécule de citrate (11). À un pH neutre, ou encore dans le sérum, l'aérobactine est plus efficace que l'entérobactine pour capter le fer, et ce en plus petite quantité (11, 112, 126, 146, 147). Elle est fortement associée à la famille des ExPEC pour leur virulence, surtout pour les APEC (sur plasmide) et les UPEC (dans le chromosome) (4, 11, 112, 126, 142, 144-149). Selon les études, de 73 à 98 % des souches APEC expriment l'aérobactine (53). Par contre chez les UPEC, par exemple chez la souche CFT073 (6), la redondance des systèmes d'acquisition du fer fait que l'importance de l'aérobactine est tempérée par les autres systèmes qui prennent le relai en cas de dysfonctionnement (mutation/délétion) de ce sidérophore (150).

La yersiniabactine

Comme son nom l'indique, la yersiniabactine fut pour la première fois isolée chez une souche virulente de *Yersinia enterocolitica*. Cette molécule fait partie de la famille des catéchols (Fig. 7) (151). Elle est capable, facilement, de récupérer le fer séquestré par les transferrines et les lactoferrines grâce à une forte affinité de $K_d \approx 10^{-36}$ M (132, 152). Ce sidérophore est très souvent associé à la virulence des souches qui le possède (132, 151). Tout d'abord, les gènes responsables de sa synthèse se situent sur un IPA. De plus, il aide à prévenir (en captant le fer environnant) la production de radicaux hydroxyles par les leucocytes polynucléaires, les monocytes et les macrophages, afin de détruire le microorganisme (153). Enfin, il a un rôle important dans la colonisation du tractus urinaire par les UPEC en participant à la formation des biofilms (154).

Ces quatre systèmes de séquestration du fer sont régulés par la protéine Fur et la concentration intracellulaire en fer. Lorsque le taux de fer dans la bactérie est suffisant, une majorité des protéines Fur sont couplées avec le fer. Ainsi, elles sont activées et pourront se fixer sur les promoteurs des gènes cibles (ici les gènes de synthèse des sidérophores) et réprimer leur expression. En outre, pour conserver une certaine homéostasie du fer, un petit ARN RyhB

dégradera les ARNm des protéines d'entreposage du fer. Il est lui-même régulé par Fur (11, 126, 129, 146, 151, 152).

2.2.2 *Les flagelles*

Comme décrits dans le chapitre 1.1.2, les flagelles servent à la mobilité des bactéries. Ils se composent de trois structures distinctes : le corps basal, le crochet et le filament flagellaire composé de flagelline (antigène H). Leur taille est d'environ 150 µm de longueur et de 20 nm de largeur. De cinq à dix flagelles peuvent être observés, répartis autour de la bactérie (flagelles péritriches). Plus d'une quarantaine de gènes sont impliqués dans la production des flagelles (155).

La direction empruntée par la bactérie est donnée par chimiotactisme. En effet, différents signaux seront émis aux flagelles par l'intermédiaire de protéines membranaires. Ces protéines détectent, par exemple, des nutriments, des changements de l'environnement, etc. Ceci permet à la bactérie de se diriger vers les sources de nourriture ou encore de fuir un environnement hostile. En dehors de tous stimulus, aucune direction précise ne sera fournie aux microorganismes (52).

La capacité à se mouvoir est considérée comme un facteur de virulence. Comme pour les bactéries entériques, le fait de posséder des flagelles leur permet de pénétrer la barrière que constitue le mucus (52, 156-159). De même, il a été rapporté que pour les NMEC, les flagelles sont nécessaires à la traversée de la barrière hématoencéphalique. Paradoxalement, ce n'est pas la mobilité qui est associée avec les flagelles, mais une fixation par la flagelline sur les cellules de la BHE (160). Cette observation a aussi été faite avec des InPEC et des cellules épithéliales humaines non sécrétrices de mucus (161). Chez les UPEC, sans être essentiel à la colonisation du tractus urinaire, c'est un facteur qui confère un avantage certain pour la migration vers les reins (74, 162, 163).

2.2.3 *Les toxines*

Plusieurs toxines, exprimées par *Escherichia coli*, ont été répertoriées. Mais leur association avec certains pathotypes, leur mode d'action et leurs cibles spécifiques font que seulement les toxines présentes chez les ExPEC, et plus particulièrement chez les UPEC, seront décrites dans ce chapitre.

L'hémolysine

Dès les années 1900, il a été observé que certaines souches d'*Escherichia coli* pouvaient lyser les érythrocytes. De plus, dans les années 1920, la production d'hémolysine par des souches UPEC fut constatée (164). Il existe trois types d'hémolysines, mais seule l' α -hémolysine est largement représentée chez les UPEC, à hauteur de 50 % de souches positives (127, 164-167).

Cette hémolysine est une toxine calcium dépendante qui est relâchée par les bactéries dans l'environnement (11, 168, 169). Elle a été sûrement acquise, par les souches pathogènes, lors d'un événement de transposition de l'opéron *hlyCABD* (164). Ces quatre gènes sont suffisants pour la formation de l'hémolysine : le gène *hlyA* exprime l'hémolysine en tant que telle, *hlyC* rendra actif la protéine et *hlyBD* forme un système de sécrétion de type I, ce système sert au transport de la toxine à l'extérieur de la cellule bactérienne (11, 167, 168, 170, 171). L'hémolysine forme des pores de 2 nm dans la membrane plasmique des cellules de l'hôte. Cela induit la lyse des cellules cibles et de cette manière, un relargage de nutriments se produira et donc favoriser la croissance bactérienne. Par exemple, la lyse des érythrocytes permet la libération de fer, qui est un élément limité dans le système sanguin (11, 167, 169, 171). Le fer particulièrement régule l'expression de l'opéron *hly* : lorsque la concentration intracellulaire augmente, les gènes sont réprimés (11). C'est un facteur de virulence reconnu chez les ExPEC et principalement chez les UPEC. En effet, cette toxine donne un avantage significatif : elle crée une exfoliation des cellules de l'urothélium et ainsi une hémorragie. De même, elle permet une attaque plus importante des tissus rénaux et ainsi un passage plus facile dans le système sanguin, pour provoquer une septicémie. Enfin, une diminution de la phagocytose est observée lors de sa production (11, 127, 164, 167, 172, 173).

Le facteur nécrosant cytotoxique 1

Décrite pour la première fois par Caprioli *et al.* (1983), la libération de cette toxine induisait une altération de la morphologie des cellules eucaryotes (174, 175). La production du facteur nécrosant cytotoxique 1 (CNF1) est dirigée par un seul gène *cnf1*. Chez les UPEC, ce gène se situe sur un IPA et tout de suite après l'opéron de l' α -hémolysine. Même si la régulation de son expression n'a pas été encore précisée, un seul transcrit, contenant les gènes de

l'hémolysine et le gène du CNF1, a été détecté (176, 177). Cette association génétique est très conservée parmi les souches criblées (176, 178).

Le CNF1 est retrouvé dans plus d'un tiers des UPEC et près de 50 % des souches hématuriques le possède. C'est un facteur de virulence non négligeable pour la colonisation et l'invasion des cellules du tractus urinaire (176, 178, 179). Il augmente les dommages commis aux tissus par apoptose des cellules les composant. De plus, le CNF1 aide à la survie des bactéries en diminuant la faculté de phagocytose des leucocytes polymorphonucléaires et la libération d'espèces réactives de l'oxygène. Cela minimisera la réponse inflammatoire aiguë générée par l'hôte (175, 176, 180). Son mode d'action est assez complexe, mais sommairement, le CNF1 entre dans le cytosol de la cellule cible et provoque une activation constitutive des protéines Rho fixant le GTP. Ainsi, plusieurs fonctions de la cellule seront perturbées comme le cytosquelette ou encore l'expression d'immunomodulateurs (18, 176, 179).

Les toxines Sat et Vat

La grande famille des autotransporteurs (système de sécrétion de type V) décrit des protéines qui possèdent déjà toute la machinerie pour pouvoir être exprimées à l'extérieur de la cellule bactérienne.

La première étape est le recrutement des protéines par la machinerie de sécrétion Sec, au niveau de la membrane interne, grâce à un peptide signal en N-terminal des protéines. Ensuite, la partie C-terminale des protéines forme des feuilletts bêta qui s'inséreront dans la membrane externe de la bactérie, afin de constituer un pore. Enfin, le domaine passager (ici les toxines Sat et Vat), relié à la partie C-terminale par un domaine de liaison, traversera le pore et se retrouvera dans l'environnement extracellulaire. Sat et Vat, étant des toxines sécrétées, le domaine de liaison sera clivé pour les libérer de la cellule. Elles font partie de la sous-famille des SPATE, pour sérine protéase autotransporteur des *Enterobacteriaceae*. C'est-à-dire que ces protéines présentent un motif de sérine ayant une activité protéase (170, 181).

La toxine sécrétée Sat a été découverte en premier lieu chez la souche UPEC CFT073. Elle possède une activité cytopathogène sur les cellules de l'urothélium et des reins. Le gène, responsable de son expression, est sur un IPA (181). Par l'intermédiaire d'une attaque du cytosquelette des cellules cibles, Sat induit des modifications et des anomalies de la structure cellulaire. Cela se traduit par la perte des jonctions cellulaires, la production de vacuoles et une

disparition spécifique de la membrane glomérulaire (182, 183). Les dommages infligés aux tissus permettraient une meilleure dissémination des bactéries dans le système sanguin. Cependant, les modèles *in vivo* n'ont pas encore permis de montrer un rôle significatif de la toxine Sat dans la virulence des UPEC (183).

La toxine sécrétée Vat a aussi un effet cytopathogène, par la formation de vacuoles intracellulaires. Tout comme *sat*, le gène *vat* est présent sur un IPA. Il est reconnu que cette toxine est importante chez les APEC, même si son expression a été observée pendant des ITU, aucune preuve n'a été encore rapportée de sa contribution dans ce type d'infection (170, 184).

Des études phylogénétiques ont montré la forte association de ces deux toxines avec les groupes phylogénétiques B2 et D, constatés comme étant les groupes de souches ExPEC (185). Étant donné la conservation remarquable de Sat et Vat parmi les souches UPEC, il est fort probable que leur rôle est important pour la virulence. Des investigations plus poussées doivent être entreprises pour découvrir leur fonction spécifique dans les ITU.

2.2.4 L'adhésine afimbriaire *Iha*

Chez une souche non adhérente, l'expression de cette adhésine afimbriaire lui permet d'exhiber un phénotype d'adhérence sur des cellules de reins et de vessie (186). C'est une protéine membranaire présente parmi 37 à 55 % des souches UPEC (187, 188). L'adhésine *Iha* est aussi impliquée comme récepteur de l'entérobactine. Malgré le fait que différentes études lui donnent un rôle dans la virulence des UPEC, le résultat est peut-être biaisé vu que les mutants, qui n'expriment plus *Iha*, perdent leur capacité d'adhérence et ont aussi une altération de leur capacité d'acquisition du fer (18, 189).

2.3 Les fimbriae chez les UPEC

2.3.1 Caractéristiques générales

De multiples filaments, provenant des bactéries et qui ne correspondaient pas à des flagelles, ont été observés pour la première fois en 1950 (190). Ce n'est que bien plus tard que le rôle de ces appendices fut décrit. Ils sont responsables de l'adhérence des bactéries (191). Après la découverte de plusieurs filaments (pilus « sexuels », flagelles, filaments pour l'adhérence), ayant des propriétés différentes, Ottow (1975) proposa le nom de fimbriae pour tous les

appendices impliqués dans l'adhérence des bactéries (192). Depuis, seulement dans les familles majeures de la voie d'assemblage du « *chaperonne-usher* » et de la voie d'assemblage alternative, plus d'une soixantaine de fimbriae ont été révélés (193).

Étant des organelles complexes, leur production demande une coordination très précise entre le repliement, la sécrétion et l'ordre d'assemblage des protéines qui les composent (194). Cela passe par quatre grandes voies d'assemblage des fimbriae : comme mentionnée précédemment, la voie du « *chaperonne-usher* » est la principale et la mieux caractérisée, la voie alternative est fortement ressemblante à cette première. Il existe aussi la voie de sécrétion générale et enfin la voie de précipitation nucléaire extracellulaire (195).

La plupart des fimbriae des voies du « *chaperonne-usher* » et alterne sont composés : d'une sous-unité mineure, qui fait souvent office de protéine initiatrice de la production du fimbria et de protéine d'« adhérence » ; d'une répétition de la sous-unité majeure, elle forme le corps et la longueur du fimbria ; d'une protéine « *usher* », cette protéine sert de point d'ancrage aux fimbriae dans la membrane externe et elle permet l'exportation des sous-unités à l'extérieur de la bactérie ; d'une chaperonne, elle est responsable du repliement des protéines dans le périplasme et de l'apport des sous-unités vers l'*usher* pour l'assemblage du fimbria (Fig. 8) (26).

Les fimbriae sont des acteurs importants pour l'établissement et le maintien de l'infection. Ils médient l'interaction entre l'hôte et le pathogène, en se fixant sur des récepteurs spécifiques au niveau des cellules de l'hôte (195). Outre leur rôle principal d'adhérence, il est reconnu qu'ils peuvent être impliqués dans l'invasion des cellules et la formation des biofilms (81, 193). Les fimbriae primordiaux chez les UPEC sont les fimbriae Type 1 et les fimbriae P (26, 127, 193-196). Mais plus d'une dizaine de clusters de gènes, codant pour des fimbriae, peuvent être retrouvés chez les bactéries uropathogènes. Par conséquent, cela demande une synchronisation dans l'expression de chacun des fimbriae qui dépendra, principalement, de l'environnement. C'est ce qu'on appelle la variation de phase (127). De plus, le fait d'avoir une population mixte de bactéries, ayant des états de variation de phase différents, permet d'avoir toujours une sous-population avec le meilleur phénotype pour répondre à un changement brusque d'environnement (197).

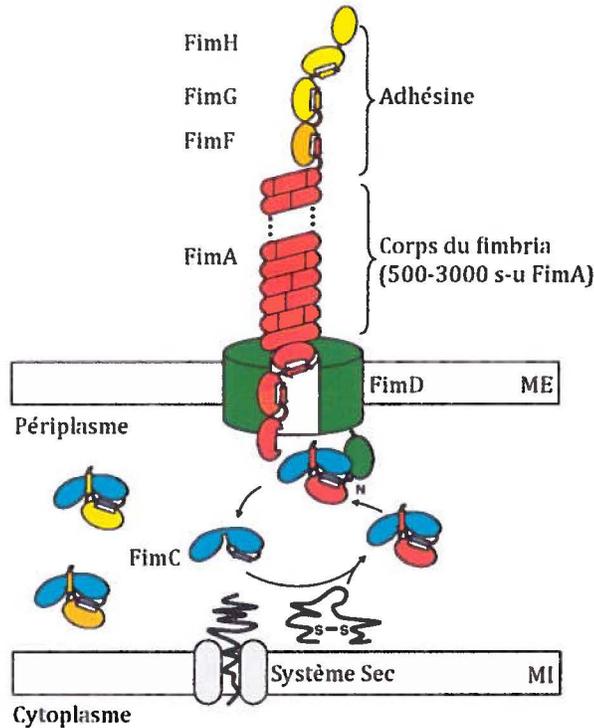


Figure 8 : Modèle de l'assemblage du fimbria Type 1 par la voie du « chaperonne-usher »

La protéine FimH est l'adhésine, elle reconnaît les récepteurs à la surface des cellules de l'hôte. Les protéines FimG et FimF servent comme des connecteurs entre FimH et le corps du fimbria composé d'une répétition de la sous-unité majeure FimA. FimD est la protéine « usher » et enfin FimC est la chaperonne. Ces protéines sont sécrétées au niveau du périplasm grâce au système Sec. Abréviations : s-u pour sous-unité ; ME pour membrane externe ; MI pour membrane interne. Modifiée de Nishiyama *et al.* (2005) (198).

2.3.2 Classification

Comme expliqué dans le chapitre « 1.1.2 : Classification sérologique », les fimbriae peuvent être distingués selon leur groupe sérologique (25). Cela passe par une agglutination des bactéries par différents anticorps monoclonaux, spécifiques à un groupe sérologique. Par contre, ce classement s'applique exclusivement à *E. coli* et non pas aux autres *Enterobacteriaceae*. Par exemple, le fimbria P compte 12 groupes sérologiques : F7 à F16, F40 et F48. Ou encore, des fimbriae représentent seuls un groupe sérologique. Cette technique est trop spécifique à la séquence : plusieurs groupes sont attribués au fimbriae P alors que génétiquement et phénotypiquement trois types sont connus uniquement. De plus, les bactéries expriment plus qu'un type de fimbria à leur surface et en des temps différents, ce qui complique la détermination des groupes sérologiques (26).

En 1966, Duguid *et al.* (1966) proposa une classification des fimbriae, selon leur morphologie et leur fonction, basée sur l'hémagglutination. Il différençia sept types de fimbriae (incluant le pili sexuel F). Ainsi, le type 2 ne provoque pas d'hémagglutination, mais ressemble morphologiquement au type 1 ; le type 3 induit une hémagglutination mannose résistante des érythrocytes traités avec l'acide tannique, etc (199). Malgré une mise à jour effectuée par Ottow (1975), les analyses génétiques ont démontré que cette classification est trop arbitraire et empirique pour être utilisée (192). Cependant, des reliquats subsistent tel que la nomenclature pour les fimbriae codés par le cluster *fim*, nommé encore les fimbriae Type 1 (26).

Avec l'avènement de la génétique et du séquençage, d'autres classifications ont vu le jour. C'est le cas notamment, des classements basés sur les séquences des sous-unités majeures et sur les mécanismes d'assemblage des fimbriae (26). Mais tous ces classements ne décrivent pas toute la diversité de morphologies et de fonctions des fimbriae, qui font partie de la voie d'assemblage du « *chaperonne-usher* » et de la voie alterne. L'étude de Nuccio *et al.* (2007) propose de clarifier cette diversité en utilisant la séquence des protéines *usher*. En effet, les fimbriae peuvent être composés de plusieurs sous-unités majeures. Par ailleurs, elles sont beaucoup plus sujettes aux mutations, étant donné qu'elles sont en contact direct avec le système immunitaire. Donc cela pose un problème pour la classification par cette méthode. En revanche, les fimbriae possèdent seulement une seule protéine *usher* conservée, qui de plus, n'est pas accessible aux défenses de l'hôte. Ce qui en fait un classement plus précis dans la représentation de la relation phylogénétique des fimbriae de ces voies d'assemblage (200).

2.3.3 Les fimbriae Type 1

C'est le fimbria le plus commun chez *E. coli*. Plus de 70 % des souches possède le cluster de gènes *fim* (201, 202). Comme une grande majorité des fimbriae exprimés chez *E. coli*, l'assemblage des fimbriae Type 1 (et des fimbriae P) passe par la voie du « *chaperonne-usher* » (Fig. 8). Donc dans le cluster *fim*, tous les gènes nécessaires à sa production sont présents (Fig. 9). C'est à dire : *fimA*, situé au début de l'opéron *fim*, code pour la sous-unité majeure ; *fimC*, est le gène de la chaperonne ; *fimD*, exprime l'*usher* ; *fimF* et *fimG* sont des protéines qui lieront entre eux FimA et l'adhésine FimH (provenant du gène *fimH*) (196, 203-209). Le gène *fimI* ne possède, à ce jour, aucun rôle défini dans la production du fimbria Type 1. Malgré cela, plusieurs observations ont été effectuées. Par exemple, la protéine est exportée dans le périplasme et la

délétion du gène *fimI* inhibe totalement l'expression des fimbriae Type 1. De plus, chez *Salmonella enterica*, cette protéine joue le rôle de terminateur de l'élongation du fimbria, elle possède la même fonction que PapH (voir le chapitre suivant) et aurait peut-être le même rôle chez *Escherichia coli* (210, 211). Les gènes *fimB* et *fimE* sont impliqués dans la régulation de l'expression du fimbria et ne font pas partie de l'opéron *fim* (voir le chapitre « 2.3.5 : Régulation de leur expression ») (212-214).

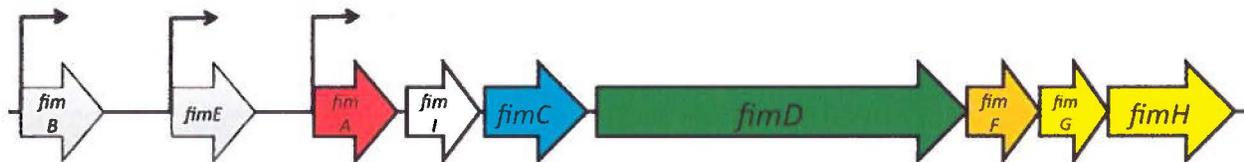


Figure 9 : Schéma du cluster de gènes *fim*

Chaque flèche de couleur représente un gène du cluster *fim*. Les couleurs reprennent celles de la figure 8. *fimBE* sont des gènes de régulation, ils possèdent chacun un promoteur (représenté par une fine flèche noire) ; *fimA* pour la sous-unité majeure ; *fimI* code pour une protéine qui aurait le rôle d'ancrage et de terminateur (voir le texte pour plus d'explications) ; *fimC* pour la chaperonne ; *fimD* pour l'*usher* ; *fimFG* codent pour les protéines de liaison et enfin *fimH* pour l'adhésine. La taille des flèches est dépendante de la taille des gènes et de la région génomique représentée.

Les fimbriae Type 1 ont une longueur pouvant aller jusqu'à 2 μm et un diamètre de 6 à 7 nm (195, 201, 213). Le corps des fimbriae est sous forme d'hélice et est composé seulement d'une répétition de la sous-unité majeure FimA (195, 201, 209). Les protéines FimFGH provoquent l'initialisation de l'extension du fimbria et se retrouvent au sommet. En effet, pour toute la famille fimbriaire de la voie du « chaperonne-usher » (et de la voie alterne), l'assemblage s'effectue par la base du fimbria. Donc, la production commence par l'adhésine FimH, puis vient le tour des protéines FimFG et de la protéine FimA (204, 207, 215). Aucune énergie, telle que l'ATP, n'est nécessaire pour l'assemblage et l'extension du fimbria (216). La protéine FimH est responsable de la spécificité de la fixation du fimbria aux récepteurs mannosylés. Elle reconnaît les structures ayant des groupements D-mannose (196, 201, 205). Il a été démontré que les bactéries, exprimant les fimbriae Type 1, se fixeront sur plusieurs récepteurs, notamment, l'uroplakine Ia présente à la surface de la première couche cellulaire dans la lumière de la vessie, le collagène de type I et IV, la laminine et la fibronectine (202, 217).

Toutes les protéines structurales de ces fimbriae sont sécrétées dans le périplasme par le système Sec. Ce qui demande la présence d'un peptide signal sur chacune des séquences natives

(195, 203, 209, 216). La protéine chaperonne FimC est composée de deux domaines de types immunoglobulines et à un rôle important (avec FimD) dans la fabrication du fimbria (194, 195, 203). En effet, lors de la sécrétion des protéines par le système Sec et par reconnaissance de séquences en C-terminale et N-terminale, la chaperonne va directement prendre en charge ces protéines. Ceci étant afin de pouvoir effectuer le relargage effectif dans le périplasme, par le système Sec (195, 216). La formation des complexes avec la chaperonne empêche la formation d'agrégats des sous-unités dans le périplasme, ainsi que leur dégradation par des protéases. La première fonction pour la protéine FimC est, bien entendu, le repliement des protéines, pour qu'elles adoptent leur configuration tertiaire. En effet, elles sont sécrétées de façon linéaire dans le périplasme (203, 216). Passé toutes les étapes de sécrétions et repliements, les protéines structurelles sont prêtes à être prises en charge par la protéine *usher* FimD. Une forte interaction se produit entre l'*usher* et le complexe chaperonne-protéine. La partie N-terminale de l'*usher* va reconnaître la chaperonne chargée, s'ensuit une modification de sa conformation spatiale pour permettre le transfert de la protéine chargée par la chaperonne à l'*usher* et ainsi faire l'assemblage du fimbria (195, 198, 216, 218). Un fait intéressant c'est quand absence de FimC, la protéine PapD (des fimbriae P) peut la remplacer dans l'assemblage du fimbria Type 1 (203). Les protéines structurelles passent à travers l'*usher*, qui forme un pore de 2 à 3 nm, de façon linéaire : la taille du pore est trop petite pour faire traverser la membrane externe à une hélice de sous-unités majeures d'un diamètre de 6 nm. En somme, l'hélice se forme à l'extérieur de la cellule bactérienne (195, 219). Enfin, la taille du fimbria dépend essentiellement de la concentration en sous-unités présentes dans le périplasme. De même, le nombre de fimbriae exprimés dépend du nombre de protéines FimFGH disponibles pour l'initialisation de la production (195).

Les fimbriae Type 1 sont essentiels dans la virulence chez les UPEC. Le récepteur D-mannose est présent à la surface de la plupart des cellules de la vessie (11, 220, 221). L'adhérence des bactéries aux cellules de l'hôte entraîne une réponse inflammatoire, ainsi qu'une exfoliation des cellules et leur apoptose. C'est l'un des mécanismes primaires du système immunitaire inné (73, 222). De plus, l'invasion des UPEC, dans les cellules de la vessie, est médiée par ces fimbriae. La fixation de FimH provoque un réarrangement du cytosquelette des cellules de l'urothélium. Ainsi, les *E. coli* seront en mesure de former des communautés bactériennes intracellulaires (CBI). Ceci expliquerait la récurrence des ITU chez la femme par

les mêmes clones : la formation des CBI protège les bactéries de l'attaque du système immunitaire et de la prise d'antibiotiques (73, 80, 127, 218).

2.3.4 Les fimbriae P

Le nom de ces fimbriae vient tout simplement du fait qu'ils sont fortement associés aux souches causant des pyélonéphrites. De 70 à 80 % des UPEC responsables des ITU sévères possèdent ces fimbriae, contrairement aux bactéries fécales (19 %), aux bactéries des BAS (24 %) et à celles des cystites (36 %) (11, 223, 224). De plus, les isolats des personnes souffrant d'une infection systémique par *E. coli* sont souvent porteurs du cluster de gènes *pap*, codant pour ces fimbriae. En effet, les infections systémiques sont couramment précédées d'une pyélonéphrite (225).

La morphologie des fimbriae P est très semblable à celle des fimbriae Type 1, avec un diamètre de 5 à 7 nm et une longueur variant de 0,5 à 2 µm (225). Onze gènes sont impliqués dans la formation et l'expression des fimbriae P. Ils proviennent certainement d'un transfert horizontal de gènes, vu qu'ils se situent sur un IPA (226). Étant donné que la voie d'assemblage est la même que pour les fimbriae Type 1, les fonctions pour la plupart des gènes du cluster *pap* sont les mêmes que pour les gènes du cluster *fim*. Par exemple, *papD* expriment la protéine chaperonne ; *papC* l'*usher* ; *papA*, la sous-unité majeure et forme le corps des fimbriae ; les gènes *papEF* servent à la production de deux protéines effectuant le lien entre PapG (l'adhésine) et l'hélice produite par PapA (11, 194, 227-232). Contrairement aux fimbriae Type 1, les fimbriae P possèdent une protéine (PapH) qui médie l'allongement des fimbriae et sert d'ancrage (233).

Les fimbriae P reconnaissent spécifiquement des glycosphingolipides à la surface des cellules de l'hôte. Mais il existe trois variantes de l'adhésine PapG, qui distinguent des récepteurs légèrement différents, particulièrement pour les groupements voisins en position distale ou proximale au disaccharide de galactose (234, 235). La fixation des fimbriae sur ses récepteurs entrainera une libération de céramide de leur partie lipidique. Cette molécule se fixe sur les TLR4, ce qui induira une production de cytokines, de chémokines et un recrutement de neutrophiles au site de l'infection. Cela va inévitablement aboutir à des dommages aux tissus dus à la réponse inflammatoire (226, 236). En raison de la prédominance des récepteurs pour les fimbriae P au niveau des reins (principalement pour le variant II), c'est surtout à cet endroit que

les dommages sont les plus importants (226). De plus, une étude a montré la faible implication des fimbriae P dans la colonisation de la vessie (202). C'est donc un facteur de virulence important pour les souches menant à des pyélonéphrites, voir des septicémies.

2.3.5 Régulation de leur expression

Chez une même souche pathogène, près de 10 clusters de gènes, codant pour des fimbriae, peuvent être retrouvés dans son génome (237). Pour éviter des interférences entre les différentes structures fimbriaires et pour répondre correctement à l'environnement de la bactérie, une régulation fine et croisée entre les fimbriae est nécessaire. Par exemple, deux mécanismes différents de régulation sont à l'œuvre pour les fimbriae Type 1 (réarrangement de séquence ADN) et pour les fimbriae P (régulation épigénétique).

Régulation d'expression des fimbriae Type 1

La variation de phase pour les fimbriae Type 1 est assez radicale dans le sens que pour réprimer la production du fimbria, c'est directement une séquence de 314 pb, appartenant au promoteur de *fimA*, qui se « retourne » et donc inactive le promoteur (Fig. 10) (238, 239). Ce réarrangement d'ADN est sous l'égide de quatre recombinaisons. FimB peut effectuer le changement de la phase ouverte à la phase fermée ou inversement. Malgré la forte homologie avec cette dernière, la recombinaison FimE, quant à elle, fait le changement de la phase ouverte à la phase fermée seulement (11, 212, 238, 240, 241). Ceci est dû à une plus grande affinité pour la séquence de reconnaissance quand le promoteur est en phase ouverte (242). De plus, par la stabilisation de transcrits lorsque le fimbria est en phase ouverte, une plus grande production de protéines FimE est observée. Ces recombinaisons reconnaîtront deux séquences répétées inversées (SRI) de 9 pb. Ces SRI se situent aux extrémités du fragment de 314 pb (243). Les deux autres recombinaisons, indépendantes de FimB et FimE, impliquées dans le réarrangement d'ADN sont : IpuA, elle possède les mêmes attributs que FimB (activité « phase ouverte à fermée » et « phase fermée à ouverte ») et IpbA change la phase de fermée à ouverte uniquement. Mais ces deux recombinaisons ne sont pas présentes chez toutes les souches d'*E. coli* (244, 245).

En plus de la régulation effectuée par les recombinaisons, plusieurs protéines vont interagir avec la variation de phase. C'est le cas notamment de la protéine H-NS qui réprime la production de Type 1, contrairement aux protéines Lrp et IHF qui régulent positivement l'opéron *fim* (246, 247). Également, par des mécanismes non encore élucidés, des facteurs environnementaux

peuvent modifier l'expression des fimbriae : le fait de faire croître une culture bactérienne de façon statique et avec un pH basique, d'avoir une température autour de 37 °C ou encore d'utiliser un milieu minimal plutôt qu'un milieu riche augmentera sa production (247-252). Enfin, les expressions de flagelles et du fimbria Type 1 ne sont pas mutuellement exclusives, malgré leur rôle opposé. L'augmentation de l'expression de Type 1 affecte l'expression des flagelles en la diminuant. Cependant, l'inverse n'est pas observé (253, 254).

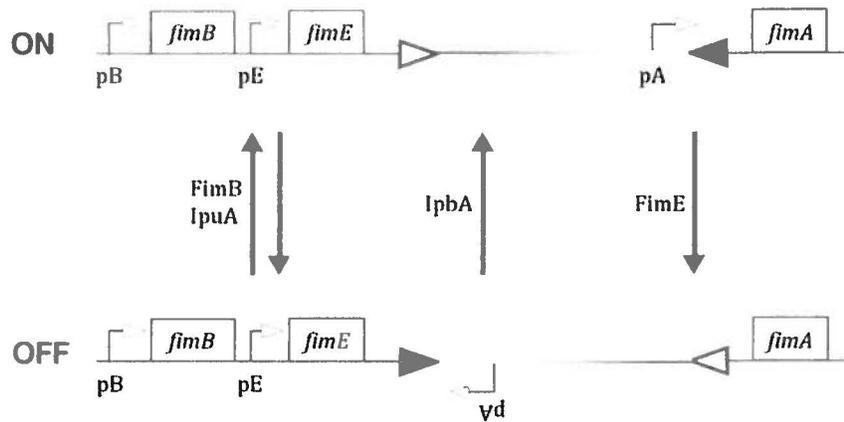


Figure 10 : Schéma de la régulation de la variation de phase du fimbria Type 1 par les recombinases

La position relative des promoteurs (flèches blanches), des gènes (rectangles blancs) et des SRI (triangle blanc et noir) sont représentés. L'inversion du fragment entre les SRI est médiée par plusieurs recombinases : FimB, IpuA, IpbA et FimE. Chacune favorisant un sens particulier de la réaction (sauf pour FimB et IpuA) (flèches noires). La phase « ON » pour la phase ouverte et la phase « OFF » pour la phase fermée. Modifiée et mise à jour de van der Woude *et al.* (2004).

Régulation d'expression des fimbriae P

Contrairement aux fimbriae Type 1, les fimbriae P sont régulés de façon épigénétique (255). C'est une modification de la méthylation de deux sites « GATC » et une différence de fixation de plusieurs protéines qui détermineront la phase ouverte ou fermée du promoteur de l'opéron *pap* (Fig. 11).

La variation de phase du fimbria P repose sur un équilibre entre la méthylation des deux sites « GATC » (GATC I et GATC II) par la protéine Dam, la fixation de la protéine Lrp, qui est elle-même modifiée pour son affinité avec les sites « GATC » par la protéine PapI et l'action des protéines PapB et CAP. Lorsque Lrp est fixée sur le site GATC II, la variation de phase est fermée. Cet état est conservé par la méthylation du site GATC I par Dam (diminution de

l'affinité de Lrp pour le site, en présence de groupements méthyles). Le changement de phase peut s'effectuer lors de la réplication de l'ADN, où il est hémiméthylé. En effet, quand le site GATC I n'est pas méthylé, le complexe Lrp-PapI pourra se fixer sur ce site et ainsi procéder à l'expression de l'opéron *pap*. Par PapI, Lrp a plus d'affinité pour le site GATC I. De même, la méthylation du site GATC II permet la conservation de cet état de phase. Les protéines CAP et PapB augmentent la stabilité de la phase ouverte en se fixant sur l'ADN (243, 255-260).

Par ailleurs, la production des fimbriae P est aussi régulée par : la température, le mode de culture (bouillon versus gélose, statique versus agité), le pH, la source de carbone et la concentration pour certains acides aminés (11, 243, 255, 260).

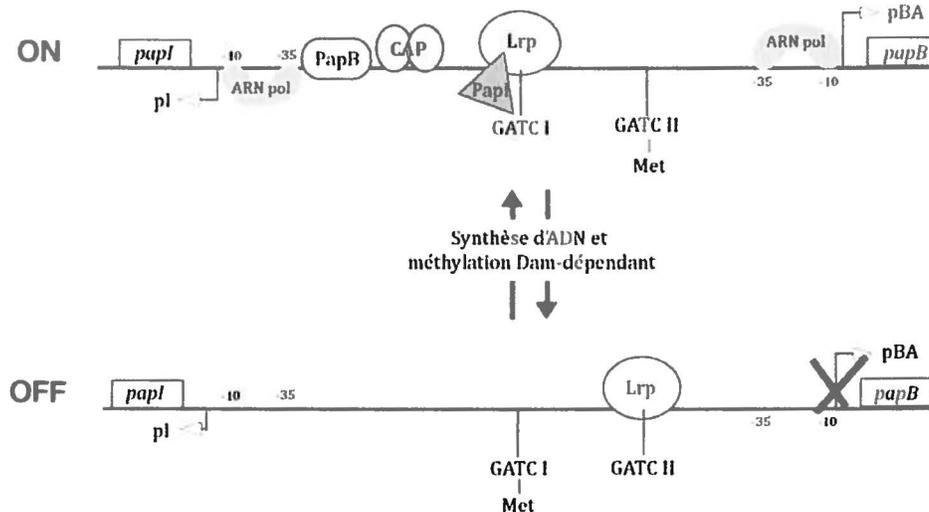


Figure 11 : Schéma simplifié de la régulation de la variation de phase du fimbria P

La position relative des promoteurs (flèches blanches), des sites « GATC » et de leur méthylation (Met), des gènes (carrés blancs) et des protéines impliquées dans la variation de phase (cercles et triangle) est représentée sur cette figure. La phase « ON » est pour la phase ouverte et la phase « OFF » pour la phase fermée. Le changement s'effectue par la différence de méthylation des sites GATC I et GATC II, ainsi que par la fixation de la protéine Lrp et de l'action des protéines PapI, PapB et CAP. Modifiée de van der Woude *et al.* (2004).

La régulation croisée ou « cross-talk »

Il est très important pour la bactérie d'exprimer les adhésines en fonction de la niche colonisée. En effet, dans le cas d'une ITU, c'est plus de quatre types de tissus qui sont rencontrés et donc une multitude de récepteurs différents présents à la surface des cellules. Le fait d'exprimer un type de fimbriae à un certain moment de l'infection plutôt qu'à un autre permet une optimisation de la colonisation. De plus, le fait de ne pas produire à la surface tous les types

de fimbriae évite les interférences physiques entre les structures et permet un emploi optimal de fimbriae distincts. Enfin, les fimbriae étant des structures qui se trouvent à la surface de la bactérie, ils sont reconnus par le système immunitaire de l'hôte. Par conséquent, une diminution de leur production induit une réaction du système immunitaire modérée. Il existe donc une hiérarchisation de l'expression des fimbriae (243, 254, 261). Par exemple, la protéine PapB, du fimbria P, augmentera l'expression du gène *fimE*. Or, la protéine FimE est reconnue pour changer l'état de la variation de phase de la position ouverte à la position fermée. C'est donc une régulation directe de l'expression des fimbriae Type 1 par des protéines impliquées dans la production des fimbriae P (243, 254).

2.3.6 Comparaison des familles fimbriaires de la voie du « chaperonne-usher » et de la voie alternative

La voie alterne d'assemblage est un exemple d'évolution convergente avec la voie d'assemblage classique du « chaperonne-usher » (262). D'un point de vue mécanistique générale, il n'y a pas de grandes différences entre ces deux voies. Les fimbriae, faisant partie de cette famille (CFA/I, CS, Tcf, Cbl...), possèdent le strict minimum pour leur production (200, 263, 264). Soit une protéine chaperonne, une protéine *usher*, une sous-unité majeure et une sous-unité mineure (26). De plus, leur organisation génétique est différente des fimbriae Type 1 et P : l'ordre des gènes est chaperonne/sous-unité majeure/*usher*/sous-unité mineure. Par ailleurs, aucune des séquences de reconnaissance pour la chaperonne et des domaines de liaisons, conservés chez Type 1 et P, ne sont retrouvés chez cette famille fimbriaire. Ils possèdent des domaines et séquences qui leur sont uniques (200, 262). Enfin, la protéine chaperonne est nécessaire à la stabilisation de l'*usher*. Ce qui n'est pas le cas dans la voie classique d'assemblage (26, 195).

2.3.7 Les autres types de fimbriae importants pour les ITU

Outre les fimbriae principaux, Type 1 et P, deux autres types d'adhésines sont importantes chez les UPEC. Tout d'abord, la famille d'adhésines Dr regroupe plusieurs fimbriae et adhésines non fimbriaires (11). Leur récepteur est l'antigène Dr, composant du facteur DAF qui est retrouvé sur plusieurs types de tissus comme les érythrocytes et les cellules du tractus urinaire (11, 265-268). D'ailleurs, c'est jusqu'à 50 % des souches UPEC, provoquant des cystites, qui possèdent le cluster de cinq gènes permettant l'expression de l'hémagglutinine Dr (11, 202). Ces

fimbriae permettent une invasion des cellules épithéliales de la vessie. C'est pourquoi, ils sont fortement impliqués dans la récurrence des ITU (202, 269).

Le deuxième fimbria est F1C. Il médie surtout l'attachement des bactéries aux cellules des tubules rénaux et aux érythrocytes (11, 202, 270). Par la fixation des fimbriae à ces récepteurs, un relargage d'interleukine 8, cytokine de la réponse inflammatoire, est effectué. Par conséquent, il joue un rôle important dans l'inflammation des tissus rénaux et donc dans l'établissement et l'entretien des pyélonéphrites (271). Un fait intéressant à noter est l'observation d'un échange de composants entre les fimbriae F1C et les fimbriae Type 1. En effet, plusieurs études ont montré la création d'hybrides entre ces deux fimbriae, par l'utilisation de sous-unité mineure des deux systèmes fimbriaires (272, 273).

2.3.8 Conclusion sur les fimbriae

Comme décrits tout au long de ce chapitre, les fimbriae ont un rôle essentiel dans la virulence des souches UPEC. Chaque système fimbriaire aura un rôle particulier, en reconnaissant des récepteurs spécifiques et en étant exprimé à des moments différents de l'infection. Ceci demande une régulation fine et croisée entre ces fimbriae. Cette faculté à exprimer de façon séquentielle les fimbriae explique la capacité des bactéries à coloniser des niches aussi différentes. Cela passe notamment par les signaux environnementaux qui sont captés par les bactéries. Étant donné la multitude des systèmes fimbriaires, beaucoup de ces systèmes n'ont pas été encore caractérisés. Découvrir et clarifier leur rôle permettra de dresser un schéma complet de toutes les étapes d'une ITU et des mécanismes en jeux.

3 L'arginine et le régulateur ArgR

La L-arginine fait partie des dix plus abondants acides aminés, en terme de quantité par cellule, chez *Escherichia coli*. Outre le fait qu'il est l'un des vingt acides aminés essentiels qui composent l'ensemble des protéines, il joue un rôle de stockage de l'azote (274), étant l'acide aminé qui possède le plus de groupements amines. Ainsi, il minimise l'impact osmotique de l'azote et permet sa mobilisation efficace (275). Par la voie des succinyltransférases, l'arginine peut être aussi utilisée par les bactéries comme seule source d'azote dans un milieu pauvre (276). Il possède d'autres rôles connexes comme, la modification des phosphatidylglycérols par

l'intermédiaire de l'ornithine (277), la résistance à l'acide (278) ou encore comme composant d'un peptide signal (279).

La synthèse *de novo* de l'arginine se fait à partir du glutamate, lui-même appartenant au cycle de Krebs et du carbamoylphosphate. Différentes enzymes, au nombre de huit, sont impliquées dans cette synthèse. Ces enzymes sont codées par le régulon de l'arginine (280), qui est composé de 26 gènes ayant un rapport avec le métabolisme de l'arginine. La plupart de ces gènes codent pour les enzymes de synthèse, mais seulement une partie est sous forme d'opéron. Ce qui diffère de ce que l'on peut observer pour la plupart des voies métaboliques. De plus, l'arginine régule ces gènes par l'intermédiaire du régulateur transcriptionnel principal *argR* (280-284). La fixation de l'arginine active la protéine ArgR, qui va, à son tour, se fixer sur les séquences ARG (285, 286) présentes au niveau des différents promoteurs des gènes du régulon de l'arginine (287-289) et ainsi réguler négativement, dans la plupart des cas, leur expression.

3.1 Les différentes sources d'arginine

La bactérie *Escherichia coli* a différentes manières d'acquérir l'arginine : elle peut soit l'importer directement via des transporteurs membranaires, soit le synthétiser *de novo*.

3.1.1 Le transport de l'arginine

Il existe, à ce jour, trois transporteurs différents de l'arginine, qui présentent chacun des caractéristiques qui leur sont propres, notamment pour leur affinité avec l'arginine, leur régulation et leur activité. L'un des tout premiers transporteurs découverts et caractérisés fut le système de transport LAO, pour Lysine Arginine et Ornithine (290, 291). Le second fut le système de transport AO, pour Arginine et Ornithine (292-294). Un troisième et dernier transporteur, spécifique à l'arginine, fut caractérisé bien plus tard (295), malgré les différentes preuves de son existence (294, 296).

Bien que ces transporteurs possèdent des caractéristiques physico-chimiques différentes, ils sont basés sur la même structure que les transporteurs ABC (« *ATP-Binding Cassette* ») (297). C'est-à-dire qu'ils sont composés : de deux protéines perméases qui s'insèrent dans la membrane interne ; une protéine qui s'associe en dimère, responsable de l'apport énergétique par l'hydrolyse de l'ATP (298-301) ; et enfin une protéine qui fixe le substrat, soit ici un ou des acides aminés (302, 303) (Fig. 12).

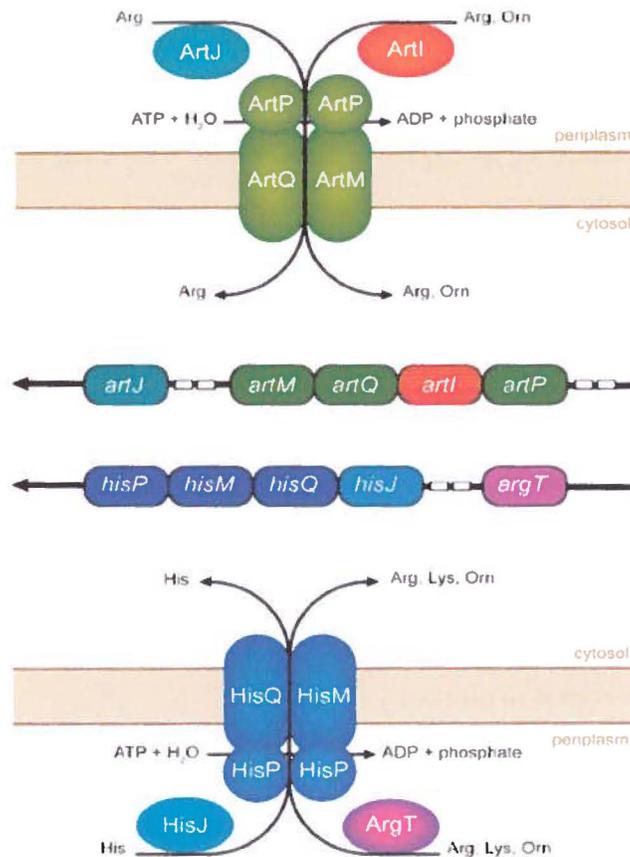


Figure 12 : Schéma des trois transporteurs de l'arginine chez *Escherichia coli*.

Tirée de Caldara *et al.* (2007) avec autorisation. Les flèches indiquent la direction de transcription des gènes. Les boîtes vides indiquent des sites de fixation pour ArgR (302).

Le système de transport LAO

Ce système a une forte affinité avec l'arginine (290, 304), mais il est plus particulièrement utilisé lors d'une carence en azote (276, 305). Il est composé de quatre protéines (303) : HisQ et HisM qui compose la sous-unité membranaire ; HisP, associé en dimère et au complexe HisQM, fournit l'énergie au système par l'hydrolyse de l'ATP (299, 300, 306) et elle est située du côté périplasmique de la membrane (307) ; et enfin la protéine qui est responsable de la spécificité du transporteur pour les acides aminés LAO, ArgT (308). Il est à noter que le complexe HisQMP est aussi employé pour le transport de l'histidine via HisJ (308, 309), ceci peut s'expliquer par le fait que la protéine ArgT et HisJ proviennent d'une duplication d'un seul et même gène ancestral (309, 310). Ce système de transport est codé par les gènes *hisJQMP* et *argT* (309, 311-314) (Fig. 12) et ils font partie du régulon de l'arginine, sauf pour *argT* (287, 302).

Le système de transport LAO est principalement utilisé par la bactérie quand elle est en état de carence d'azote. En effet, les gènes *argT* et *hisJQMP* sont sous le contrôle d'un activateur transcriptionnel NtrC (305), qui est lui-même activé lors de la carence en azote (315). Ce système lui permet donc d'utiliser l'arginine comme seule source d'azote (308, 316).

Le système de transport AO et le système de transport spécifique à l'arginine

Ces systèmes partagent le même complexe protéinique ArtPQM. Ces protéines sont assez similaires aux protéines du système LAO (295) : ArtQ et ArtM formeraient la sous-unité membranaire et la protéine ArtP, qui se situe au niveau du périplasme (295) et associée en dimère, fournira l'énergie au système par l'hydrolyse de l'ATP (298, 299, 301). De même que HisJ et ArgT, ArtJ et ArtI font la spécificité des systèmes par la fixation de leur ligand soit l'arginine et l'ornithine pour ArtI (292-294, 302, 317) et l'arginine pour ArtJ (290, 294-296). Le système AO a une plus faible affinité contrairement au système de transport spécifique pour l'arginine. Par conséquent, à une concentration supérieure à 10^{-6} M, c'est le système avec la plus faible affinité, soit par ArtI, qui est privilégié (284).

Ces protéines ont été récemment définies comme le résultat de la transcription des gènes *artPIQM* et *artJ* (292, 295, 302, 317-319). De même que pour les gènes codant pour le système LAO, ils sont tous sous le contrôle du régulateur transcriptionnel ArgR (Fig. 12) (287, 302). De plus, il a été observé une augmentation de l'expression de *artJ*, lorsque les bactéries sont cultivées en urine. Ceci indiquerait que l'arginine est importante pour la virulence des UPEC et que, surtout, l'arginine est en faible quantité dans le tractus urinaire (320).

3.1.2 *La synthèse de novo de l'arginine et sa dégradation*

Les voies de synthèse et de dégradation de l'arginine ont été depuis longtemps étudiées et caractérisées. Contrairement à la majorité des procaryotes et quelques eucaryotes, *Escherichia coli* ne possède pas l'enzyme de dégradation de l'arginine qui lui permettrait d'accomplir le cycle de l'urée. Néanmoins, l'arginine joue un rôle majeur dans le stockage de l'excès d'azote (275).

La biosynthèse de l'arginine se décompose en deux voies principales (Fig. 13) : la première consiste en la production ou modification du glutamate, la seconde par la transformation de la glutamine en carbamyl-phosphate. Ces deux voies se rejoignent ensuite pour former la citrulline

et enfin l'arginine. Plusieurs des intermédiaires peuvent être directement importés par la bactérie. C'est le cas notamment du glutamate et de l'ornithine (304, 321, 322).

Tous les gènes, qui codent pour les protéines responsables de la production d'arginine à partir de glutamine, font partie du régulon de l'arginine. C'est un système de rétrocontrôle négatif dirigé par ArgR (280, 282-284, 322).

Il existe différentes voies de dégradation de l'arginine, mais une seule est notable par le fait que les gènes codant pour les protéines enzymatiques, font aussi partie du régulon de l'arginine. C'est la voie des AST (Fig. 13). Cependant, l'arginine avec ArgR, va réguler positivement l'expression de ces gènes. En effet, la voie des AST est utilisée pour la production d'azote et de glutamate, lorsque la cellule est en carence d'azote, mais sur un milieu contenant de l'arginine (276, 308, 316, 323).

Finalement, l'arginine, par son catabolisme, est un précurseur d'un grand nombre de composants essentiels à la bactérie. Notamment la putrescine qui est utilisée comme source de carbone et d'azote ou encore le succinate pour une chaîne respiratoire (322).

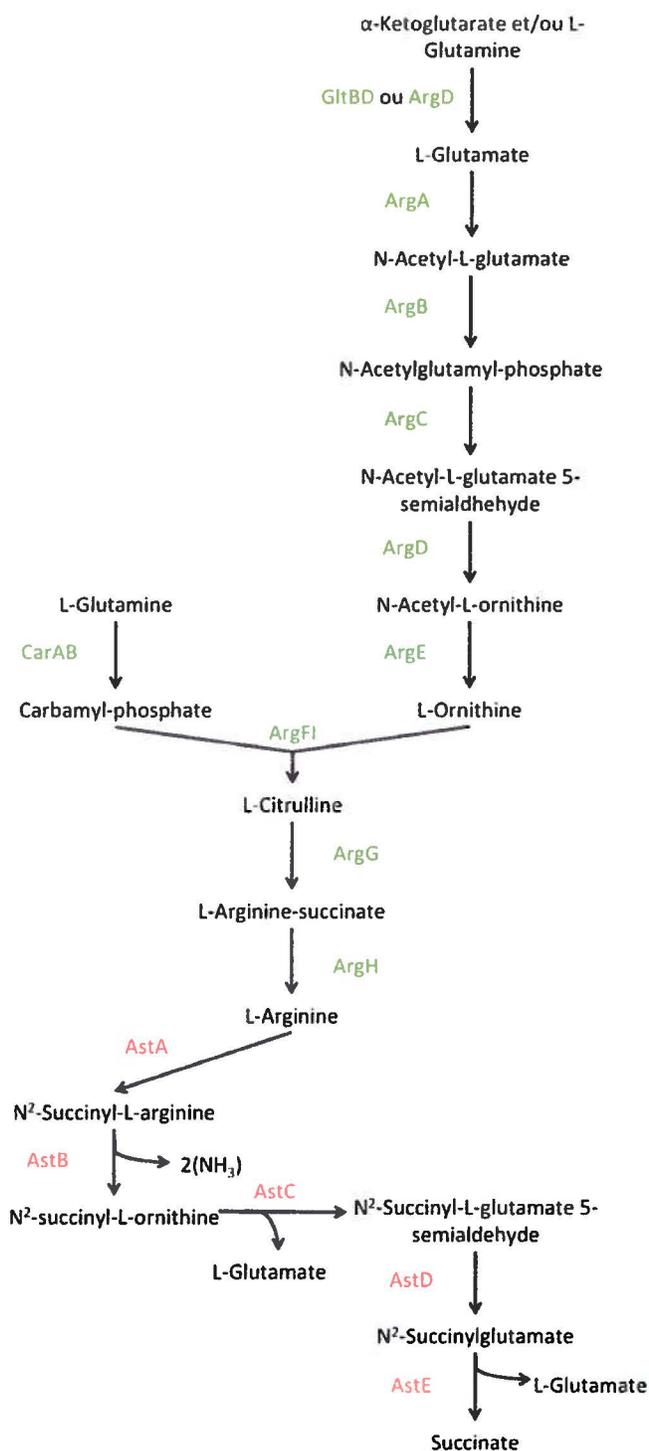


Figure 13 : Biosynthèse de l'arginine et son catabolisme par la voie des AST

Le nom des protéines enzymatiques, effectuant la biosynthèse de l'arginine, est noté en vert ; le nom des protéines enzymatiques, effectuant le catabolisme, est noté en rouge. Chaque molécule produite est en noir.

3.2 La protéine ArgR

La protéine ArgR (281) est le régulateur principal du régulon de l'arginine (280, 324). Les principaux gènes retrouvés dans ce régulon sont impliqués dans le transport et le métabolisme de l'arginine. Cette protéine agit, majoritairement, comme un régulateur négatif. Pour cela, elle se fixera sur des séquences de régulation situées en amont des gènes.

3.2.1 Structure et site de reconnaissance de ArgR

Découverte initialement comme un répresseur de la voie de biosynthèse de l'arginine, la protéine ArgR se présente sous forme hexamérique (325) et est stabilisée par l'arginine, afin de l'activer (286, 326, 327). Ceci est particulier, car la majorité des protéines qui se fixent à l'ADN, chez les procaryotes, sont sous forme de dimère et occasionnellement en trimère ou tétramère. La quantité de cette protéine est assez variable du fait que plusieurs facteurs peuvent induire ou réprimer sa production. Les estimations varient de 40 à 200 hexamères par cellule (322, 328).

La protéine ArgR reconnaît spécifiquement une séquence ADN riche en bases A et T sous forme d'un palindrome imparfait : les séquences ARG (285). Elles se composent d'une séquence de 18 nucléotides, mais n'ont pas une identité parfaite lorsqu'on compare les séquences présentes chez les gènes du régulon. En revanche une séquence consensus peut se détacher : WNTGAATAAWAATNCANT (où W = A ou T et N = n'importe quelle base). Pour que la protéine puisse se fixer, elle a besoin de la coopération de deux séquences ARG, séparées communément de 3 pb (329, 330) et que ces deux séquences sont alignées sur la même face de la double hélice d'ADN (288, 289). Il est possible que ArgR ne reconnaisse qu'une seule séquence ARG, toutefois l'affinité de la protéine pour l'ADN sera grandement diminuée de près de 100 fois (288, 289, 324, 331). C'est le cas notamment du gène *argG* qui possède deux séquences ARG constitutives et une séquence située plus en amont du gène (289, 324, 329). Généralement, une des deux séquences chevauche la région -35 du promoteur, ainsi le complexe ADN-ArgR couvre la séquence de Pribnow et le site d'initiation de la transcription (286, 288, 289, 326, 332). Cela incite à penser que ArgR joue son rôle de régulateur négatif par le principe d'encombrement stérique avec, dans ce cas, l'ARN polymérase (302, 322, 332). Pour les séquences ARG qui sont plus éloignées du début de la transcription, la fixation de ArgR sur l'ADN provoque un changement de conformation du promoteur et donc diminue l'affinité de l'ARN polymérase pour l'ADN, ou encore, peut inhiber l'interaction d'un activateur de la

transcription avec sa séquence spécifique (302, 321). Cependant, dans le cas de la régulation des gènes de la voie des AST, où ArgR induit leur activation, les séquences ARG sont positionnées de telle façon que cela facilite l'interaction de l'ARN polymérase avec l'ADN (323).

3.2.2 Fonctions de *ArgR*

Régulation des gènes métaboliques de l'arginine et de son transport

En premier lieu, c'est en observant la répression catabolique par l'arginine que Maas *et al.* et Gorini *et al.* ont attribué à un facteur R, dénommé plus tard comme la protéine ArgR, le rôle de régulateur métabolique (280, 282-284, 333). En effet, par l'ajout d'arginine dans le milieu, ils observèrent une répression de différentes enzymes impliquées directement dans la synthèse de l'arginine, comme l'ornithine transcarbamylase (325). Plusieurs séquences ARG ont pu être mises en évidence par des expériences d'empreintes ADN sur tous les gènes de biosynthèse (288, 289, 321). De même, les gènes du transport de l'arginine et de son catabolisme par la voie des AST ont été reconnus comme faisant partie du régulon de l'arginine (287, 302, 323, 329).

Il est à noter que ArgR réprime sa propre synthèse, toujours par l'intermédiaire de séquences ARG (285, 287, 325). Mais sa régulation est plus fine. En effet, le gène *argR* est sous le contrôle de deux promoteurs différents. Le promoteur le plus proche du site de traduction possède les séquences ARG, contrairement au deuxième promoteur situé plus en amont (285, 327). Il n'y a, par contre, pas encore d'explication, démontrée expérimentalement, sur le fait que la protéine ArgR, lorsqu'elle inhibe le premier promoteur en se fixant sur l'ADN, n'interfère pas avec la transcription provenant du second promoteur. Néanmoins, une hypothèse peut être émise : l'existence d'une balance entre la quantité d'ARN polymérase et de régulateur ArgR actif. Ceci expliquerait la production d'un minimum de ArgR.

L'exception du répresseur chez *Escherichia coli* B

La souche B d'*Escherichia coli* possède un allèle différent du gène *argR* (noté *argR_B* et ArgRB pour la protéine) (334), qui ne diffère, dans la séquence traduite, que d'un seul et unique acide aminé en position 70. Ce seul changement d'acide aminé donne des propriétés physico-chimiques à la protéine ArgR légèrement différentes comparées à celle de la souche K-12, par exemple. En absence d'arginine, ArgRB a une grande affinité pour l'ADN alors qu'en présence cette affinité diminue, donc ArgRB est active sous ses deux formes : liée et non liée à l'arginine. Ceci est le contraire de ce qui est observé habituellement, où l'arginine active et accroît l'affinité

de la protéine pour l'ADN. Par conséquent, *argR_B* code pour une protéine se comportant comme un super-répresseur (327).

Rôles annexes de ArgR

Outre les régulations des gènes directement en lien avec le métabolisme de l'arginine, ArgR intervient aussi dans la synthèse des protéines par l'intermédiaire du gène *metY*. Ce gène code pour l'un des quatre ARN de transfert qui sont requis pour l'initiation de la synthèse des protéines. Cependant, ArgR réprime cet ARNt seulement en présence du complexe Crp-AMPc, qui est lui-même produit lorsque la bactérie est dans état de stress métabolique par le manque de nutriment. Ce qui suggère une relation entre le métabolisme de la cellule et l'initiation de la synthèse des protéines, mais la relation directe entre métabolisme de l'arginine et l'ARN de transfert n'est pas encore connu (335).

Il a été démontré que ArgR est une protéine accessoire dans la résolution des multimères en monomères, des plasmides ayant pour origine de répllication ColE1, de même que pour l'origine ColK. Comme pour les régions promotrices des gènes du régulon de l'arginine, on retrouve, proche des sites de recombinaison, une séquence ARG. ArgR, en partenariat avec la protéine PepA, provoque une modification de la conformation tridimensionnelle du plasmide, le rendant ainsi plus stable et à même d'être résolu en monomère (331, 336-338). En outre, ces deux protéines accessoires sont aussi requises pour le maintien du prophage P1 (339).

3.3 Aperçu de la protéine ArgR chez les autres espèces bactériennes

Les éléments de régulation, la protéine ArgR et les séquences ARG, ont été étonnamment bien conservés à travers différentes lignées bactériennes et archaebactériennes. On retrouve ces éléments chez *Deinococcus radiodurans*, *Mycobacterium tuberculosis*, *Bacillus subtilis*, ou encore *Chlamydia pneumoniae*. Naturellement, il existe des différences, par exemple sur la séquence consensus qui est reconnue par ArgR (329). De plus, les gènes qui sont sous le contrôle de ArgR peuvent différer, c'est le cas pour un homologue de *artJ* d'*E. coli*, qui, chez *Chlamydia*, n'est pas régulé par ce système (340).

Les différentes études sur le régulon arginine, et plus particulièrement sur le régulateur principal ArgR, ont montré une interrelation entre différents métabolismes et mécanismes cellulaires, ainsi que leur régulation fine. Cela révélerait que le régulateur transcriptionnel ArgR

joue un rôle plus important que la simple régulation des gènes impliqués dans le métabolisme de l'arginine.



HYPOTHÈSES ET OBJECTIFS

Suite à la détermination, grâce à la méthode SCOTS, de séquences exprimées *in vivo* uniques à la souche APEC χ 7122 (et absentes de la souche K-12 χ 289), un nouveau cluster, possédant quatre cadres de lecture ouverts et codant potentiellement pour un fimbria, fut découvert (4). Cependant, après le séquençage du cluster de gènes, une IS2 a été trouvée, interrompant de façon nette (pas de délétion de séquence) la fenêtre de lecture ouverte (ORF) *efoC*. Ceci peut être avancé, car une analyse *in silico* a permis de révéler cette IS (voir Annexe I), et qu'après un criblage par PCR de différentes souches du laboratoire, la souche ECOR72, notamment, possède ce cluster de gènes non interrompu par une IS. Dans ces deux cas, le cluster *efoABCD* est retrouvé entre les gènes *exuT* et *exuR* (Fig. 14). Le nom choisi pour ce cluster *efo* provient de « exu-linked fimbrial operon ».

Le criblage, effectué par M. Lymberopoulos (2004), a révélé que plusieurs souches, de différentes niches écologiques, possèdent potentiellement ces gènes (voir Annexe II) (341). Étant donné que la présence de cette séquence ne semble pas être associée avec un groupe phylogénétique ou avec un pathotype particulier, il a été choisi de caractériser *efoABCD* chez la souche ECOR72.

Vu que *efoC* n'est pas interrompu par une IS, les hypothèses émises sont que : d'une part *efoABCD* code pour un fimbria fonctionnel et que d'autre part, ce fimbria a un rôle dans la virulence de la souche ECOR72 lors de l'établissement d'une ITU.

Les différents objectifs de ce mémoire sont :

- La caractérisation génomique du cluster de gènes *efoABCD*
- L'établissement de la distribution et de la relation phylogénétique de ces gènes avec les autres fimbriae bactériens
- La détermination de la fonctionnalité et de la régulation de ces gènes
- La détermination de leur rôle dans la virulence de la souche ECOR72

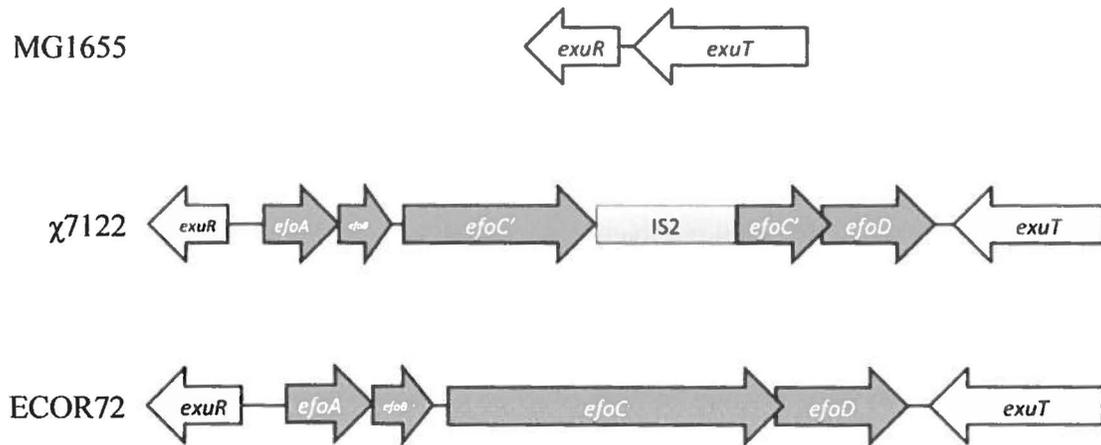


Figure 14 : Organisation génétique de la région entre les gènes *exuR* et *exuT* chez trois souches différentes d'*Escherichia coli*

Les flèches représentent les gènes de la région entre les gènes *exuRT*. Les gènes *exuRT* sont conservés chez *Escherichia coli*, ils sont responsables du métabolisme des hexuronates. La région de la souche MG1655 ne possède pas le cluster de gènes *efoABCD*. Chez la souche χ7122, une IS2 interrompt le gène *efoC*, sans délétion de séquence. La souche ECOR72 présente un cluster « intact », sans IS. La taille des flèches est dépendante de la taille des gènes et de la région génomique représentée.

MATÉRIEL ET MÉTHODES

1 Liste des souches et plasmides utilisés ainsi que les milieux et les conditions de croissance

Souches ou plasmides	Caractéristiques	Sources ou références
Souches		
χ 7122	APEC ; O78:K80:H9 ; <i>gyrA</i> Nal ^R	(342)
BW25142	<i>lacIⁿrrnB3 ΔlacZ4787 hsdR514 DE(araBAD)567 DE(rhaBAD)568 ΔphoBR580 rph-1 galU95 ΔendA9 uidA(ΔMlu)::pir-116 recA1</i> ; Souche de clonage	(343)
DH5α	F ⁻ (Δ80d <i>lacZΔM15</i>) Δ(<i>lacZYA-argF</i>)U169 <i>endA1 recA1 hsdR17(rk⁻ mk⁺) deoR thi-1 supE44 gyrA96 relA1</i> ; Souche de clonage ; Tc ^R	Laboratoires de recherche de Bethesda
ECOR72	UPEC ; Souche de référence de la collection ECOR	(7)
MG1655	F ⁻ lambda- <i>ilvG- rfb-50 rph-1</i>	(3)
MGN-617	<i>thi thr leu tonA lacY glnV supE DΔasdA4 recA::RP4 2-Tc::Mu [pir]</i> ; Km ^R	(344)
ORN172	<i>thr-1 leuB thi-1 Δ(argF-lac)U169 xyl-7 ara-13 mtl-2 gal-6 rpsL tonA2 supE44 Δ(fimBEACDFGH)::kan pilG1</i>	(345)
<i>Serratia liquefaciens</i>	ATCC27592	(346)
Tuner (DE3)	Kit Novagen n°70726-3 pour production de protéines	Novagen
QT48	χ 7122 Δ <i>lacZYA::cat</i> ; Cm ^R	Cette étude
QT885	ORN172 + pBC-SK+	(189)
QT1282	DH5α + pIJ160	Cette étude
QT1306	ORN172 + pIJ164	Cette étude
QT1307	ORN172 + pIJ165	Cette étude
QT1430	ECOR72 Δ <i>efoABCD::kan</i> ; Km ^R	Cette étude
QT1465	ECOR72 Δ <i>efoABCD::FRT</i>	Cette étude
QT1469	ECOR72 Δ <i>lacZYA::FRT</i>	Cette étude
QT1516	ECOR72 Δ <i>fimBEACDFGH</i>	(347)
QT1995	ECOR72 Δ <i>fimBEACDFGH ΔefoABCD::FRT</i>	Cette étude
QT2921	ECOR72 Δ <i>lacZYA::FRT</i> + pRS415	Cette étude
QT2922	ECOR72 Δ <i>lacZYA::FRT</i> + pIJ407	Cette étude
QT2967	ECOR72 Δ <i>lacZYA::FRT ΔargR::kan</i> ; Km ^R	Cette étude
QT2970	ECOR72 Δ <i>lacZYA::FRT ΔargR::kan</i> + pRS415 ; Km ^R	Cette étude
QT2971	ECOR72 Δ <i>lacZYA::FRT ΔargR::kan</i> + pIJ407 ; Km ^R	Cette étude
QT2996	MGN617 + pIJ422	Cette étude

QT3021	QT2967:: <i>Tn7T-Gm::argR</i>	Cette étude
QT3026	QT3021 + pRS415	Cette étude
QT3027	QT3021 + pIJ407	Cette étude
QT3044	ORN172 + pIJ425	Cette étude
QT3045	ORN172 + pIJ426	Cette étude
QT3057	Tuner (DE3) + pIJ427	Cette étude
QT3103	ECOR72 Δ <i>lacZYA::FRT</i> + pSIF194	Cette étude
Plasmides		
pBC-SK+	Plasmide à haut nombre de copies ; Cm ^R	Stratagene
pCP20	Plasmide assistant avec les gènes de la recombinaison Flp ; Thermosensible ; Ap ^R Cm ^R	(348)
pCR-XL-TOPO	Vecteur de clonage à haut nombre de copies ; Km ^R	Invitrogen
pET-28	Vecteur d'expression avec un marqueur His en fusion ; Km ^R	Novagen
pGP-Tn7-Gm	Vecteur d'intégration simple copie dans le chromosome, au site <i>attTn7</i> ; Ap ^R Gm ^R	(349)
pKD3	Plasmide matrice pour l'amplification de la cassette de résistance <i>cat</i> flanquée de sites FRT	(348)
pKD4	Plasmide matrice pour l'amplification de la cassette de résistance <i>kan</i> , flanquée de sites FRT	(348)
pKD46	Plasmide ayant les gènes pour la λ Red recombinaison ; Thermosensible ; Ap ^R	(348)
pRS415	Vecteur de fusion pour un promoteur ; Ap ^R	(350)
pSTNSK-Cm	Plasmide possédant les gènes de la transposase Tns ; Thermosensible ; Km ^R Cm ^R	(349)
pIJ160	pCR-XL-TOPO:: <i>efoABCD</i> ; Km ^R	Cette étude
pIJ164	pBC-SK+:: <i>efoABCD</i> ; Cm ^R	Cette étude
pIJ165	pBC-SK+:: <i>efoABCD</i> ; Cluster inversé par rapport à pIJ164 ; Cm ^R	Cette étude
pIJ407	pRS415:: <i>Pefo-efoA'</i> ; Ap ^R	Cette étude
pIJ422	pGP-Tn7-Gm:: <i>argR</i> ; Ap ^R Gm ^R	Cette étude
pIJ425	pBC-SK+:: <i>efoACD</i> ; Produit par PCR inverse ; Cm ^R	Cette étude
pIJ426	pBC-SK+:: <i>efoABC</i> ; Produit par PCR inverse ; Cm ^R	Cette étude
pIJ427	pET-28:: <i>argR</i> ; Km ^R	Cette étude

Tableau 1 : Liste des souches et des plasmides utilisés dans cette étude

Ap^R résistance à l'ampicilline, Cm^R résistance au chloramphénicol, Gm^R résistance à la gentamicine, Km^R résistance à la kanamycine, Nal^R résistance à l'acide nalidixique, Tc^R résistance à la tétracycline.

De façon routinière, les souches ont été mises en croissance sur un milieu LB, à 37 °C, sur gélose ou bouillon. Lorsque nécessaire, les antibiotiques ont été utilisés à des concentrations

variables : pour l'ampicilline, 100 µg/mL ; pour le chloramphénicol, 30 µg/mL ; pour la gentamicine, 15 µg/mL ; pour la kanamycine, 30 µg/mL.

2 Liste des amorces utilisées

Amorces	Direction	Caractéristiques	Séquences 3' → 5'
CMD26	Sens	Dans le gène <i>glmS</i> pour intégration au site <i>attTn7</i> (utilisée avec CMD1416)	GAT CTT CTA CAC CGT TCC GC
CMD33	Sens	ChuA.1 (méthode Clermont)	GAC GAA CCA ACG GTC AGG AT
CMD34	Anti-sens	ChuA.2 (méthode Clermont)	TGC CGC CAG TAC CAA AGA CA
CMD35	Sens	YjaA.1 (méthode Clermont)	TGA AGT GTC AGG AGA CGC TG
CMD36	Anti-sens	YjaA.2 (méthode Clermont)	ATG GAG AAT GCG TTC CTC AAC
CMD37	Sens	TspE4C2.1 (méthode Clermont)	GAG TAA TGT CGG GGC ATT CA
CMD38	Anti-sens	TspE4C2.2 (méthode Clermont)	CGC GCC AAC AAA GTA TTA CG
CMD46	Sens	Début du promoteur T7 (utilisée avec CMD327)	GTA ATA CGA CTC ACT ATA GGG
CMD177	Sens	Pour produire un mutant <i>lac</i>	CAG CAA ATC GCG CTG TTA GC
CMD178	Anti-sens	Pour produire un mutant <i>lac</i>	GCA TTG ATG ACG CTT GGG CA
CMD266	Sens	Cassette de résistance à la kanamycine (utilisée avec CMD1123)	CAG TCA TAG CCG AAT AGC CT
CMD267	Anti-sens	Cassette de résistance à la kanamycine (utilisée avec CMD707)	CGG TGC CCT GAA TGA ACT GC
CMD327	Anti-sens	Termineur T7 (utilisée avec CMD46)	GCT AGT TAT TGC TCA GCG G
CMD345	Sens	Au début du gène <i>exuT</i>	TAA TCG AAA ATC GGC TGG TC
CMD346	Anti-sens	Dans le gène <i>exuR</i>	GTG GAA CTG CAA ATC CCA CT
CMD368	Sens	Amorce de séquençage du cluster <i>efo</i>	CGG ACG AAG TTA TCA TAC CAC TTT A
CMD369	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	TAA CTA ATT TTT GGG CAG CAT TTA C
CMD370	Sens	Amorce de séquençage du cluster <i>efo</i>	GTA GTA GCA AAA CTG TGC GCT TTG
CMD371	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	TTT TTA GCC AGA GAA CCG TTA GTT A
CMD372	Sens	Amorce de séquençage du cluster <i>efo</i>	TCT TAC CAG TTG ATG ACC AAA ATC T

CMD373	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	AGA TTA TTT CCA GCG TAT CGG TAT T
CMD374	Sens	Amorce de séquençage du cluster <i>efo</i>	TCT CGG TAA ATA TTA TGC AAT GGT T
CMD375	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	GTA GAC TTT TAA TGC CAC GCT GTA G
CMD376	Sens	Amorce de séquençage du cluster <i>efo</i>	CTT TAC CTT TAA CTT TCT GCC ACT C
CMD377	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	CGC TAA GTT GTA CAG TTT CAG AGA T
CMD378	Sens	Amorce de séquençage du cluster <i>efo</i> (aussi utilisée avec CMD1096)	AAC GCC TTA AGC TGT TAT GAA AGT A
CMD379	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	ATA AGG GAA CAA CAC ACT GTC ACT T
CMD380	Sens	Amorce de séquençage du cluster <i>efo</i>	ACT TAT AGC TCT ACG CTG GCT GTC
CMD381	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	ATG CCA AAT CAT ACG TAA TTT CAC T
CMD382	Sens	Amorce de séquençage du cluster <i>efo</i>	TAC AGC GGA GAT GTT ACT TGT AGT G
CMD383	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	TTC ATC TTT GAG CAT GAT TTC ATA A
CMD384	Sens	Amorce de séquençage du cluster <i>efo</i>	CTG GAA GGG GAA AGT TAA ATT ACA T
CMD385	Anti-sens	Amorce de séquençage du cluster <i>efo</i>	CAG TTC TGG CAG TAT TTG ACC TG
CMD386	Sens	Amplification du cluster <i>efo</i> avec site XbaI	GCT CTA GAC CGA TGT GTT TGG TCG TAA C
CMD387	Anti-sens	Amplification du cluster <i>efo</i> avec site XbaI	TCT CTA GAG CGG CAA GTT GTT GAT ACA AAC
CMD707	Sens	Dans le gène <i>exuT</i> (utilisée avec CMD267)	GTC GTC AGG TAA TGT CCC TA
CMD783	Sens	Mutation du cluster <i>efo</i> par la méthode Datsenko (site NotI)	AAT GCG GCC GCA AAA TGC TTT TTT ATA TTG ACT TAC AAT AAA TCT ACA TCT CTA ATA CAC CGT GTA GGC TGG AGC TGC TTC
CMD784	Anti-sens	Mutation du cluster <i>efo</i> par la méthode Datsenko (site NotI)	TAA GCG GCC GCT GGC AAG GAA GGC GTA AAG ATT ACT GAC AAT TTG CCC ATG TAG TAC CCG GAT GGG AAT TAG CCA TGG TCC
CMD1096		Dans le gène <i>efoC</i> (utilisée avec CMD378)	AAC TGT GCC TGA ACC TGA AT
CMD1123		Dans le gène <i>exuR</i> (utilisée avec CMD266)	GAT CGT GGT CAT CAC ACC AG

CMD1416	Anti-sens	Dans un des bords du transposon Tn7 (utilisée avec CMD26)	GCT TTT TCA CAG CAT AAC TGG A
CMD1480	Sens	Mutation de <i>argR</i> par la méthode de Datsenko	ATG CGA AGC TCG GCT AAG CAA GAA GAA CTA GTT AAA GCA TTT AAA GCA TTG TGT AGG CTG GAG CTG CTT C
CMD1481	Anti-sens	Mutation de <i>argR</i> par la méthode de Datsenko	TTA AAG CTC CTG GTC GAA CAG CTC TAA AAT CGC TTC GTA CAG CTC TTT GAT GAA TAT CCT CCT TAG TTC C
CMD1482	Sens	Amplification du gène <i>argR</i> avec son promoteur pour complémentation avec site SmaI	CAG AAT CCC GGG GCC GTG ACG CAG GCA TGT TTC
CMD1483	Anti-sens	Amplification du gène <i>argR</i> avec son promoteur pour complémentation avec site KpnI	TTT CCC GGT ACC CAG AAG CGA CGG GGC AGA GA
CMD1504	Sens	PCR inverse sur pJ164, suppression du gène <i>efoB</i>	GTT CTG CCT AGG GGC ATC TAC CGT GGT GTA GTA AGC AT
CMD1505	Anti-sens	PCR inverse sur pJ164, suppression du gène <i>efoB</i>	CTG CAA CCT AGG ACA TCG CTG CGA CCA GAA GAG ATT
CMD1506	Sens	PCR inverse sur pJ164, suppression du gène <i>efoD</i>	ATT GTC CCT AGG CTT TAC GCC TTC CTT GCC AAC CAT
CMD1507	Anti-sens	PCR inverse sur pJ164, suppression du gène <i>efoD</i>	CAG AAG CCT AGG TCC GCT CTA CCA GAT GCC AAA TCA
CMD1535	Sens	Détermination de l'opéron <i>efoABCD (efoA)</i>	GCG AGT TTG AAC TGG CGG TCC ATT A
CMD1536	Anti-sens	Détermination de l'opéron <i>efoABCD (efoB)</i>	TGC TAG CAT CCA GGC TTT GAA CCA
CMD1537	Sens	Détermination de l'opéron <i>efoABCD (efoB)</i>	GGC ATC TAC CGT GGT GTA GTA AGC AT
CMD1538	Anti-sens	Détermination de l'opéron <i>efoABCD (efoC)</i>	AAT AGCC TCA AAC CCT GCG GGA AT
CMD1539	Sens	Détermination de l'opéron <i>efoABCD (efoC)</i>	ACT GGT AGC TGA AGA TGG CGG ATT
CMD1540	Anti-sens	Détermination de l'opéron <i>efoABCD (efoD)</i>	TAT TTC CTG CCG ACG CTT CAT CCT
CMD1557	Sens	Amplification du promoteur <i>efo</i>	GTA TTG AAT TAA TAT TGT GAA TAT TTA TAA ATT C
CMD1558	Anti-sens	Amplification du promoteur <i>efo</i>	ATA AAT TTG ATC TTA CAA TAT TGT TTC C

CMD1559	Sens	Amplification du promoteur <i>efo</i> avec biotine	biot*GTA TTG AAT TAA TAT TGT GAA TAT TTA TAA ATT C
CMD1560	Anti-sens	Amplification du promoteur <i>efo</i> avec biotine	biot*ATA AAT TTG ATC TTA CAA TAT TGT TTC C
CMD1561	Sens	Amplification du promoteur de <i>argR</i>	CCT ACC GTT CAG CCA AAA CA
CMD1562	Anti-sens	Amplification du promoteur de <i>argR</i>	AAG TCA CCC GAT ATG GTG GT
CMD1567	Sens	Pour la production de ArgR avec site NdeI	GTT CAT GTC ATA TGC GAA GCT CGG CTA AGC AAG A
CMD1568	Anti-sens	Pour la production de ArgR avec site NotI	CAT TAG TAG CGG CCG CTT AAA GCT CCT GGT CGA ACA

Tableau 2 : Liste des amorces utilisées dans cette étude

« biot » signifie qu'une molécule de biotine est ajoutée. Les bases soulignées correspondent à la modification de séquence pour ajouter un site de restriction enzymatique.

3 Manipulations génétiques et de l'ADN

Les enzymes de restriction et la ligase, utilisées dans cette étude, proviennent de la compagnie Fermentas. La Taq polymérase provient de la compagnie Bio Basic Inc. La polymérase à ADN à haute fidélité, Herculase II Fusion, provient de la compagnie Agilent Technologies. Les produits de PCR, les fragments digérés et les extractions plasmidiques ont été extraits et purifiés à l'aide des kits de biologie moléculaire de la compagnie Bio Basic Inc. La préparation des bactéries électrocompétantes et chimiocompétantes a été faite selon (351). Dans la majorité des cas, les plasmides initialement créés ont été introduits dans la souche de clonage BW25142.

4 Construction des mutants pour ECOR72

La majorité des mutants utilisés dans cette étude ont été produits par la technique développée par Datsenko et Wanner (348). Brièvement, en utilisant la λ Red recombinaison (pKD46), le ou les gènes, que l'on veut muter, sont remplacés par une cassette de résistance. La cassette de résistance dépend du plasmide qui est utilisé comme matrice pour la PCR : l'utilisation du plasmide pKD3 donne la cassette de résistance au chloramphénicol (Cm^r) et le plasmide pKD4 à la kanamycine (Km^r). Les amorces utilisées ont un fragment 3' qui ne

s'apparie pas avec le plasmide, mais qui est une séquence d'homologie se trouvant de part et d'autre du ou des gènes qui seront mutés. Après la mutation effectuée et la sélection des clones résistants, la cassette de résistance peut être, ou non, enlevée grâce à la recombinaison FLP (plasmide pCP20). Cette recombinaison reconnaîtra des séquences FRT qui flanquent cette cassette.

— *ECOR72Δlac::FRT (QT1469)*

Cette souche a été produite par la méthode de Datsenko. La cassette de résistance a été amplifiée à partir d'un lysat de la souche QT48 ($\chi7122\Delta lac::Cm^r$), avec les amorces CMD177/178. Les amorces ont été prises à l'extérieur de la zone de mutation afin d'avoir de larges séquences d'homologie avec la souche ECOR72. La cassette a été ensuite enlevée à l'aide du plasmide pCP20. La mutation a été vérifiée par croissance sur une gélose MacConkey.

— *ECOR72ΔlacΔargR::Km^r (QT2967)*

Elle a été fabriquée à partir de la souche QT1469 par la méthode de Datsenko. La cassette de résistance a été amplifiée à partir du plasmide pKD4, avec les amorces CMD1480/1481. La mutation a été vérifiée par PCR avec les amorces CMD1482/1483 et par repiquage pour vérifier les résistances aux antibiotiques.

— *ECOR72ΔlacΔargR::Km^r Tn7::argR-Gm^r (QT3021)*

QT3021 a été créée à partir de la souche QT2967 par la méthode d'intégration simple copie dans le chromosome, au site *attTn7*, développée par notre laboratoire (349). Brièvement, cette méthode consiste à introduire au site unique *attTn7* le ou les gènes que l'on veut rajouter. Elle se base sur une conjugaison entre une souche donneuse (MGN617) auxotrophe pour l'acide méso-diaminopimélique (DAP), transformée préalablement avec le plasmide pGP-Tn7-Gm (transposon Tn7) qui contient le fragment d'ADN qui sera introduit dans le chromosome, et la souche receveuse qui possède déjà le plasmide codant pour les transposases (pSTNSK-Cm). Après la conjugaison, les transposases intégreront le fragment contenu dans le transposon Tn7 qui a été fourni par la souche donneuse. La dissociation entre la souche donneuse et la souche receveuse se fait par un isolement sur une gélose ne contenant pas de DAP, donc seule la souche receveuse peut croître sur cette gélose. Le plasmide pSTNSK-Cm est éliminé par des passages à 42 °C des clones.

La première étape est l'amplification du gène *argR* de ECOR72 à partir des amorces CMD1482/1483. Cet amplicon a été digéré SmaI et KpnI et inséré dans les mêmes sites de restriction dans le plasmide pGP-Tn7-Gm. Ceci a produit le plasmide pIJ422 qui a été introduit par électroporation dans la souche MG1655, ce qui a donné la souche QT2996. La souche QT2967 a été transformée avec le plasmide pSTNSK-Cm, puis la conjugaison a eu lieu afin de donner la souche QT3021. La vérification de l'intégration du gène *argR* a été faite par deux PCR avec les amorces CMD26/1416 et CMD1482/1483.

— *ECOR72Δefo::FRT (QT1465)*

Cette souche a été produite par la méthode de Datsenko. La cassette de résistance a été amplifiée à partir du plasmide pKD4, avec les amorces CMD783/784. Elle a été ensuite enlevée à l'aide du plasmide pCP20. La vérification de la mutation a été faite par PCR avec les amorces CMD386/387

— *ECOR72ΔfimΔefo::FRT (QT1995)*

Cette souche a été produite par la méthode de Datsenko, avec comme base génomique la souche QT1516 (347). Cependant, la cassette de résistance n'a pas été amplifiée à partir d'un des plasmides matrices pour cette technique, mais à partir de l'ADN génomique de la souche QT1430, avec les amorces CMD345/346. QT1430 est une souche intermédiaire dans la production de la souche QT1465 : la cassette de résistance à la kanamycine n'a pas été enlevée par l'utilisation du plasmide pCP20. Donc en prenant des amorces dans les gènes *exuT* et *exuR*, les séquences d'homologie avec la zone à muter sont assez grandes pour une reconnaissance par la λ RED recombinase. La vérification de la mutation a été exécutée après l'insertion de l'allèle *efoABCD::Km^r*, avec les amorces CMD266/1123 et CMD267/707.

5 Clonage et séquençage des gènes du cluster *efo*

Le cluster de gènes *efo* a été récupéré par PCR, avec la polymérase à haute fidélité Herculase II Fusion, à partir de l'ADN génomique de ECOR72 (obtenu par lyse alcaline). Les amorces utilisées sont CMD386/387. L'amplicon a été digéré par XbaI et inséré dans le plasmide pCR-XL-TOPO, aux mêmes sites de restriction. Le plasmide ainsi créé (pIJ160) a été introduit par électroporation dans une souche d'*Escherichia coli* DH5 α (QT1282). Après une culture de

nuit, à 37 °C en LB liquide et avec de la kanamycine, le plasmide a été extrait et purifié. Le plasmide purifié et les amorces de séquençage CMD368 à CMD385 ont été envoyés au Service Séquençage du Centre Innovation Génome Québec de l'Université McGill.

6 Détermination de la formation d'un opéron par les gènes *efoABCD*

L'extraction d'ARNm a été faite sur 1 mL d'une culture bactérienne, incubée pendant la nuit à 37 °C agitée et dans 5 mL de LB, grâce aux kits d'extraction d'ARN de la compagnie Qiagen (RNAprotect Bacteria Reagent ; RNeasy Plus Mini Kit ; RNase-Free DNase Set). La production de l'ADN complémentaire a été effectuée avec le kit de Bio-Rad : iScript cDNA Synthesis Kit. Les protocoles qui ont été utilisés sont ceux des compagnies. Six couples d'amorces (de I à VI) ont été utilisés. Le couple I (CMD1535/1536) amplifie une région de 346 pb entre les gènes *efoA* et *efoB*. Le couple II (CMD1537/1538) amplifie une région de 231 pb entre les gènes *efoB* et *efoC*. Le couple III (CMD1539/1540) amplifie une région de 408 pb entre les gènes *efoC* et *efoD*. Le couple IV (CMD1535/1538) amplifie une région de 739 pb entre les gènes *efoA* et *efoC*. Le couple V (CMD1535/1540) amplifie une région de 3481 pb entre les gènes *efoA* et *efoD*. Le couple VI (CMD1537/1540) amplifie une région de 2973 pb entre les gènes *efoB* et *efoD*. Les PCR sur l'ADN complémentaire et l'ADN génomique de ECOR72 ont été exécutées avec les conditions suivantes : pour les couples d'amorces I, II, III et IV le nombre de cycles était de 35, avec une température d'hybridation de 55 °C et un temps d'élongation de 1 min. Pour les couples d'amorces V et VI, le nombre de cycles était le même, avec une température d'hybridation de 58 °C et un temps d'élongation de 4 min.

7 Préparation de l'antisérum spécifique à EfoB

Le peptide CTKIWSNDVTKDVK a été synthétisé et purifié. Étant une séquence interne à la protéine EfoB, la partie C-terminale et N-terminale ont été acétylée et amidatée respectivement. Ce peptide a été couplé avec un haptène, la protéine KLH (pour « Keyhole Limpet Hemocyanin »). Ce peptide a servi à immuniser deux lapins blancs de Nouvelle-Zélande. Le service de design et de la synthèse du peptide, ainsi que de la production du sérum polyclonal,

a été fourni par la société New England Peptide, Inc. Le sérum a été absorbé contre un lysat de la souche ORN172.

8 Visualisation par MET et immunobuvardage de type Western

8.1 Visualisation par microscopie électronique à transmission (MET)

Pour effectuer la visualisation par microscopie électronique à transmission du fimbria Efo chez la souche afimbriaire ORN172, plusieurs plasmides ont été créés. L'opéron *efoABCD*, avec son propre promoteur, a été amplifié par PCR avec les amorces CMD386/387. L'amplicon et le plasmide pBC SK+ ont été digérés avec l'enzyme de restriction XbaI, puis les deux ont été ligués ensemble pour former les plasmides pIJ164 et pIJ165. La souche ORN172 a été transformée par électroporation avec ces plasmides et avec le plasmide pBC SK+ vide, ce qui a produit les souches QT885, QT1306 et QT1307. Les plasmides, ne contenant plus, soit le gène *efoB* (pIJ425), soit le gène *efoD* (pIJ426), ont été construits à partir d'amplicons (obtenus par des PCR inverses à partir des amorces CMD1504/1505 pour *efoB*- et CMD1506/1507 pour *efoD*-) ligués sur eux-mêmes, après une digestion par l'enzyme AvrII.

Les bactéries ont été centrifugées, à 120000 x g pendant 5 min, sur des grilles de nickel. Puis une coloration négative, par l'acétate d'uranyle, a été appliquée. Les grilles ont été séchées et observées au microscope électronique à transmission Philips EM300.

8.2 Extraction fimbriaire

Après une préculture de nuit dans 5 mL de LB, 1 mL de cette culture a été transféré dans 100 mL de LB, pour une croissance agitée de nuit à 37 °C. Après cette croissance, les cultures ont été centrifugées à 2057 x g pendant 15 min. Le culot a été ensuite repris dans une solution de 75 mM NaCl avec 0,5 mM Tris-HCl à pH 7, pour obtenir une concentration finale de 500 mg/mL de culture bactérienne. Il a été chauffé pendant 60 min à 60 °C et puis vortexé pendant 2 min. Une centrifugation a été effectuée à 3000 x g à 4 °C, pendant 15 min. Le surnageant a été récolté dans un autre tube, auquel il a été ajouté de l'acide trichoroacétique à une concentration finale de 1/10 du volume total, ceci afin d'effectuer la précipitation des protéines. Il a été incubé pendant 20 min sur glace et centrifugé à 20000 x g à 4 °C, pendant

15 min. Le culot a été lavé deux fois, avec 1 mL d'une solution de 0,05 M de Tris-EDTA à pH 12 et a été resuspendu dans 200 µL de 0,05 M de Tris-EDTA à pH 8,5.

8.3 Immunobuvardage de type Western

Cent µL de solution de charge bleue 3x (0,24 M Tris-HCl pH 6,8, 6 % SDS, 30 % glycérol v/v, 16 % β-mercaptoéthanol v/v et 0,06 % bleu de bromophénol p/v) ont été ajoutés aux échantillons. 25 µL de HCl 1 M ont été rajoutés et les échantillons ont été chauffés dans un bain-marie à ébullition, pendant 5 min. Après refroidissement, 25 µL de NaOH 1 M ont été ajoutés. Les extractions fimbriaires ont été séparées par électrophorèse en gel de polyacrylamide, contenant du laurylsulfate de sodium (SDS) selon (351). Deux gels sont formés, le premier dit de concentration à 5 % et un second dit de séparation à 15 %. La migration a été effectuée pendant 15 min à 100 V et pendant 45 min à 200 V. Les gels ont été, soit colorés avec du bleu brillant de Coomassie, soit transférés sur membrane de nitrocellulose (Bio-Rad) pendant 60 min à 100 V, en utilisant une cellule électrophorétique Mini Trans-Blot (Bio-Rad). Les membranes ont été bloquées avec une solution de blocage (25 mM Tris, 0,9 % NaCl p/v, 0,05 % Tween 20 (Pierce) v/v, 3 % de lait écrémé p/v, pH 7,2). L'incubation avec l'anticorps primaire (1:5000) et l'anticorps secondaire (1:25000) a été effectuée pendant 60 min chacun, à température pièce. La détection par substrat chimiluminescent a été pratiquée avec le kit de détection Supersignal West Pico (Pierce).

9 Présence de *efoC* parmi les souches d'*Escherichia coli*

L'ADN génomique, des souches testées, a été extrait par lyse alcaline. Un fragment spécifique à *efoC* et incluant ou non une IS, a été amplifié par PCR, avec les amorces CMD378/1096. Ce gène est le plus conservé chez les fimbriae de la voie « *chaperonne-usher* » (26, 352). Le protocole utilisé est celui du fabricant de la Taq polymérase. 30 cycles ont été effectués, avec une température d'hybridation de 56 °C et un temps d'élongation de 2 min et 30 sec. La spécificité des amorces est vérifiée par l'utilisation d'une souche négative (MG1655) et d'une souche positive (ECOR72) pour le fimbria Efo. Les fragments obtenus font 1235 pb ou 2571 pb pour les souches positives, sans ou avec une IS. La liste des souches criblées est présentée en annexe (Annexe III).

10 **Analyses *in silico***

10.1 **Analyse par recherche de similitudes**

Les outils utilisés sont disponibles en ligne sur le serveur de NCBI (www.ncbi.nlm.nih.gov). Le Blastn, ainsi que la recherche par similitude des domaines conservés (353), a été utilisé avec les valeurs par défaut.

10.2 **Production des arbres phylogénétiques**

Tous les arbres phylogénétiques ont été produits grâce au logiciel MEGA5 (354). L'algorithme d'alignement des séquences est ClustalW (355, 356). La méthode statistique est le « Neighbor Joining ». Le nombre de réplifications pour le bootstrap est de 1000, sauf pour la figure 23 où le nombre est réduit à 100, afin de diminuer le temps de calcul. Le modèle de substitution pour les acides nucléiques est « Jukes-Cantor » et pour les acides aminés c'est le modèle « Jones-Taylor-Thornton ».

10.3 **Analyse du promoteur de *efo***

L'analyse du promoteur *efo* a été effectuée à partir des 200 premières paires de bases en amont du gène *efoA*, grâce au logiciel en ligne BROM de SoftBerry (<http://linux1.softberry.com/berry.phtml>).

11 **Expériences de β -galactosidase**

Les tests de β -galactosidase ont été effectués en accord avec le protocole de J.H. Miller (1992) (357). Après une préculture de nuit dans 5 mL de milieu LB, 10 μ L de la culture bactérienne ont été étalés sur une gélose, le type de milieu est dépendant de l'expérience. Pour le milieu liquide, c'est une dilution au 1/100, dans 5 mL finaux, qui a été effectuée. La culture en microaérobie a été obtenue grâce aux sachets de production de CO₂ de la compagnie « Becton, Dickinson and Company » (GasPak EZ Campy Container System, BD). Après une croissance de nuit, les bactéries ont été resuspendues dans 1 mL de tampon Z (60 mM NaHPO₄•7H₂O, 40 mM NaH₂PO₄•H₂O, 10 mM KCl, 1 mM MgSO₄•7H₂O, 50 mM β -mercaptoéthanol, pH 7) et la DO_{600 nm} a été ensuite ajustée à 0,25, dans un volume final de 1 mL. Pour les cultures en liquide,

les absorbances ont été directement ajustées avec du tampon Z. Il a été ajouté 2 μ L de SDS à 1 % et 40 μ L de chloroforme. Pour lyser les cellules, un vortex de 15 sec a été effectué. Une incubation de 5 min à 28 °C, avant l'ajout de 200 μ L d'ONPG à 4 mg/mL pour démarrer la réaction enzymatique, a été effectuée. Le temps d'incubation, après le début de la réaction, est variable selon les conditions et les souches testées (<5 min pour les souches possédant le plasmide pIJ407 et 30 min pour celle possédant le plasmide vide pRS415). La réaction a été stoppée avec l'ajout de 500 μ L de Na₂CO₃ à 1 M. Une centrifugation à vitesse maximum, pendant 2 min, a été exécutée pour culoter les débris cellulaires. L'absorbance a été mesurée à une longueur d'onde de 420 nm (357). L'activité de la β -galactosidase, exprimée en unité Miller, a été calculée selon la formule : $\text{Activité} = (1000 * \text{DO}_{420 \text{ nm}}) / (t * v * \text{DO}_{600 \text{ nm}})$. Où t correspond au temps d'incubation de la réaction avec l'ONPG ; v , le volume de la culture utilisée (dans cette expérience ce volume est de 1 mL) ; $\text{DO}_{600 \text{ nm}}$ est l'absorbance, prise à la longueur d'onde de 600 nm, de la culture bactérienne (dans cette expérience elle est ajustée à 0,25) et $\text{DO}_{420 \text{ nm}}$ est l'absorbance du milieu réactionnel, prise à la longueur d'onde de 420 nm.

12 Expériences de retard sur gels de polyacrylamide

12.1 Expression et purification d'une protéine recombinante ArgR

Pour construire le plasmide pET-28::*argR* (pIJ427), le gène *argR* a été amplifié à partir de l'ADN génomique de ECOR72 par PCR. Les amorces utilisées sont CMD1567/1568. L'amplicon a été digéré avec NdeI et NotI et inséré dans les mêmes sites de restriction de pET-28. Le plasmide ainsi produit a été introduit par électroporation dans la souche d'*Escherichia coli* Tuner (DE3), pour exprimer la protéine (QT3057). La construction a été vérifiée en utilisant les amorces CMD46/327.

La protéine ArgR a été exprimée avec un marquage composé de six histidines en N-terminal de la protéine. Cinquante mL de culture de la souche QT3057 ont été amenés jusqu'à une $\text{DO}_{600 \text{ nm}}$ de 0,6, en LB-Lennox avec kanamycine à 37 °C et supplémentés avec de l'IPTG (isopropyl β -D-1-thiogalactopyranoside) à une concentration finale de 1 mM. Après deux heures de croissance, les cellules ont été récoltées et resuspendues dans 1 mL de tampon de lyse (50 mM Tris pH 8, 150 mM NaCl, 20 mM imidazole, 1 mM EDTA, 0,1 % Triton X-100, 1 mM PMSF, 0,25 mg/mL lysozyme) et incubées toute la nuit à -80 °C. Après décongélation, les lysats

ont été complétés avec 10 µg/mL de DNaseI et 20 mM de MgSO₄. Ils ont été incubés pendant 1 heure, à température pièce, sous agitation et ensuite centrifugés à 14000 x g à 4 °C, pendant 30 min. Le surnageant a été passé à travers une colonne d'acide nitrilotriacétique de nickel (Ni-NTA), de la compagnie Qiagen. La colonne a été lavée six fois avec du tampon de lavage (50 mM NaH₂PO₄, 300 mM NaCl, pH 8) qui contient une concentration croissante d'imidazole (deux fois avec 50 mM, deux fois avec 75 mM et deux fois avec 150 mM d'imidazole). La protéine recombinante ArgR a été finalement éluée avec un tampon d'éluée (50 mM NaH₂PO₄, 300 mM NaCl, 250 mM imidazole, pH 8) et analysée par SDS-PAGE (Protocole du kit HisProbe-HRP de la compagnie Thermo Scientifique). L'imidazole a été éliminé par dialyse à 4 °C et la concentration en protéine a été déterminée ensuite en utilisant le kit de réactifs pour le dosage des protéines par Coomassie (Coomassie Plus Protein Assay Reagent Kit (Pierce)).

12.2 Expérience de retard sur gel

Les fragments d'ADN des promoteurs *efo* et *argR* ont été obtenus par PCR, avec les amorces CMD1557/1558 et CMD1561/1562 respectivement. De même, les fragments biotinylés du promoteur *efo* ont été obtenus avec les amorces CMD1559/1560.

Les retards sur gels ont été effectués à partir du protocole de Schaumburg *et al.* (2006), avec quelques modifications (340). Un gel non dénaturant à 5 % d'acrylamide, 0,5 X de TBE et 7,5 mM de L-arginine a été utilisé. Avant de charger les échantillons sur le gel, une prémigration a été faite à 100 V, pendant 60 min. Ensuite, la migration a été faite à 4 °C et à 100 V, dans un tampon composé de TBE 0,5 X et 7,5 mM de L-arginine. La fixation de ArgR sur l'ADN a été effectuée, pendant 20 min à température pièce, dans un tampon contenant 20 mM Tris (pH 7.6), 10 mM MgCl₂, 5 mM CaCl₂, 100 mM KCl, 10 mM β-mercaptoéthanol, 20 mM L-arginine et 5 % de glycérol. Pour le dépôt des échantillons, une solution de charge a été utilisée (Bleu de dépôt 5 X : 60 % 0,25 X TBE, 40 % de glycérol, 0,2 % de bleu de bromophénol).

Les gels, n'ayant pas subi de transferts sur membrane, ont été réalisés avec 30 ng d'ADN (les promoteurs) et révélés par le SYBR Safe (Invitrogene). L'expérience avec ADN compétiteur non spécifique poly d(I-C) a été réalisée avec 20 fmol d'ADN biotinylé (=2,65 ng), selon le protocole du kit LightShift Chemiluminescent EMSA de la compagnie Thermo Scientifique. Brièvement, le transfert sur membrane de nitrocellulose a été effectué à 380 mA (≈ 100 V)

pendant 30 min, à 10 °C. Ensuite, les protéines et l'ADN ont été liés à la membrane par une exposition de 60 sec aux UV, à une puissance de 120 mJ/cm². Enfin, un conjugué de streptavidine et de peroxydase de raifort a été utilisé pour la révélation de l'ADN biotinylé, en présence de H₂O₂ et de luminol (substrat chimiluminescent).

13 Tests phénotypiques

13.1 Mesure de la production de biofilm

Les tests de biofilm ont été réalisés dans des plaques de 96 puits, faits en chlorure de polyvinyle flexible. Les plaques ont été stérilisées par de l'éthanol à 70 % et 30 min aux UV. Les précultures ont été faites dans 5 mL de LB, toute la nuit à 37 °C et agitées. Leur DO_{600nm} a été ajustée à 0,05. Ensuite, les plaques ont été inoculées avec 150 µL des cultures. Les conditions de culture sont variables selon les expériences, mais elles ont été toutes faites pendant 24 °C, sans agitation. Après, les plaques ont été lavées trois fois par immersion dans de l'eau distillée et séchées par tapotement. 175 µL de cristal violet à 1 % (p/v) ont été ajoutés. Les plaques ont été incubées pendant 15 min puis lavées trois fois par immersion dans de l'eau distillée. La dissolution des biofilms marqués a été faite avec 200 µL d'un mélange d'éthanol et acétone, à une proportion de 80 % d'éthanol pour 20 % d'acétone. 150 µL ont été transférés dans une plaque 96 puits pour lecteur de plaque. La DO a été prise à 595 nm.

13.2 Mesure de la nage des bactéries

Des géloses de LB à 0,25 ou 0,35 % d'agar ont été utilisées pour déterminer la nage des bactéries. Les cultures bactériennes ont été faites initialement sur des géloses LB, à partir des stocks de glycérols à -80 °C. Les colonies ont été repiquées avec un cône sur les géloses semi-solides. Après 24 h de croissance à 37 °C, le diamètre des halos a été mesuré dans deux différentes directions et leur moyenne a été utilisée.

13.3 Modèle d'infection du tractus urinaire murin

Toutes les souches utilisées pour les infections ont été cultivées premièrement dans 5 mL de bouillon LB à 37 °C, agitées et toute la nuit. Un mL de cette préculture a servi à inoculer

100 mL de bouillon LB. La croissance a été effectuée dans les mêmes conditions que précédemment. Après cela, les souches ont été mélangées ensemble uniformément, concentrées et les mélanges ont été maintenus sur glace jusqu'à l'infection des souris. Un échantillon de chaque mélange a été conservé, afin de déterminer la concentration bactérienne pour chaque souche. Les souris CBA/J, de 12 semaines, ont été anesthésiées avec un mélange de kétamine, de xylazine et d'acépromazine et inoculées par cathéter directement dans la vessie avec 20 μ L de l'inoculum ($\approx 5 \times 10^9$ bactéries). Après 48 h, les souris ont été sacrifiées (anesthésiées et ensuite euthanasiées par CO₂) et la vessie et les reins ont été récupérés. Puis, les organes ont été broyés dans un tampon de BSG (1.5 M NaCl, 20 mM KH₂PO₄, 50 mM Na₂HPO₄ (pH 7,7), 0,1 % (p/v) gélatine) et dilués pour inoculation de géloses MacConkey. Après 18 h de croissance à 30 °C, un décompte bactérien a été effectué.

14 Tests statistiques

Les tests statistiques utilisés dans cette étude sont : le test du χ^2 , avec la procédure de Morasculo, afin de déterminer avec quel groupe phylogénétique et avec quel séro groupe le fimbria Efo est associé. Concernant les résultats de β -galactosidase, ils ont été analysés avec le test *t* de Student non apparié et un ANOVA à deux facteurs, avec un test post-hoc de Bonferroni. Pour les résultats de nage, c'est seulement un test *t* de Student non apparié qui a été utilisé. Enfin, le test *t* de Wilcoxon a été employé pour les résultats des infections *in vivo* de souris.

RÉSULTATS

1 Identification du cluster de gènes *efoABCD*

Plusieurs transcrits, exprimés *in vivo* chez le poulet, ont été identifiés comme étant spécifiques à la souche APEC χ 7122 (4). Un de ces ARNm capturés code potentiellement pour une protéine *usher*. Après un séquençage des régions environnantes, quatre cadres de lecture ouverts, au total, ont été découverts. Cependant, l'un de ces cadres est interrompu par une IS2 (Annexe I). Ces cadres de lecture sont retrouvés aussi chez la souche ECOR72, entre les gènes conservés *exuRT* comme pour χ 7122, mais sans interruption de séquence (Fig. 14). Ces quatre cadres de lectures ont été nommés *efoABCD*, pour « exu-linked fimbrial operon ». Un fait remarquable est le pourcentage de bases G+C que compose ce cluster de gènes. En effet, il est de 44 % en moyenne (avec un taux de seulement de 38 % pour *efoA* et de 46 % pour les autres gènes). En comparant la région intergénique des gènes *exuRT* chez MG1655, et les régions *exuR-efoA* et *exuT-efoD*, il est à noter qu'un fragment de 31 pb est absent chez ECOR72. Le gène *efoA* fait 717 pb de long et produirait une protéine (EfoA) avec une taille prédite native de 238 aa. Le gène *efoB*, avec 501 pb, produirait une protéine (EfoB) de 166 aa. Le gène *efoC* avec une longueur de 2703 pb donnerait une protéine (EfoC) de 900 aa et enfin le gène *efoD*, de 1089 pb, coderait pour une protéine (EfoD) de 362 aa. Ces quatre protéines présentent un signal peptide en partie N-terminal de leur séquence (voir Annexe IV). Donc, les poids moléculaires des protéines EfoABCD sont, sans le peptide signal, de : 24,01 kDa pour EfoA, 15,42 kDa pour EfoB, 96,49 kDa pour EfoC et 38,26 kDa pour EfoD.

En effectuant une recherche de similitudes à partir des séquences prédites en acides aminés des protéines EfoABCD, plusieurs domaines conservés ont été révélés. Notamment pour EfoA, un domaine PRK15308 appartenant à la protéine TcfA (protéine du fimbria Tcf, chez *Salmonella enterica*) (264) et plus généralement avec les protéines chaperonnes ayant les domaines PFAM00345 et COG3121 (358, 359). Pour EfoB, c'est le domaine PFAM04449 de la sous-unité majeure du fimbria CS1 d'*Escherichia coli* qui possède la plus forte homologie (360). Le domaine PRK15310, appartenant à la protéine TcfC, donne le meilleur résultat pour l'homologie de séquence avec EfoC. De même, le domaine PFAM00577 de la famille des protéines *usher*

ressort de cette analyse (361). En dernier lieu, EfoD possède de fortes similitudes avec le domaine PFAM07434 de l'initiateur de la biogenèse du fimbria Cbl, la protéine CblD, de *Burkholderia cepacia* (362). Ces résultats permettent d'assigner des rôles putatifs pour chacune des protéines : EfoA serait la chaperonne, EfoB la sous-unité majeure, EfoC la protéine *usher* et enfin EfoD la protéine initiatrice du fimbria Efo.

Des expérimentations ont été effectuées pour déterminer si les gènes *efoABCD* sont en un seul opéron, chez la souche ECOR72. Les résultats obtenus ont pu démontrer que tous les gènes *efoABCD* sont sur le même ARNm (Fig. 15 et 16). Une bande d'amplification est observée pour le couple I (Fig. 16) aux alentours de 350 pb (puits « E »), qui démontre que les gènes *efoA* et *efoB* sont sur un seul transcrit. Il en va de même avec les couples II et III (Fig. 15) : une bande d'amplification aux alentours de 250 pb pour le couple II (puits « E » du couple II) et une bande aux alentours de 400 pb pour le couple III (puits « E » du couple III) sont observés, ce qui démontre que les gènes *efoBCD* sont sur un même ARNm. La taille de ces bandes correspond aux tailles attendues pour ces paires d'amorces (voir chapitre 6 du « Matériel et Méthodes »). Pour ces couples, les résultats des témoins négatifs sont valides. Mais, plusieurs problèmes sont apparus pour les autres paires d'amorces : des amplifications non spécifiques sur l'ADN génomique, en plus des amplifications attendues, sont observées pour les couples IV (bande à 400 pb au lieu de 739 pb), V (bande à 2000 pb au lieu de 3481 pb) et VI (bande à 550 pb au lieu de 2973 pb). Le témoin positif sur l'ADN génomique du couple III (entre les gènes *efoC* et *efoD*) n'a pas fonctionné dans cette expérience, cependant dans des expérimentations ultérieures ce témoin a été validé. De plus, plusieurs témoins négatifs (puits « T ») présentent une bande d'amplification, ce qui indique une contamination avec de l'ADN des échantillons d'extraction d'ARNm.

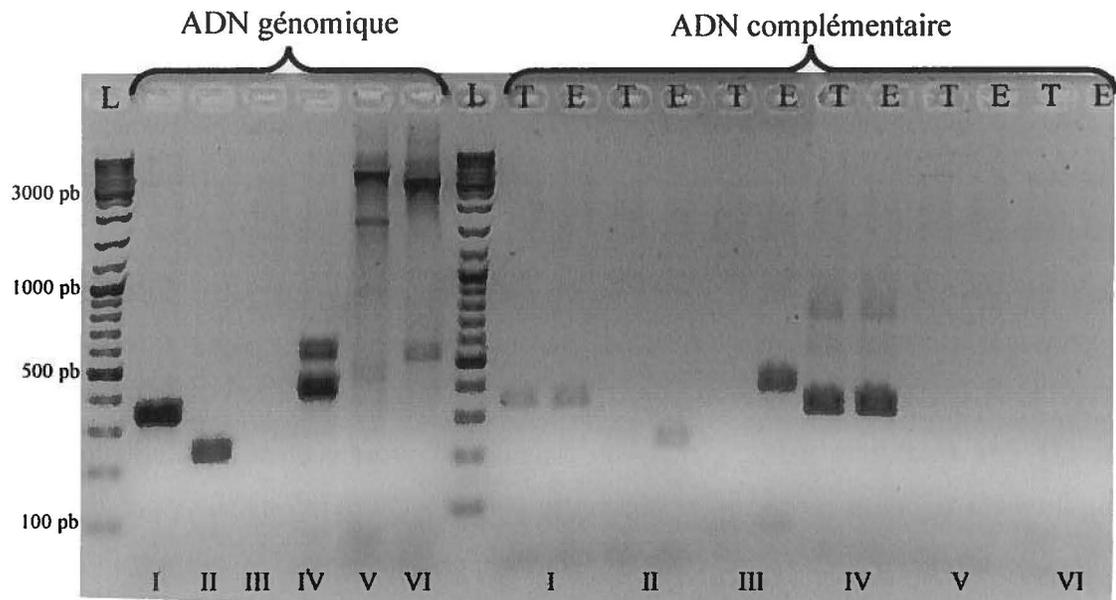


Figure 15 : Migration sur gel d'agarose des PCR, pour la détermination de la présence d'un opéron *efoABCD*
 Photographie d'une migration d'ADN sur gel d'agarose à 0,8%, avec deux sections : la première correspond à des PCR effectuées sur de l'ADN génomique pour représenter les tailles attendues des amplicons (témoins positifs). La seconde partie correspond à l'expérimentation à proprement dit, c'est-à-dire sur les ARNm traités ou non avec la reverse transcriptase. Les puits « L » pour le marqueur de taille. Les puits « T » sont les PCR « témoin négatif » sur l'ARNm sans reverse transcription. Les puits « E » sont les PCR sur l'ADN complémentaire. Les chiffres romains correspondent à chaque couple d'amorces utilisé : « I » pour CMD1535/1536 (*efoA-efoB*), « II » pour CDM1537/1538 (*efoB-efoC*), « III » pour CDM1539/1540 (*efoC-efoD*), « IV » pour CMD1535/1538 (*efoA-efoC*), « V » pour CMD1535/1540 (*efoA-efoD*) et « VI » pour CMD1537/1540 (*efoB-efoD*).

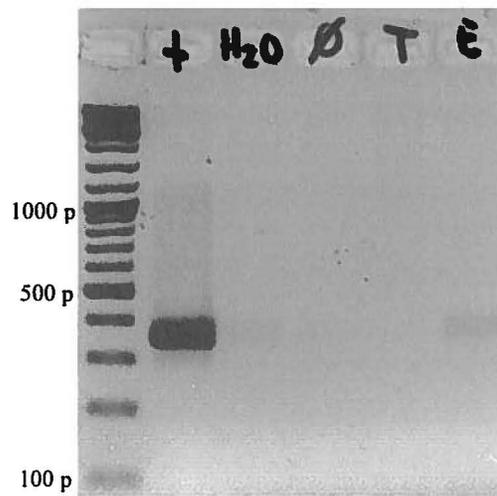


Figure 16 : Migration sur gel d'agarose des PCR, pour la détermination de la présence des gènes *efoAB* sur un même transcrit

Photographie d'une migration d'ADN sur gel d'agarose à 1%. Le puits « + » est le contrôle positif sur l'ADN génomique. Le puits « H₂O » est un contrôle négatif sur de l'eau. Le puits « ø » est un contrôle négatif sans Taq polymérase. Le puits « T » est un contrôle négatif sur l'ARNm sans reverse transcription. Le puits E est la PCR sur l'ADN complémentaire, avec les amorces du couple « I » soit CMD1535/1536, entre les gènes *efoA* et *efoB*.

2 Visualisation par MET et par immunobuvardage de type Western

Afin de déterminer si les gènes *efoABCD* produisent des fimbriae à la surface des bactéries, le cluster a été cloné dans un vecteur à haut nombre de copies (pIJ164) chez la souche afimbriaire ORN172 (363). En outre, deux autres vecteurs ont été produits par des PCR inverses, afin d'éliminer le gène *efoB* (pIJ425) ou le gène *efoD* (pIJ426). Plusieurs structures filamenteuses péritriches sont observables par microscopie électronique à transmission, à la surface de la souche QT1306, contenant le plasmide pIJ164 (Fig. 17A). Ce qui n'est pas le cas pour QT3044 (plasmide pIJ425), mais un flagelle est présent chez cette souche (Fig. 18A). Contrairement aux deux précédentes souches, la photo de QT3045 (pIJ246) présente un fragment de flagelle (des flagelles rattachés aux bactéries ont été observés), ainsi qu'une ou deux structures filamenteuses (Fig. 18B). La largeur mesurée de ces structures est de 6 nm. De plus, lorsque *efoABCD* est exprimé constitutivement, ces structures se regroupent en fuseaux et semblent relier les bactéries entre elles (Fig. 17B).

Deux autres photographies furent prises avec la souche QT1307, ceci afin d'observer l'élimination de l'effet du promoteur *lac* sur l'expression du cluster *efoABCD* (Fig. 18). En effet, contrairement à la souche QT1306 (pIJ164), dans le plasmide pIJ165, le cluster est cloné à l'inverse du promoteur *lac*. Cela permet de prendre en compte seulement l'expression par le promoteur natif. Sans l'effet du promoteur *lac*, moins de structures sont visibles à la surface des bactéries (Fig. 19A), mais comme pour QT1306, elles semblent relier les bactéries entre elles (Fig. 19B).

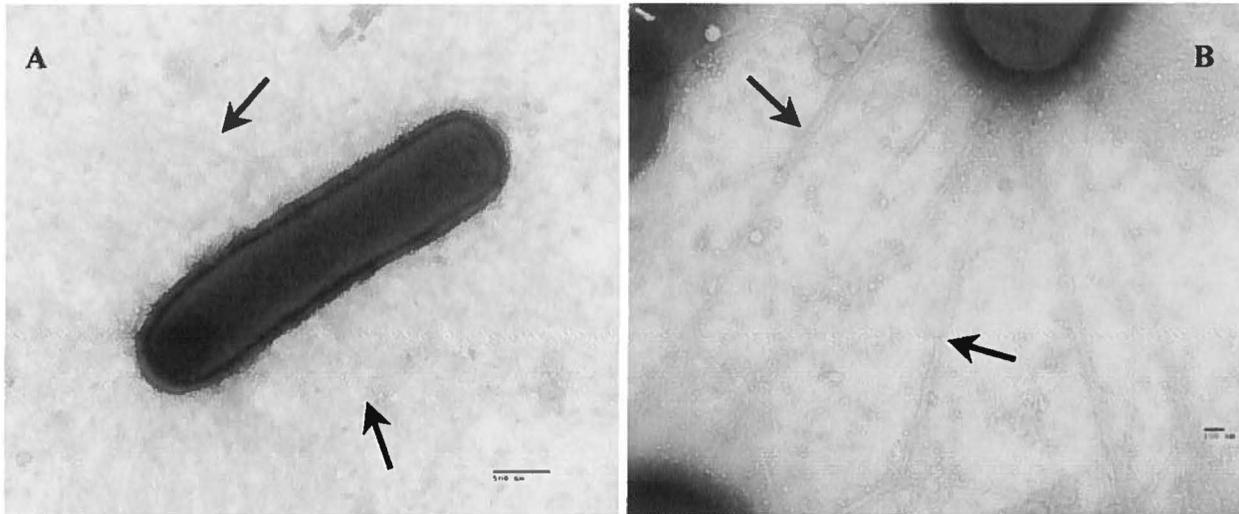


Figure 17 : Photos de microscopie électronique à transmission de la souche QT1306

A : Photo à un grossissement de 9000 fois. L'échelle, figurée par la barre noire, représente 500 nm. Des structures filamenteuses sont présentes tout autour de la bactérie (flèches noires). **B** : Photo à un grossissement de 17000 fois. L'échelle représente 100 nm. Les structures filamenteuses se regroupent en fuseaux et relient les bactéries entre elles (flèches noires).

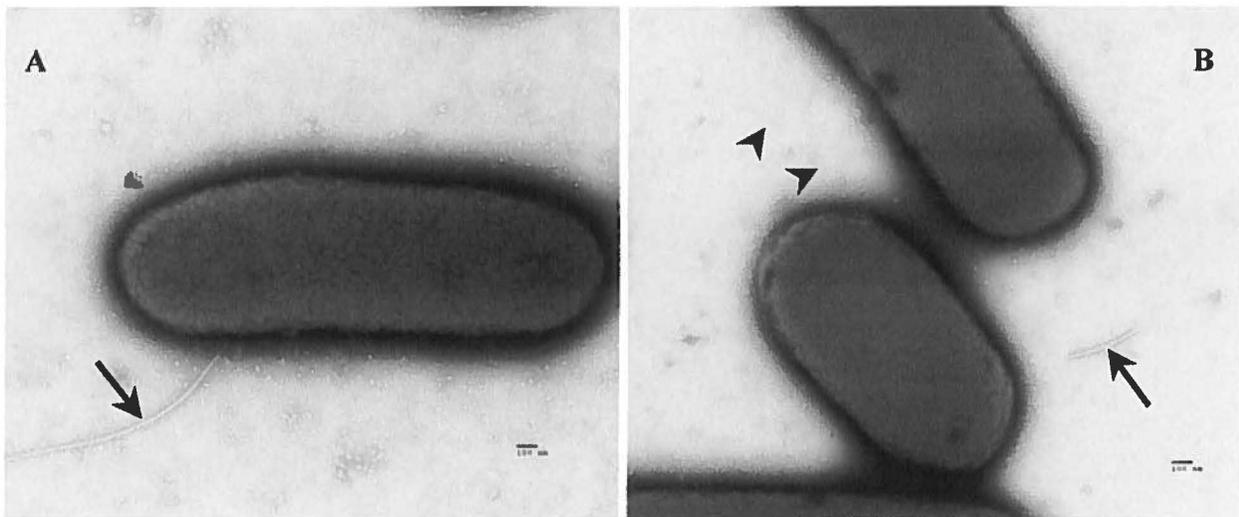


Figure 18 : Photos de microscopie électronique à transmission des souches QT3044 et QT3045

Les photos sont prises à un grossissement de 17000 fois. La barre noire représente une échelle de 100 nm. **A** : Aucune structure n'est observée sauf la présence d'un flagelle (flèche noire). **B** : Très peu de structures filamenteuses sont observables (têtes de flèches noires). Un morceau de flagelle est aussi présent (flèche noire).

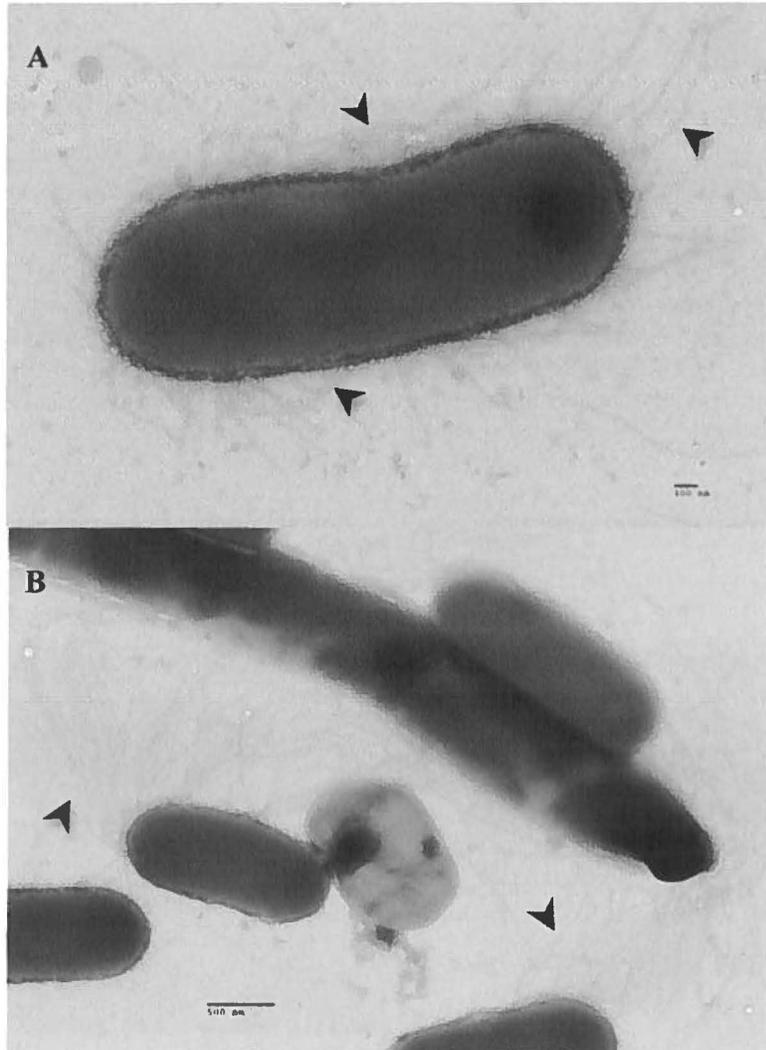


Figure 19 : Photos de microscopie électronique à transmission de la souche QT1307

A : La photo est prise à un grossissement de 17000 fois. La barre noire représente une échelle de 100 nm. Contrairement à la souche QT1306, moins de structures filamenteuses sont observables (têtes flèches noires). Aucun flagelle n'a été observé. **B :** La photo est prise à un grossissement de 9000 fois. La barre noire représente une échelle de 500 nm. Les structures filamenteuses semblent relier les bactéries entre elles (têtes flèches noires).

Pour confirmer que les structures visibles correspondent au fimbria Efo, un sérum polyclonal, reconnaissant un peptide de la sous-unité majeure (EfoB), a été utilisé sur une extraction fimbriaire et un lysat bactérien, par un immunobuvardage de type Western (Fig. 20). Par ailleurs, un contrôle négatif, soit une souche (QT885) ayant seulement le plasmide vide pBC SK+, a été rajouté. Comme attendu, une seule bande, aux alentours de 15 kDa, est révélée pour l'extraction fimbriaire sur la souche QT1306. De même, une autre bande pour QT1306 et QT3045 est observée, sur les échantillons de lysats bactériens. De plus, aucune bande n'est présente pour les souches QT885 et QT3044, ceci montre la spécificité de l'anticorps anti-EfoB.

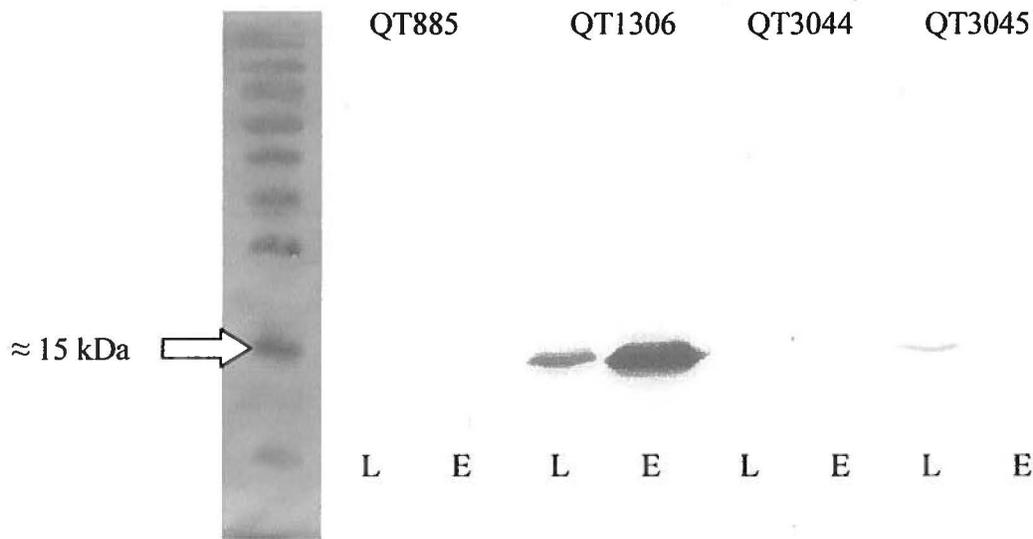


Figure 20 : Révélation de l'immunobuvardage de type Western avec l'anticorps anti-EfoB

L'échelle de poids moléculaire est la bande grise à gauche de l'image (PageRuler Plus Prestained Protein Ladder, Fermentas). Les puits « L » correspondent aux lysats bactériens et les puits « E » aux extractions fimbriaires. Le nom des souches utilisées est écrit en haut, chacune ayant un puits « L » et un puits « E ». Au total trois bandes sont révélées par l'anticorps anti-EfoB, aux environs de 15 kDa : deux bandes, lysat et extraction, pour la souche QT1306 et une seule bande pour le lysat bactérien sur QT3045. Aucune bande n'est observée pour les souches QT885 et QT3044, démontrant la spécificité de l'anticorps.

3 Distribution de *efoC* parmi la collection ECOR, des souches porcines (PPEC) et des souches aviaires (APEC)

Afin de déterminer la prévalence de l'opéron *efoABCD*, un criblage de la collection ECOR, ainsi que de diverses souches PPEC et APEC disponibles au laboratoire, a été effectué par PCR. Les amorces employées sont CMD378 et CMD1096. La région amplifiée est spécifique au gène

efoC. De plus, les paramètres de la PCR prennent en compte la présence potentielle d'une IS2. Donc les fragments obtenus sont de 1235 pb pour le gène natif et de 2571 pb pour un gène *efoC* interrompu par une IS. Les résultats sont présentés dans le tableau 3.

Les souches de la collection ECOR peuvent être regroupées en cinq groupes phylogénétiques par MLST (le groupe E étant un groupe à part, non typable par la méthode de Clermont, voir chapitre « 1.4.1 : Méthode Clermont »). La présence de *efoC* est significativement associée avec le groupe B1 comparativement aux groupes A, B2 et D (Tableau 3). Concernant les souches PPEC et APEC, leur groupe phylogénétique a été déterminé par la méthode de Clermont. Seul le groupe phylogénétique B1 est significativement lié avec *efoC*, pour les PPEC. Un résultat remarquable est l'absence de souches du groupe B1 pour les APEC, malgré le nombre important de souches testées. Une forte proportion de souches A positives avec et sans IS est à noter.

Groupe phylogénétique	Valeur pour <i>E. coli</i> provenant de						
	Collection ECOR		PPEC		APEC		
	Nb total de souches	Nb de souches positives pour <i>efo</i> (%)	Nb total de souches	Nb de souches positives pour <i>efo</i> (%)	Nb total de souches	Nb de souches positives pour <i>efo</i> (%)	
		Sans IS		Sans IS		Sans IS	Avec IS
A	26	7 (27)	22	7 (32)	77	28 (36)	43 (56)
B1	16	16 (100)*	3	3 (100)*	0	0 (0)	0 (0)
B2	15	2 (13)	0	0	50	0 (0)	6 (12)
D	12	6 (50)	3	0 (0)	28	1 (4)	1 (4)
E	3	0 (0)	NA	NA	NA	NA	NA
Sous-Total						29 (19)	50 (32)
Total	72	31 (43)	28	10 (36)	155	79 (51)	

Tableau 3 : Distribution des souches positives pour *efo* parmi la collection ECOR, des souches PPEC et APEC

NA : non applicable

* : Les résultats de la collection ECOR et pour les PPEC, pris ensemble, donnent une association significative du fimbria Efo avec le groupe phylogénétique B1 ($P < 0,0001$ pour le test du χ^2). Seuls les groupes A, B1, B2 et D ont été pris en compte.

En plus du regroupement phylogénétique des souches, les souches APEC ont été classées selon leur sérotype. Cela donne un résultat significatif quant à l'association du gène *efoC* avec

les APEC de séro groupe O78 (Tableau 4). Une seule souche est positive et sans IS chez les deux autres sérogroupe, sinon les autres souches positives présentent une interruption de séquence.

Groupe Phylogénétique	Valeur pour <i>E. coli</i>								
	APEC-O1			APEC-O2			APEC-O78*		
	Nb total de souches	Nb de souches positives pour <i>efo</i> (%)		Nb total de souches	Nb de souches positives pour <i>efo</i> (%)		Nb total de souches	Nb de souches positives pour <i>efo</i> (%)	
		Sans IS	Avec IS		Sans IS	Avec IS		Sans IS	Avec IS
A	3	0 (0)	1 (33)	3	0 (0)	0 (0)	71	28 (39)	42 (59)
B1	0	0 (0)	0 (0)	0	0 (0)	0 (0)	0	0 (0)	0 (0)
B2	20	0 (0)	1 (5)	25	0 (0)	2 (8)	5	0 (0)	3 (60)
D	3	1 (33)	0 (0)	17	0 (0)	1 (6)	8	0 (0)	0 (0)
Total	26	1 (4)	2 (8)	45	0	3 (7)	84	28 (33)	45 (54)

Tableau 4 : Distribution des souches positives pour *efo* parmi les différents sérogroupe des APEC

* : Le fimbria Efo est significativement associé avec le séro groupe O78, comparé aux deux autres sérogroupe ($P < 0,0001$ pour le test du χ^2).

Il est intéressant de noter que le report des souches positives pour *efo* sur l'arbre phylogénétique, conçu à partir des données de MLST pour la collection ECOR (364), montre des regroupements clonaux pour le cluster *efo* (Fig. 21). Sauf pour deux souches ECOR7 (groupe A) et ECOR63 (groupe B2).

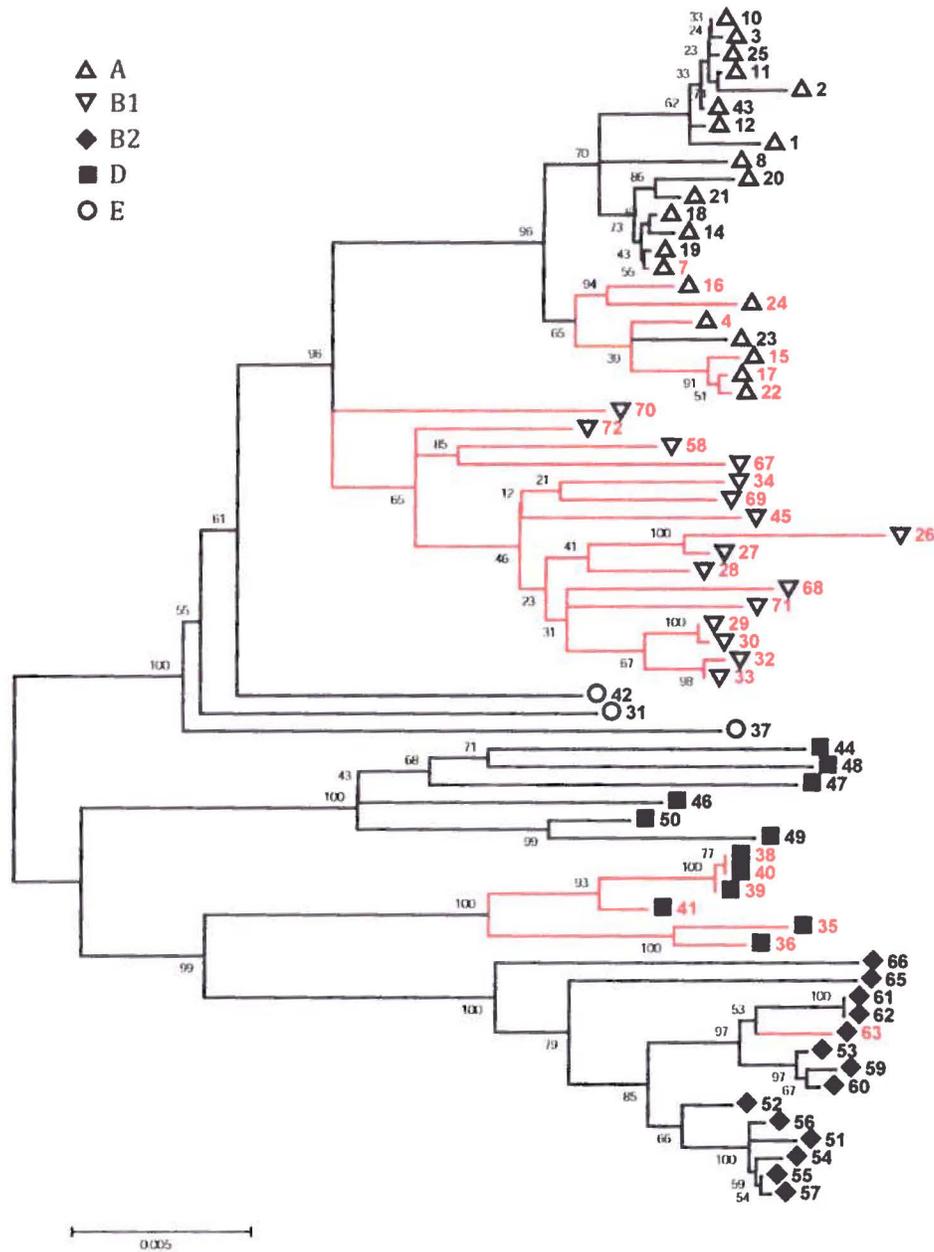


Figure 21 : Arbre phylogénétique de la collection ECOR à partir des séquences MLST

Arbre phylogénétique construit avec l'algorithme « neighbor-joining ». L'échelle représente 0,005 substitution par site. Les branches et les numéros des souches (en gras) de la collection ECOR, positives pour *efo*, sont représentés en rouge. Les numéros sont les scores du bootstrap. Les souches ECOR 5 ; 6 ; 9 ; 13 ; 64 ne sont pas représentées, car non séquencées pour le MLST (364). Seule la souche ECOR43 change de groupe phylogénétique, elle se retrouve dans le groupe A, comparé au MSLT où elle se situe dans le groupe E. Toutes les souches du groupe B1 possèdent le fimbria Efo. De plus, des groupes clonaux positifs sont observables dans le groupe A et D.

4 Relation phylogénétique des différentes séquences pour les gènes *efoABCD* et leur protéine respective

En effectuant un Blastn, à partir de la séquence nucléotidique de *efoC*_{ECOR72}, 23 souches bactériennes (22 d'*Escherichia coli* et une de *Shigella sonnei*) sont ressorties ayant une séquence avec une très forte homologie avec ce gène. Cette séquence correspond à *efoC*. Dans chacun des génomes bactériens, il fait partie d'un cluster de quatre gènes, avec des rôles présumés de protéines de fimbriae. Ces gènes se situent toujours entre les gènes *exuRT*. Pour confirmer ce résultat, quatre autres Blastn ont été effectués avec la séquence de l'opéron *efoABCD* de ECOR72 ou avec les gènes séparément. Les mêmes souches et les mêmes cadres de lecture ouverts sont ressortis de ces analyses. De plus, plusieurs souches de *Shigella* s'avèrent posséder des reliques de l'opéron, mais plusieurs IS interrompent les gènes.

Une analyse phylogénétique, à partir du résultat de recherche de similitudes avec *efoC*_{ECOR72}, permet de distinguer trois groupes de souches et donc trois allèles différents pour ce gène (Fig. 22). Le premier groupe est composé de la souche 536 et SMS-3-5, le second des souches ATCC8739, IAI39 notamment. Enfin, le troisième et dernier groupe est composé des souches ECOR72, W, IAI1, entre autres. Ce résultat est similaire aux arbres obtenus avec les trois autres gènes. La présence de l'opéron *efoABCD* ne se limite pas aux souches UPEC/APEC/PPEC : des souches environnementales, comme SMS-3-5, des souches pathogènes intestinales (E24377A, par exemple) et des souches non pathogènes (comme W ou HS) le possèdent.

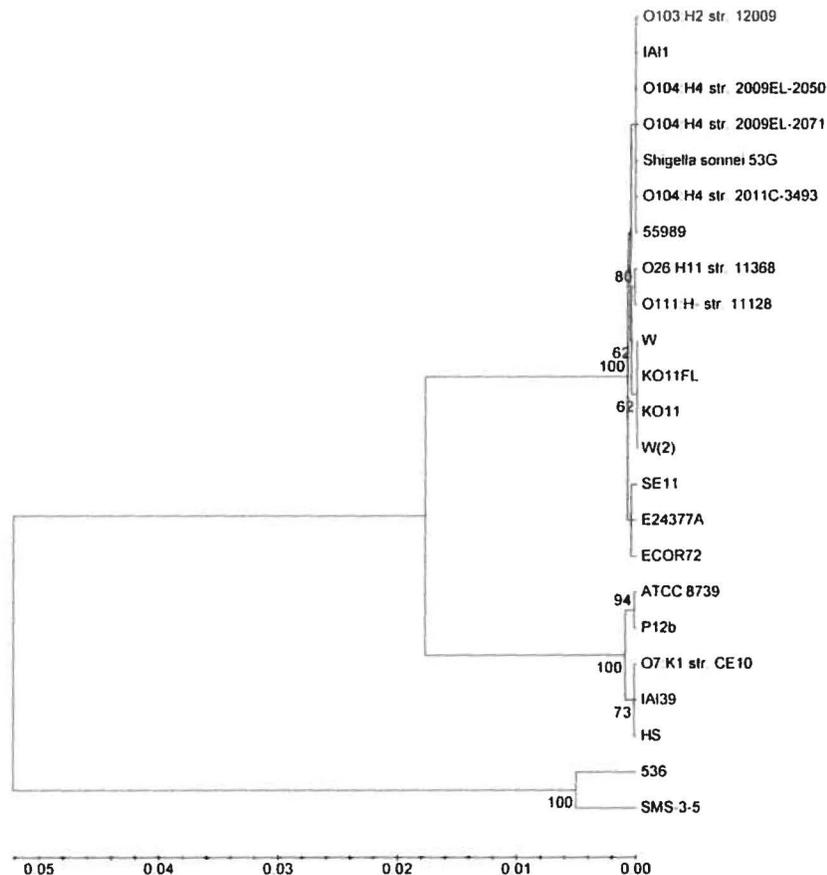


Figure 22 : Arbre phylogénétique du gène *efoC* des souches obtenues par Blastn

Arbre phylogénétique construit avec l'algorithme « neighbor-joining ». Les chiffres correspondent aux différents scores du bootstrap. L'échelle représente le nombre de substitutions par site. Trois regroupements ressortent de cette analyse. Le gène *efoC* est présent parmi plusieurs pathotypes et des souches non pathogènes.

5 Relation phylogénétique avec les autres fimbriae de la voie du « chaperonne-usher »

L'étude de Nuccio *et al.* (2007) utilise la séquence des protéines *usher* de tous les fimbriae connus de la voie « chaperonne-usher ». Les protéines *usher*, n'étant pas en contact direct avec l'environnement et le système immunitaire, évoluent plus faiblement que les autres protéines. De plus, c'est une protéine conservée dans la voie « chaperonne-usher ». Donc, en se basant sur cette même approche et en incluant les trois allèles de *efoC*, soit des souches UPEC 536, ECOR72 et IAI39, l'arbre phylogénétique des protéines *usher* a été reconstruit (Fig. 23). Les séquences, n'ayant pas de numéro d'accèsion, n'ont pas été incluses. En outre, deux séquences ne sont pas classées comme dans la précédente étude. Ceci est peut-être dû au nombre de

séquences utilisées, au nombre de répliques pour le bootstrap ainsi que des logiciels employés. Cet arbre indique que le fimbria Efo fait partie de la famille de fimbriae de la voie alterne du « *chaperonne-usher* », soit les fimbriae de la classe α .

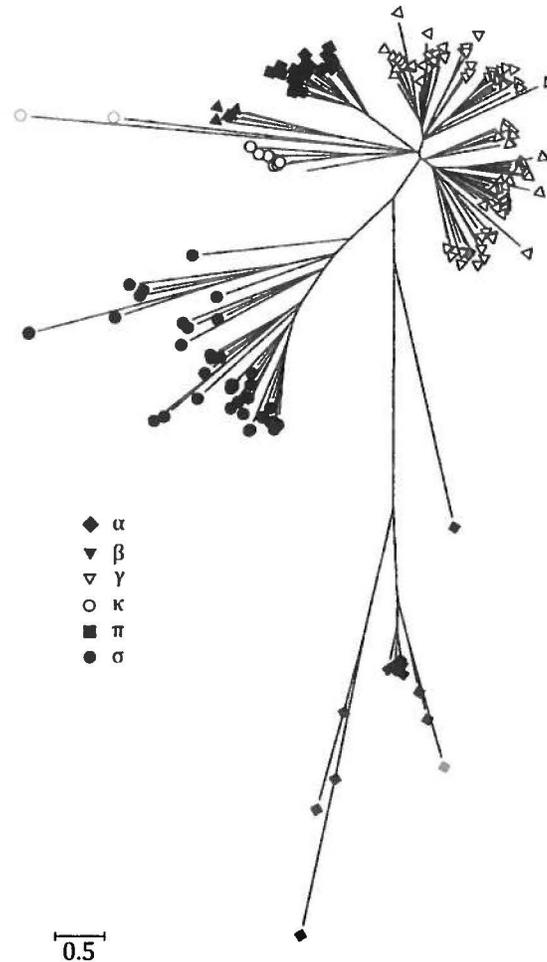


Figure 23 : Arbre phylogénétique des protéines *usher* des fimbriae de la voie « *chaperonne-usher* »

Arbre phylogénétique construit avec l'algorithme « neighbor-joining ». L'échelle représente 0,5 substitution par site. Pour alléger l'image, les scores du bootstrap ne sont pas affichés sur cet arbre, mais les scores des séparations en différentes classes sont significatifs. Les ronds verts représentent les deux séquences classées différemment par rapport à l'étude de Nuccio *et al.* (2007) (200). Le losange rouge correspond aux séquences de *efoC*, faisant partie des fimbriae de la classe α .

Un zoom sur la classe α des fimbriae, permet de montrer que le fimbria Efo a plus d'homologie avec le fimbria Tcf de *Salmonella enterica* et avec celui de *Burkholderia cepacia*, Cbl, qu'avec les fimbriae CS et CFA/I retrouvés chez *Escherichia coli* (Fig. 24). Ceci est en accord avec les domaines conservés révélés par la recherche de similitudes sur les protéines EfoABCD.

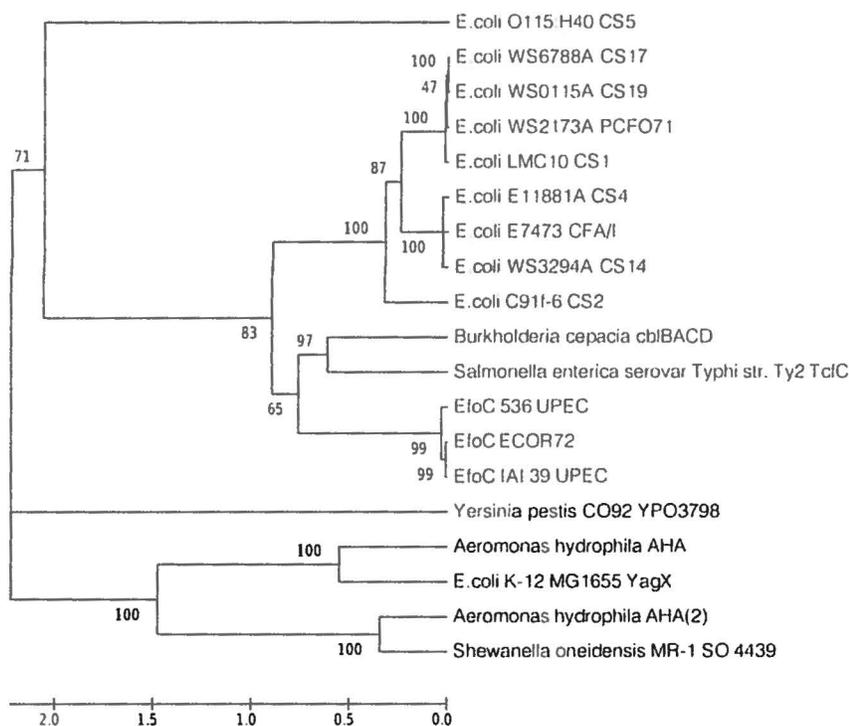


Figure 24 : Arbre phylogénétique des fimbriae- α (zoom de la Figure 23)

Arbre phylogénétique construit avec l'algorithme « neighbor-joining ». L'échelle représente le nombre de substitutions par site. Les chiffres sont les scores pour le bootstrap. Les trois allèles du gène *efoC* sont plus fortement reliés phylogénétiquement avec les fimbriae Cbl et Tcf de *Burkholderia cepacia* et de *Salmonella enterica* respectivement.

6 Analyse *in silico* du promoteur présumé du cluster *efoABCD*

Une analyse par le logiciel en ligne BPROM de SoftBerry, des 200 pb en amont du gène *efoA*, a permis de révéler les séquences présomptives -10 et -35, ainsi que des séquences de fixation pour des régulateurs transcriptionnels globaux comme Lrp et Fnr (Annexe V). Trois séquences ont été utilisées, celles des souches ECOR72, IAI39 et 536. Elles suivent la relation phylogénétique, démontrée plus haut, selon laquelle il existe trois allèles de l'opéron *efoABCD*. Un second promoteur est détecté pour la séquence de la souche 536, malgré la présence des mêmes séquences -10 et -35 des séquences de ECOR72 et IAI39 (Fig. 25) chez cette souche. Il est intéressant de noter que deux séquences pour le régulateur spécifique du métabolisme de l'arginine, ArgR, sont aussi rapportées.

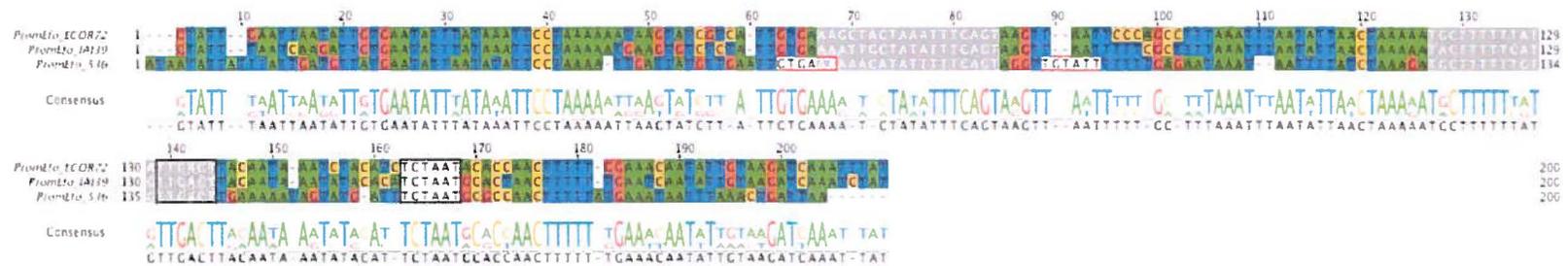


Figure 25 : Analyse des promoteurs présumés de l'opéron *efoABCD*

Alignement des trois séquences de 200 pb en amont du gène *efoA*, des souches ECOR72, IA139 et 536. La séquence consensus est affichée en dessous. Les séquences -10 et -35, conservées pour les trois séquences, sont dans les cadres noirs. Les deux autres séquences supposées pour la séquence de 536 sont dans des cadres rouges. Les sites de fixations présumés pour la protéine ArgR sont coloriés en gris.

Les séquences -10 et -35 se situent aux positions 168 et 144 respectivement. L'homologie est de 100 % (6/6) pour la -10 (séquence consensus TCTAAT) et de 83 % (5/6) pour la -35 (séquence consensus TTGACA) et cela pour les trois séquences. Alors que les séquences -10 et -35 de la souche 536, aux positions 94 et 68 respectivement, ne possèdent que 66 % d'homologie (4/6) avec les séquences consensus. L'investigation sur le rôle de la protéine ArgR est plus approfondie, étant donné qu'aucun autre rôle principal, outre la régulation des gènes du métabolisme de l'arginine, n'est connu (voir chapitre « 3.2 : La protéine ArgR »). Le calcul d'un score, donnant la conservation générale de la séquence consensus, a été établi par l'étude de Mironov *et al.* (1999). En utilisant la même table de calcul, les scores des sites de fixation ARG ont été déterminés. Pour les premières séquences ARG, en position 67 (Fig. 25), les scores sont : 2,27 pour ECOR72, 1,87 pour IAI39 et 2,15 pour 536. De même, pour la seconde séquence en position 127 : 2,03 pour ECOR72, 1,02 pour IAI39 et 0,96 pour 536. Ce sont des scores plus faibles que ceux retrouvés pour les gènes du métabolisme de l'arginine, qui sont plus proches de 4 (330) et sachant que le score maximal peut être de 4,83.

7 Régulation par différents facteurs environnementaux de l'expression du cluster *efo*

Afin de déterminer les conditions d'expression de l'opéron *efo* dans une souche native possédant cet opéron, un plasmide comprenant toute la région intergénique et avec une partie du gène *exuR* et *efoA*, en fusion avec *lacZ*, fut inséré dans une souche ECOR72 Δ *lac* (QT1469), pour donner la souche QT2922. De plus, pour obtenir un contrôle négatif (QT2921) et effectuer une normalisation des résultats, un plasmide vide (pRS415) fut aussi utilisé chez QT1469. Différentes conditions de culture ont été testées dont : l'oxygénation, le type de milieu et la température (Fig. 26). Les résultats obtenus sont une augmentation significative de l'expression de la β -galactosidase en milieu liquide par rapport en milieu solide ($P < 0,05$) (Fig. 26A), en présence d'oxygène comparé à un milieu en microaérobie ($P < 0,001$) (Fig. 26B) et à une température de 42 °C en comparaison à 30 °C ($P < 0,01$) et à 37 °C ($P < 0,0001$) (Fig. 26C). Tous ces résultats ont été obtenus en milieu riche (LB).

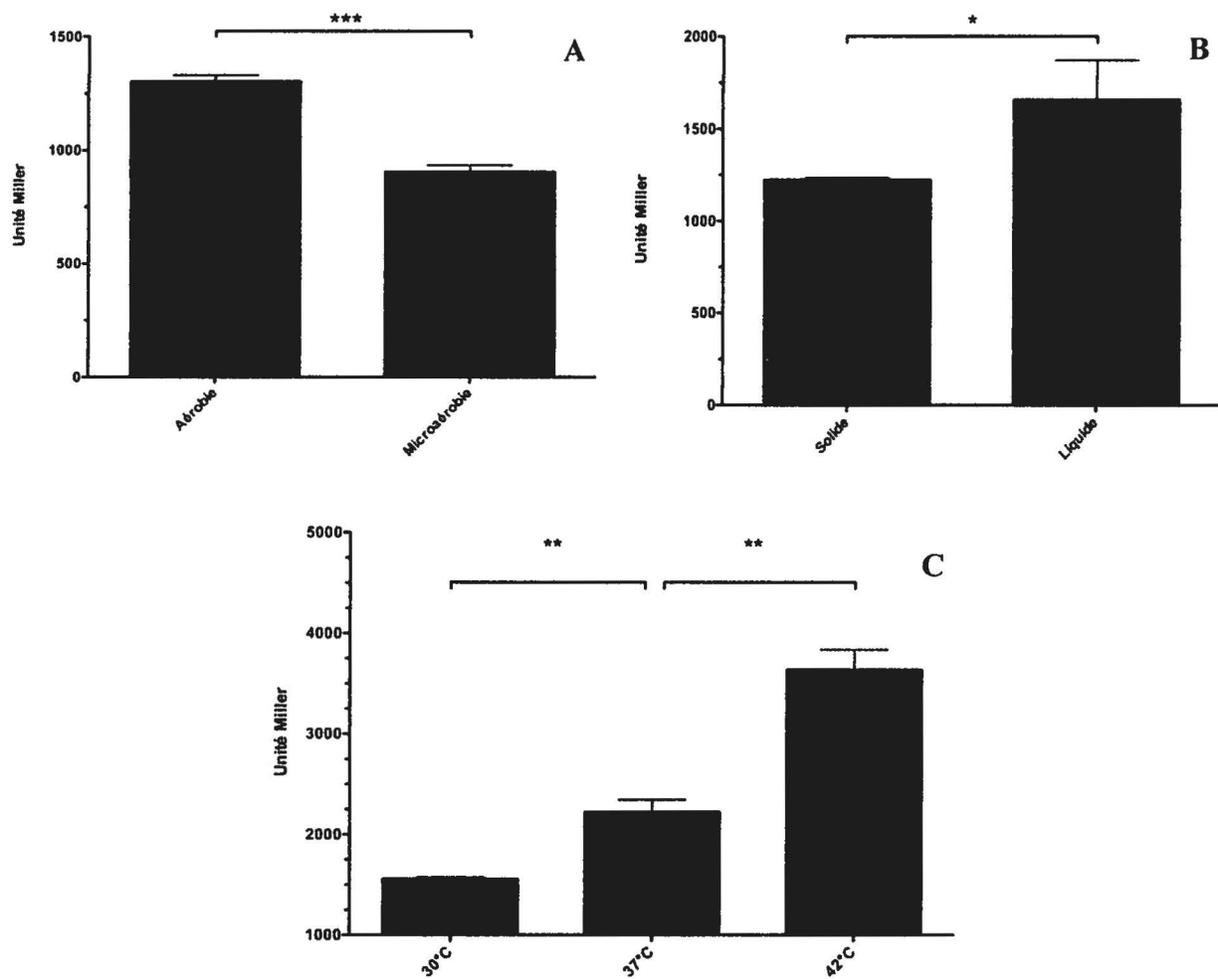


Figure 26 : Expression de la β -galactosidase sous contrôle du *PefoABCD* et selon différentes conditions de culture

Dosage de l'activité β -galactosidase (en unité Miller) exprimée par la souche QT2922, avec une normalisation des résultats par l'activité β -galactosidase exprimée chez la souche QT2921. Les bactéries ont été cultivées en milieu riche, sur milieu solide sauf indication contraire. * indique $P < 0,05$; ** indique $P < 0,01$ et *** indique $P < 0,001$. **A** : Activité de la β -galactosidase selon l'apport en oxygène dans le milieu. **B** : Activité de la β -galactosidase selon le type de milieu. **C** : Activité de la β -galactosidase selon la température de culture.

Une autre condition de culture a été testée, soit les différentes sources de carbone fournies par des sucres (Fig. 27). Les cultures bactériennes ont été effectuées en milieu minimum solide M9. La plus forte expression de la β -galactosidase est observée en présence de mannose et de ribose. La plus faible lorsque du xylose, mannitol ou arabinose est ajouté au milieu. Cela représente une différence d'activité de 1,6 fois entre le mannose et l'arabinose. En outre, une expression intermédiaire de l'activité β -galactosidase est mesurée avec le glucose, le glucuronate, le galacturonate et le maltose. Enfin, avec le glycérol, le sorbitol, le succinate, le pyruvate et l'acétate, la croissance était insuffisante, dans le temps imparti, pour effectuer un test de β -galactosidase. Le glucuronate et le galacturonate ont été utilisés, car l'opéron *efoABCD* est inséré entre deux gènes, *exuRT*, impliqués dans le métabolisme de ces sucres.

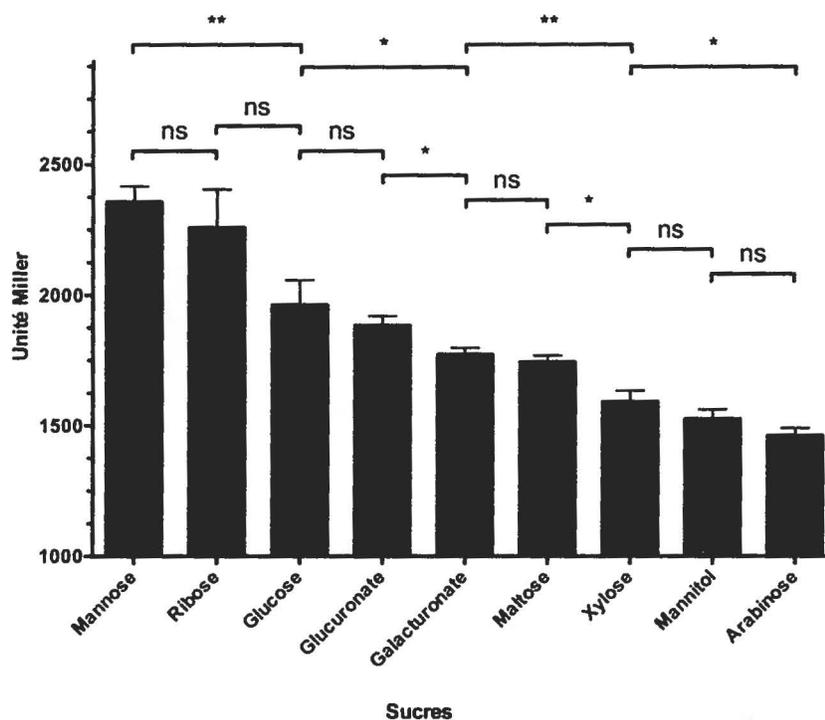


Figure 27 : Expression de l'activité β -galactosidase selon différentes sources de carbone

Les souches QT2921 et QT2922 ont été cultivées en milieu minimum M9, complété avec différentes sources de carbone. Le terme « ns » signifie « non significatif » ; * indique $P < 0,05$ et ** indique $P < 0,01$.

Par l'analyse *in silico* du promoteur putatif de *efoABCD*, deux séquences de fixation pour le régulateur ArgR ont été mises en évidence. Pour tester l'hypothèse, comme quoi cette protéine régulerait négativement l'expression du fimbria Efo, deux autres souches ont été créées. Soit une

souche ECOR72 Δ lac Δ argR::Km^r (QT2967) avec les plasmides pIJ407 et pRS415 (QT2971 et QT2970 respectivement) et une seconde, complétée simple copie dans le chromosome avec le gène *argR*, ECOR72 Δ lac Δ argR::Km::Tn7T-Gm::argR (QT3021). Elle a été aussi transformée avec les plasmides pRS415 et pIJ407 pour donner les souches QT3026 et QT3027 (349). Elles ont été cultivées sur un milieu solide M9, complété avec du glucose et avec ou sans 1 mM d'arginine (Fig. 28). Tout d'abord, les résultats obtenus pour la souche QT2922 sont une diminution significative de l'activité β -galactosidase en présence d'arginine ($P < 0,01$). De plus, lorsque le gène *argR* est muté, l'expression augmente significativement ($P < 0,001$) en moyenne de 1,2 fois, en présence ou en absence d'arginine dans le milieu. À cela s'ajoute une expression plus forte en présence d'arginine, chez cette souche mutée. Finalement, chez QT3027, où le gène *argR* a été rajouté simple copie dans le chromosome, l'activité de la β -galactosidase retrouve les valeurs mesurées pour QT2922. Ceci démontre un effet de la protéine ArgR sur l'expression de la β -galactosidase, sous le contrôle du promoteur des gènes *efoABCD*.

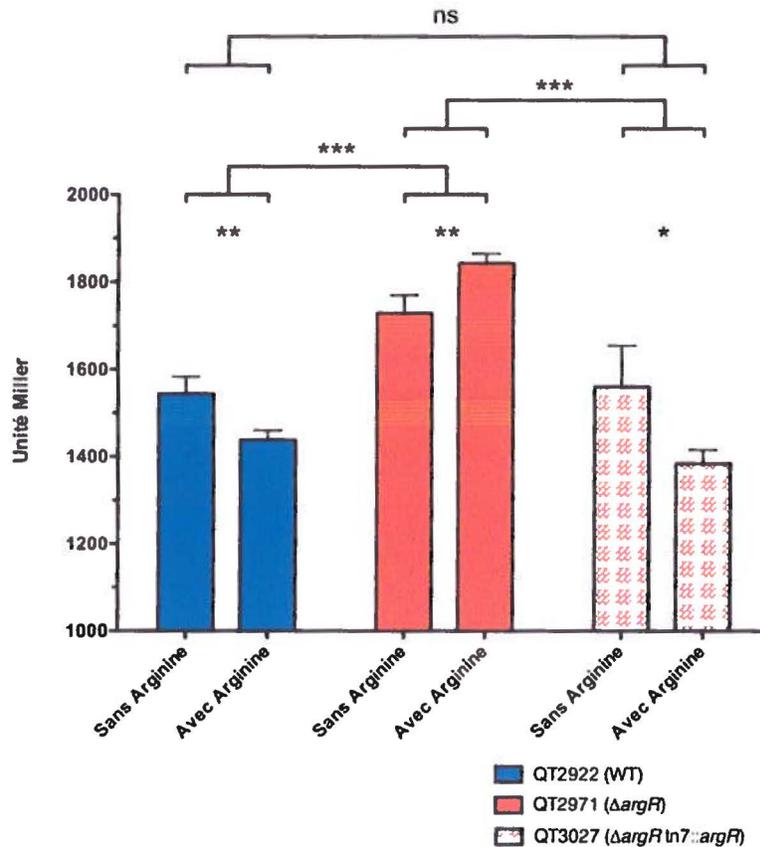


Figure 28 : Étude de l'influence de la protéine ArgR sur l'expression de l'activité β -galactosidase.
Trois souches ont été utilisées : QT2922, ECOR72 mutée pour l'opéron *lac* (WT); QT2971 qui possède, en plus de la souche QT2922, une suppression du gène *argR* ; QT3027 qui est la souche QT2971, complétée Tn7 avec *argR*. Le terme « ns » signifie « non significatif » ; * indique $P < 0,05$; ** indique $P < 0,01$; *** indique $P < 0,001$.

8 Expériences d'interactions de la protéine ArgR sur le promoteur de *efoABCD*

Par les tests de β -galactosidase, il a été montré que la protéine ArgR a une incidence sur des gènes qui sont sous le contrôle du promoteur de *efoABCD*. Pour démontrer que cette protéine interagit directement avec ce promoteur, des expériences de retard sur gels de polyacrylamide ont été effectuées. Les premières conditions sont simples : c'est la mise en contact direct du promoteur *efo* avec plusieurs quantités de la protéine ArgR. Pour cela, il est nécessaire de produire de la protéine ArgR pure. À cette fin, une protéine recombinante, flanquée d'un marqueur à six histidines, a été produite chez une souche Tuner (DE3) d'*Escherichia coli*. Des expériences d'expressions et de solubilités ont été réalisées afin de déterminer les meilleures conditions de purifications. En accord avec ces tests préliminaires, les cellules ont été mises en

croissance jusqu'à une $DO_{600\text{ nm}}$ de 0,6, ensuite elles ont été lysées et enfin la protéine ArgR a été purifiée de façon soluble par passage à travers une colonne d'acide nitrilotriacétique de nickel et par dialyse, pour éliminer l'imidazole. Les gels de vérification de la purification de la protéine sont présentés ci-dessous (Fig. 29). La protéine ArgR est obtenue avec une pureté suffisante dans les fractions d'élution (puits « E »).

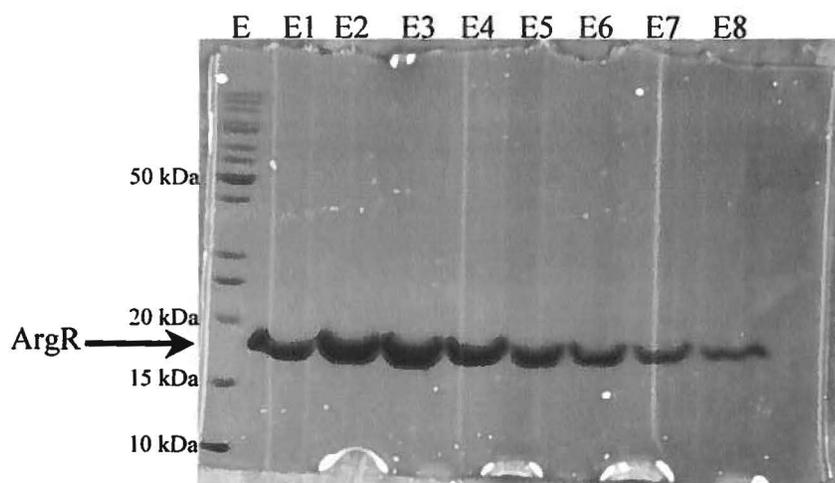
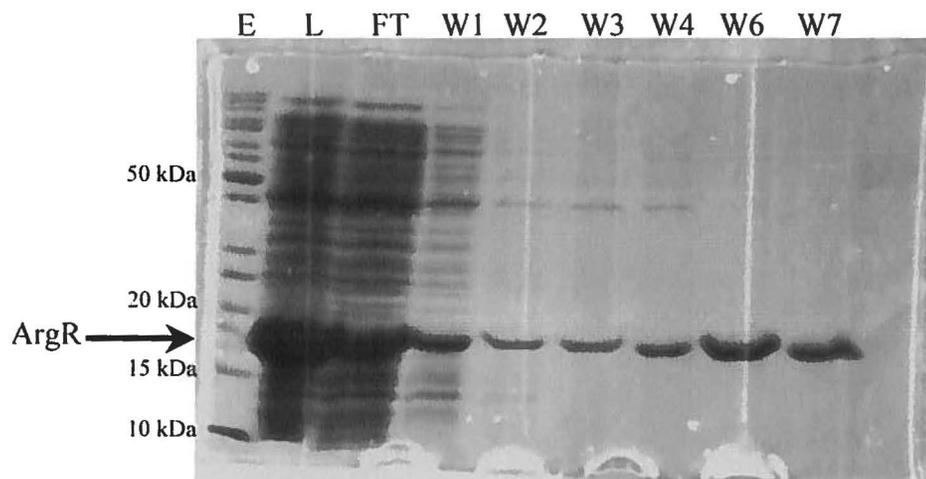


Figure 29 : Vérification de la purification de façon soluble de la protéine ArgR

Migration des fractions de lavage et d'élution sur un gel SDS-PAGE à 15%. Les puits « E » sont l'échelle de poids moléculaire (PageRuler Unstained Protein Ladder, Fermentas). Le puits « L » est le lysat bactérien. Le puits « FT » est l'effluent. Les puits « W » avec un chiffre correspondent aux différentes fractions de lavage de la colonne. Les puits « E » avec un chiffre correspondent aux différentes fractions d'élution. Seule la protéine ArgR est visible dans les fractions d'élution.

Comme contrôle de la fonctionnalité de l'expérience, le promoteur de *argR*, connu pour être régulé par cette protéine, a été inclus dans cette expérimentation. Une fixation de la protéine ArgR, représentée par un retard sur gel de l'ADN en complexe comparé à l'ADN libre, est observée dès 44 μM de protéine monomérique ajoutée. Ce retard est d'autant plus marqué à partir de 88 μM et 235 μM . Mais la bande correspondant à l'ADN libre ne disparaît pas, malgré ces fortes quantités de protéine (Fig. 30A). En comparaison avec le contrôle positif, soit le promoteur de *argR*, la quantité de protéine nécessaire pour provoquer un début de fixation sur l'ADN est plus faible (29 μM). De plus, à une quantité de 88 μM , plus aucune bande d'ADN libre n'est observable (Fig. 30B). D'autres essais ont été effectués avec des fragments d'ADN ne possédant pas de séquences ARG (séquences appartenant aux gènes *cysE* et *entB*). Or, avec ces fragments d'ADN, une fixation est observée, aux mêmes quantités de protéine ArgR que pour le promoteur de *Efo*. Donc la fixation de la protéine sur le promoteur de *efo* est possiblement non spécifique.

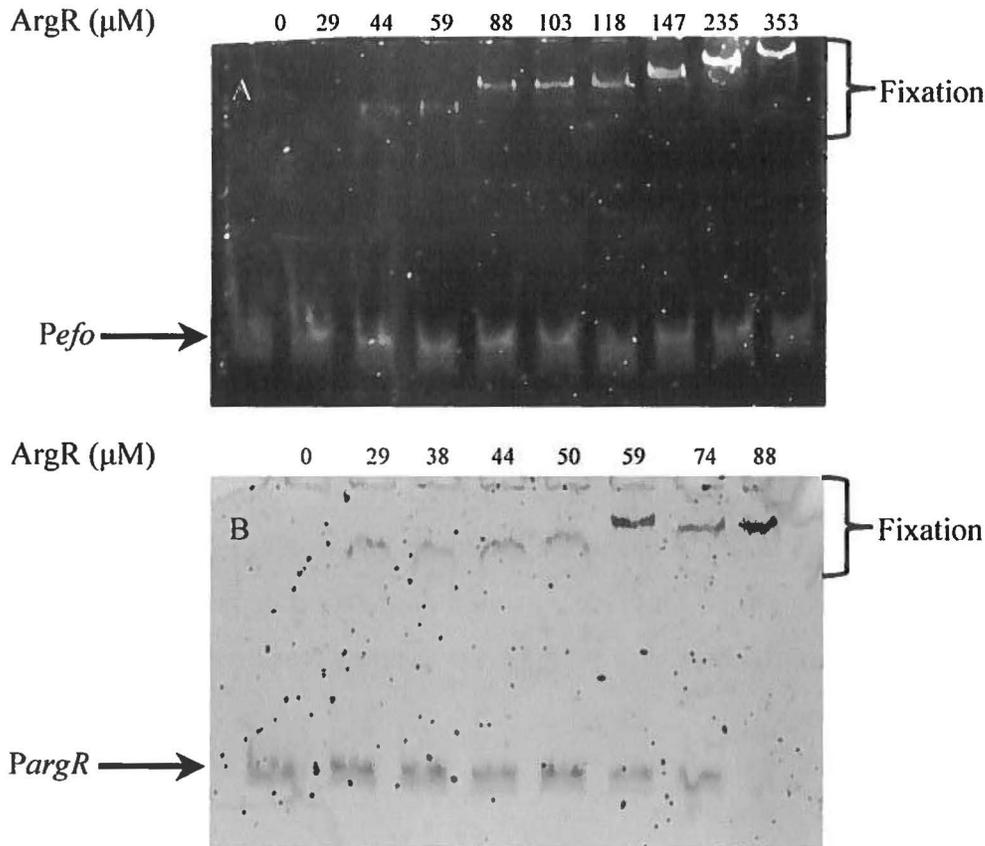


Figure 30 : Expérience de retard sur gel avec la protéine ArgR et les promoteurs de *efo* et de *argR*

Les retards sur gel ont été effectués sur des gels à 5 % de polyacrylamide. La concentration de monomère de la protéine ArgR, utilisée pour chaque puits (1 μL de chaque concentration), est indiquée au-dessus des photographies des gels. Les flèches noires indiquent la bande de l'ADN libre. Les accolades indiquent la zone où se trouvent les bandes correspondant au complexe ADN/ArgR. **A** : Révélation de la fixation de ArgR avec le promoteur de *efo*. **B** : Révélation de la fixation de ArgR avec le promoteur de *argR*.

Pour confirmer, ou infirmer, que la fixation de ArgR sur le promoteur de *efo* est spécifique, un autre retard sur gel a été entrepris avec un ADN compétiteur non spécifique (poly d(I-C)), ajouté en plus du promoteur *efo* biotinylé (voir le nouveau protocole dans le chapitre « 12 : Expériences de retards sur gels de polyacrylamide », dans la section « Matériels et Méthodes »). Cette nouvelle étude indique qu'en présence de cet ADN, la fixation de la protéine ArgR sur le promoteur est abolie et des fixations non spécifiques apparaissent (Fig. 31). Cela se traduit par des traînées le long des pistes de migration et l'absence des bandes de retard sur gel, synonyme de la formation du complexe ADN/ArgR.

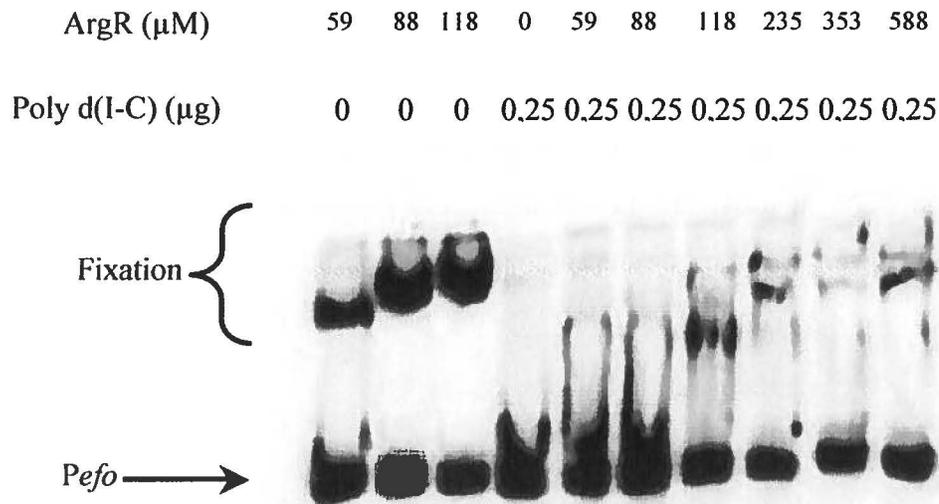


Figure 31 : Expérience de retard sur gel avec ArgR, le promoteur *efo* et de l'ADN compétiteur non spécifique
 Les retards sur gel ont été effectués sur des gels à 5 % de polyacrylamide. La concentration de monomère de la protéine ArgR, utilisée pour chaque puits (1 μL de chaque concentration), est indiquée au-dessus de la photographie du gel. En plus, la quantité d'ADN compétiteur non spécifique (Poly d(I-C)) est mentionnée sur la deuxième ligne au-dessus de la photographie. Dans les trois premiers puits, aucun ADN compétitif n'a été ajouté et il est observé des bandes de retard sur gel (contrôle positif). Mais, dès l'ajout de cet ADN, des trainées apparaissent le long des pistes de migration. C'est la marque de fixations non spécifique de ArgR.

9 Tests phénotypiques

Près de 50 % des souches testées possèdent les gènes *efo*, dont 58 % sans IS. Étant donné cette forte prévalence, la conservation du fimbria Efo doit être primordiale pour certains mécanismes de virulence ou pour la persistance dans un milieu non favorable à la croissance bactérienne. Pour déterminer le rôle du fimbria Efo, différents tests phénotypiques ont été entrepris.

9.1 Production de biofilm et le fimbria Efo

Le premier test est la détermination de l'implication du fimbria dans la formation des biofilms. Plusieurs conditions ont été testées : la température (30 °C, 37 °C et température pièce) et le type de milieu (riche ou minimum). Les souches utilisées sont les même que pour les photos de microscopie électronique et de gel pour les immunobuvarages, soit QT885 et QT1306, plus une souche de *Serratia liquefaciens* comme contrôle positif et enfin ECOR72 et ses mutants pour

les fimbriae Type 1 et Efo (Fig. 32). Aucune différence dans la production de biofilm n'a été notée, sauf entre la souche ECOR72 Δ *fim* et la souche ECOR72 Δ *fim* Δ *efo*, en milieu minimum et à température pièce, où une diminution de cette production a été observée de 1,2 fois (Fig. 32A) et lorsque les fimbriae Type 1 sont absents, à une température de 30 °C (Fig. 32B). La température de 42 °C a aussi fait l'objet d'une investigation. Mais, à cette température, la quantité produite de biofilm n'est pas mesurable, car trop faible.

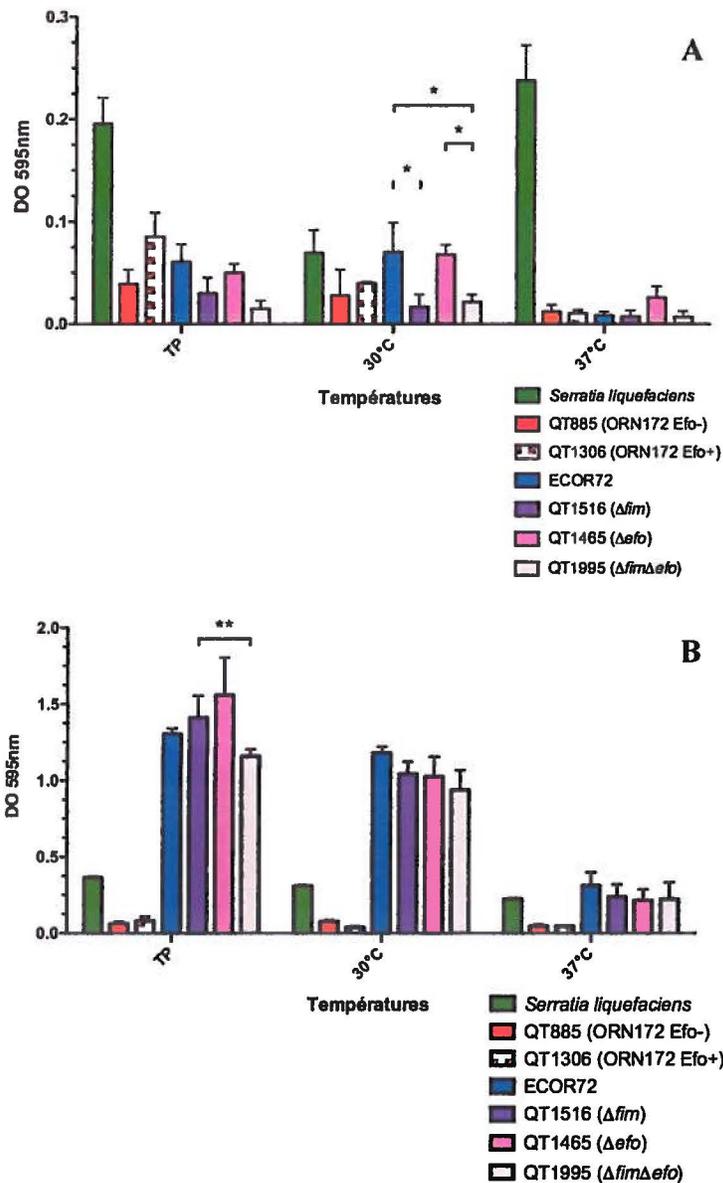


Figure 32 : Rôle du fimbria Efo dans la formation de biofilm

Évaluation de la production de biofilm, à trois températures différentes (TP pour température pièce), par mesure de l'absorption à une DO de 595 nm. La souche *Serratia liquefaciens* est un contrôle positif. A : Croissance des bactéries en milieu riche (LB). B : Croissance des bactéries en milieu minimum M9, supplémenté avec du glucose.

Aucune variation n'a été relevée entre ORN172 et QT1306. Les seules différences proviennent d'une diminution de production lorsque le fimbria Type 1 est muté, à 30 °C ($P < 0,05$) et entre QT1516 et QT1995, en milieu minimum et à la température pièce ($P < 0,01$).

9.2 Nage des bactéries et surexpression de *efoABCD*

Le second test consiste à mesurer la nage de la souche QT885 (ORN172 Efo-) et QT1306 (ORN172 Efo+), dans une gélose M9 semi-solide. Les deux concentrations d'agar utilisées sont 0,25 % ou 0,35 %. Le résultat obtenu est une diminution hautement significative de la nage des bactéries lorsqu'elles expriment les fimbriae Efo, pour les deux concentrations d'agar (Fig. 33). Le diamètre du halo bactérien est 1,6 fois inférieur quand la concentration d'agar est de 0,25 % et 1,3 fois inférieur pour 0,35 % d'agar ($P < 0,001$).

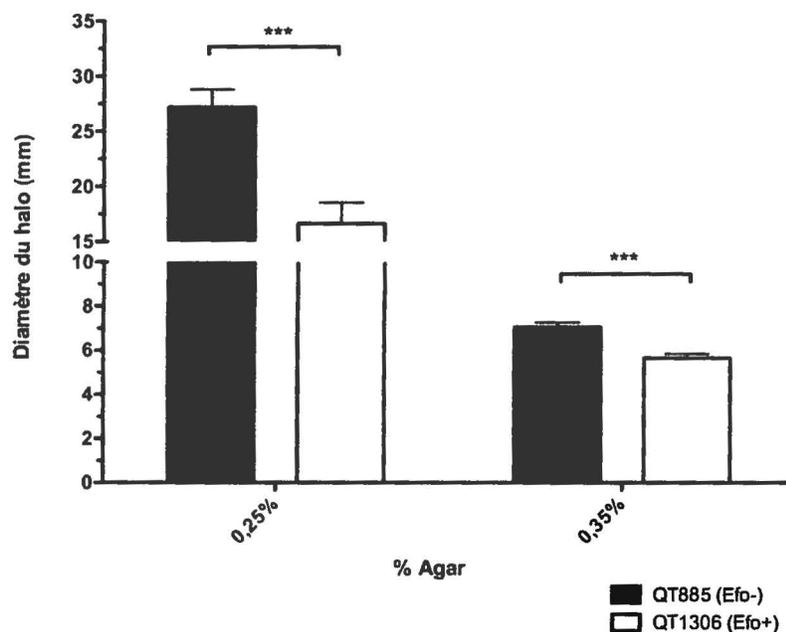


Figure 33 : Mesure de la nage des bactéries en fonction de la concentration en agar et de l'expression des fimbriae Efo

La nage des bactéries est mesurée selon le diamètre du halo qu'elles forment (27 réplicats). Deux souches sont utilisées : QT885 (ORN172 Efo-) et QT1306 (ORN172 Efo+). Une diminution significative ($P < 0,001$) est observée lorsque ORN172 exprime les fimbriae Efo et cela pour les deux concentrations d'agar 0,25 % et 0,35 %.

9.3 Rôle des fimbriae Efo dans la colonisation du tractus urinaire

Étant présent chez plusieurs souches urinaires, notamment chez ECOR72, il est intéressant de vérifier si les fimbriae Efo sont impliqués dans la colonisation du tractus urinaire. Pour cela, des co-infections chez des souris CBA/J ont été réalisées (modèle d'infection du tractus urinaire

murin). Les couples de souches utilisées sont : ECOR72/ Δfim , ECOR72/ Δefo et ECOR72/ $\Delta fim\Delta efo$. 48h après le début de l'infection du tractus urinaire, les souris sont nécropsiées et les vessies et les reins sont récupérés, pour mesurer le taux de bactéries contenu dans ces organes. La souche ECOR72 est un mutant Δlac , ceci afin de la différencier avec les souches mutantes pour les fimbriae, sur des géloses McConkey. Il a déjà été déterminé dans le laboratoire que cette mutation n'affecte pas la virulence des UPEC. Les tests ont été effectués avec cinq souris pour chaque couple de co-infection. Ce nombre est assez faible pour déterminer statistiquement des différences significatives, mais cela permet de présenter la tendance générale. Donc les résultats obtenus sont qu'aucune différence de colonisation n'est observée dans la vessie entre la souche sauvage et les mutants pour les fimbriae (Fig. 34A), malgré la tendance pour le mutant *fim*. Par contre dans les reins, une mutation des fimbriae Type 1 réduit l'aptitude du mutant à coloniser les reins et cela d'une façon significative ($P < 0,01$ pour le simple mutant et $P < 0,05$ pour le double mutant) (Fig. 34B). Aucune tendance ne se démarque lorsque le fimbria Efo est muté. Il est à noter que les différences observées ne proviennent pas d'un retard de croissance entre les souches (Annexe VI).

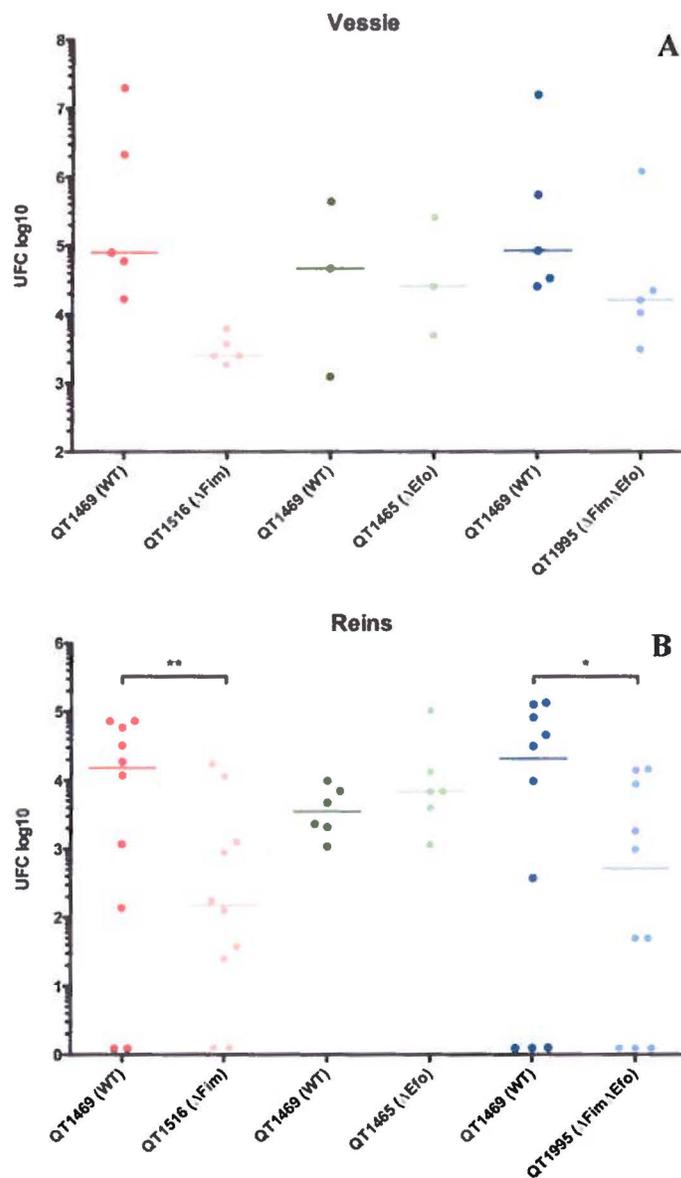


Figure 34 : Co-infections du tractus urinaire murin avec ECOR72 et ses mutants pour les fimbriae Type 1 et Efo

Le décompte bactérien est mesuré en log de dix d'unité formant colonie (UFC log10). Les couples de souches utilisées sont : QT1469/QT1516, QT1469/QT1465 et QT1469/QT1995. * signifie $P < 0,05$ et ** signifie $P < 0,01$

A : Décompte bactérien dans les vessies récoltées après les nécropsies. Aucune différence significative n'est observable. **B** : Décompte bactérien dans les reins récoltés après les nécropsies. Des différences significatives sont constatées lorsque le fimbria Type 1 est muté.

DISCUSSION

La bactérie *Escherichia coli* est présente chez la plupart des animaux. Elle fait partie intégrante de la flore commensale. Mais certaines souches d'*E. coli* peuvent provoquer des maladies infectieuses. Ceci est notamment dû à la grande plasticité de son génome, ce qui lui confère un important potentiel d'acquisition de facteurs de virulence. Afin de déterminer quels sont les gènes acquis par ces souches pathogènes, l'étude de Charles M Dozois *et al.* (2003) s'est basée sur la capture sélective de transcrits exprimés par la souche χ 7122 (APEC) *in vivo* et que l'on ne retrouve pas chez une souche non pathogène K-12 χ 289. Elle a permis de découvrir plusieurs fragments d'ADN, codant pour de potentiels fimbriae (4).

Le séquençage d'un de ces transcrits et de la zone autour, a révélé la présence de quatre cadres de lecture, entre les gènes *exuRT*. Ces gènes ont été nommés *efoABCD*, pour « *exu*-linked *fimbrial operon* ». Cependant, chez la souche χ 7122, le gène *efoC* est interrompu par une IS. Ce qui n'est pas le cas chez la souche ECOR72, qui possède les quatre gènes « intacts » et dans la région intergénique des gènes *exuRT*. De plus, chez tous les génomes séquencés et positifs pour *efo* (Fig. 22), ces gènes sont systématiquement entre *exuRT*. Une recherche des domaines conservés, parmi les protéines exprimées par les gènes *efoABCD*, a démontré une homologie avec les protéines du fimbria Cbl de *Burkholderia cepacia* et plus particulièrement avec le fimbria Tcf de *Salmonella enterica*. En effet, le domaine PFAM00345, pour la protéine chaperonne, n'est retrouvé que chez le fimbria Tcf et chez le fimbria Efo. Ceci est en accord avec le classement phylogénétique, à partir des séquences des gènes conservés des « *usher* » (Fig. 23 et référence (200)), qui catégorise le fimbria Efo dans la classe α , dont font partie les fimbriae Tcf et Cbl. De surcroît, le fimbria Efo est plus fortement relié phylogénétiquement avec ces fimbriae qu'avec d'autres types de fimbriae présents chez *Escherichia coli*, comme CFA/I et CS, qui sont véhiculés par des plasmides. De ce fait, il peut être conclu que le fimbria Efo est unique : il ne possède pas une homologie forte avec un seul fimbria et il ne provient pas d'une duplication de regroupement de gènes à l'intérieur même du génome d'*E. coli*. L'acquisition des gènes *efo* a dû se produire il y a suffisamment longtemps pour ne pas être une duplication simple des fimbriae Tcf et Cbl. En effet, le pourcentage de G+C est en moyenne de 44 %, il est donc différent de celui observé en général chez ces espèces bactériennes (≈ 60 % pour *Burkholderia* et ≈ 52 % pour *Salmonella*), mais assez proche dans le temps pour ne pas avoir eu le pourcentage

de bases G+C suffisamment modifié, afin de correspondre à ce que l'on retrouve communément dans le génome d'*Escherichia coli* (50,8 % (3)). L'insertion du cluster *efoABCD* entre les gènes *exuTR* ne modifie pas le site de fixation de la protéine ExuR sur son propre promoteur (365, 366). La recherche des domaines conservés a permis de déterminer des rôles hypothétiques pour chacune des protéines EfoABCD : EfoA serait la protéine chaperonne, EfoB la sous-unité majeure, EfoC l'*usher* et enfin EfoD serait la protéine initiatrice. Malgré des rôles hypothétiques, un clonage des gènes *efoABCD*, dans un vecteur à haut nombre de copies, a démontré que seuls ces quatre gènes sont nécessaires et suffisants pour la production de fimbriae chez une souche afimbriaire (Fig. 17, 18 et 19). La surexpression du fimbria Efo présente une organisation particulière des fimbriae : les structures semblent relier les bactéries entre elles et se rassembler en fuseaux. Néanmoins, cette disposition peut être due à une trop forte expression de Efo à la surface des bactéries. Cela amène la question suivante : est-ce que le fimbria Efo est impliqué dans l'auto-agrégation des bactéries ? Des expérimentations ont été menées dans ce sens par la mesure d'absorbance après 3h d'incubation à 4°C des souches QT1306 et QT885 et par l'observation de ces bactéries exprimant la GFP en microscopie électronique. Ces essais ont été concluants avec une augmentation significative de l'auto-agrégation chez QT1306 (Efo+). Néanmoins, lors de nouveaux tests avec, en plus, les souches QT3044 et QT3045, il n'a pas été possible de répéter les résultats précédents. La raison de ce problème n'a pas été découverte. Par ailleurs, *efoB* et *efoD* (la sous-unité majeure et la protéine initiatrice putative) sont indispensables pour une expression à la surface de la bactérie (Fig. 18). Cependant, en dépit de la délétion du gène *efoD*, quelques structures filamenteuses ont été observées. Plusieurs hypothèses peuvent expliquer cette observation : le clonage de *efoABCD* entraîne une production importante de protéines (Fig. 20). Pour pallier ce stress, la bactérie dégrade assez fortement les protéines, comme le montre la diminution de la bande correspondant à la protéine EfoB chez la souche QT3045. Mais cette dégradation n'est peut-être pas suffisante, donc l'exportation des protéines fimbriaires serait forcée, d'où l'observation de ces structures. Une autre possibilité est l'utilisation marginale d'une protéine initiatrice d'un autre système fimbriaire, comme vue précédemment avec d'autres fimbriae (203, 272, 273).

Le criblage des souches de la collection ECOR et de plusieurs PPEC démontre une association significative de *efo* avec le groupe phylogénétique B1 (Tableau 3). Ce groupe est composé essentiellement de souches commensales et de souches faisant partie des InPEC (90).

D'autre part, il est fortement relié avec le groupe A. Une représentation des souches, possédant *efo*, sur l'arbre phylogénétique de la collection ECOR (Fig. 21) montre un regroupement en clusters de ces souches au sein des groupes A et D. L'hypothèse serait une acquisition par l'ancêtre du groupe B1 du fimbria Efo par un transfert horizontal interespèces et ensuite une diffusion, toujours par ce type de transfert, parmi *Escherichia coli*. Par contre, aucune séquence répétée inversée, appartenant à des éléments transposables, n'a été retrouvée. Une grande proportion de souches APEC possède *efo* (51 %), mais dans la majorité des cas, *efoC* est interrompue par une IS (63 %) (Tableau 4). C'est une inactivation spontanée du fimbria Efo par un élément transposable. Compte tenu du fort taux d'interruptions de *efoC*, son inactivation ne doit pas influencer la virulence des APEC (ou elle doit être compensée par d'autres mutations) et même être privilégiée par la sélection naturelle (84). L'un des éléments qui peuvent agir dans ce sens est le système immunitaire de l'hôte (85, 89). Il reconnaîtrait, de façon trop importante, le fimbria Efo. En outre, *efo* est associé significativement avec le sérotype O78, qui est le plus répandu chez les APEC (52, 53, 57, 367). Ce résultat est en corrélation avec la relation du fimbria Efo et du groupe B1. En effet, les sérotypes O1 et O2 font plutôt partie du groupe phylogénétique B2, contrairement à O78 qui est plus associé avec le groupe B1 (54). Le fait d'avoir très peu de souches B1 chez les APEC et un nombre important de souches du groupe A, peut simplement s'expliquer par le fait que la méthode Clermont de détermination du groupe phylogénétique est déficiente pour les souches APEC (notamment pour l'amplification du fragment TspE4.C2). En conséquence, le fimbria Efo pourrait être utilisé comme marqueur du groupe phylogénétique B1. Selon l'étude de Walk *et al.* (2007), les bactéries du groupe B1 ont un avantage sélectif pour survivre dans l'environnement (104). La question qui se pose est : est-ce que le fimbria Efo fait partie des avantages sélectifs pour la survie à l'extérieur de l'hôte ?

L'arbre phylogénétique des souches, possédant *efoC* (ainsi que des autres gènes *efoABD*, données non présentées), a révélé la présence de trois allèles différents (Fig. 22). La souche ECOR72 possède l'allèle le plus représenté. Toutefois, cette surreprésentation peut provenir d'un biais dans le choix des souches séquencées. Les questions que cela soulève sont : est-ce que les trois allèles ont une même fonction ? Est-ce qu'ils reconnaissent des récepteurs légèrement différents, comme pour les fimbriae P (234, 235) ? Enfin, aucune association du fimbria Efo avec un pathotype en particulier n'est à noter.

Il a été déterminé que les gènes *efoBCD* sont sur un même ARNm (Fig. 15 et 16). De plus, un test préliminaire avec les amorces CMD1535/1536, sur le même échantillon d'ADNc présenté dans la figure 16, a apporté un résultat positif pour la présence des gènes *efoAB* sur le même transcrit (avec le témoin négatif sur l'ARN sans reverse transcription, qui ne produit aucune bande d'amplification). À partir de ces résultats, la conclusion qui en ressort est que les gènes *efoABCD* sont organisés en opéron. Donc, le promoteur, responsable de leur expression, se situe en amont du gène *efoA*. Une fusion transcriptionnelle de ce promoteur avec le gène *lacZ* et un clonage dans une souche sauvage pour l'opéron *efo* a permis de démontrer plusieurs régulations par des facteurs environnementaux, notamment la température, l'oxygénation et la source de carbone (Fig. 26 et 27). Ce genre de régulation a déjà été observée pour d'autres fimbriae chez *E. coli* (11, 243, 247-250, 255, 260). L'intégration des gènes *efoABCD*, entre les gènes du métabolisme des hexuronates *exuRT*, n'empêche pas l'utilisation de ces sucres par la bactérie (Fig. 27, galacturonate et glucuronate). La régulation par l'oxygène pourrait provenir du régulateur transcriptionnel ArcA, il médie la réponse de la bactérie lors d'un changement des conditions respiratoires (actif en condition de microaérobie et anaérobie (368, 369)). En effet, une séquence de régulation pour cette protéine a été détectée dans la séquence de 200 pb, en amont de *efoA*, où est le promoteur présumé (Annexe V). Mais des expérimentations complémentaires sont requises. De même, cette analyse du promoteur *efo* a révélé deux séquences de régulation pour la protéine ArgR. Or cette protéine a besoin de deux séquences séparées seulement de 2 à 3 pb. Donc, il existerait, potentiellement, deux autres séquences non détectées par le logiciel BPROM. Pour étudier l'influence de cette protéine sur l'expression de la β -galactosidase, des expérimentations ont été effectuées chez une souche mutante pour *argR* et chez une souche complémentée Tn7::*argR*. En présence d'arginine, l'expression est diminuée chez la souche sauvage et la souche complémentée (Fig. 28). C'est un résultat cohérent, étant donné que ArgR est activée par l'arginine et que c'est un régulateur négatif principalement. De surcroît, lorsque *argR* est supprimé, l'expression de la β -galactosidase augmente et d'autant plus lors de l'ajout d'arginine. Une explication serait que sans arginine dans le milieu minimum, le métabolisme de la bactérie est davantage ralenti et donc l'expression des gènes non essentiels est plus atténuée. Pour démontrer l'interaction directe de la protéine ArgR sur le promoteur du cluster de gènes *efo*, des expériences de retard sur gel ont été entreprises. Un début de fixation de ArgR est observé sur ce promoteur (Fig. 30) mais à une concentration trois fois plus élevée

(88 μM), comparée à un promoteur (*PargR*) connu pour être régulé par cette protéine (29 μM). De plus, même à de très fortes concentrations de protéine ArgR, il reste toujours de l'ADN non fixé, contrairement au témoin positif. Pour confirmer que cette fixation est spécifique, de l'ADN compétiteur non spécifique (poly d(I-C)) a été ajouté en plus du promoteur de *efo* et de la protéine ArgR (Fig. 31). Il en résulte une abolition de cette fixation et l'apparition de « trainées » le long des pistes de migration et l'absence des bandes de retard sur gel. Il est à noter une différence de hauteur avec les plus fortes concentrations, ceci est peut-être dû à la production d'agglomérats entre ArgR, qui forme des gros complexes hexamériques, et l'ADN. En conclusion de ces expériences, il est démontré que la protéine ArgR ne se fixe pas de façon spécifique avec le promoteur du cluster de gènes *efo*. Par conséquent, ArgR aurait un effet sur la régulation de l'expression de l'opéron *efo* de façon indirecte, telle que la répression d'un activateur. Par exemple, le gène *metY* est impliqué dans l'initiation de la biosynthèse des protéines et il est régulé négativement par ArgR et le complexe CAP. Or ce complexe est activé lors d'un stress par le manque de nutriments (provenant d'une augmentation de la production de AMPc), comme une culture en milieu M9. Par conséquent, en milieu minimum l'expression du gène *metY* est réduite et d'autant plus lorsque la protéine ArgR est active (304, 335). Ceci peut expliquer la diminution d'expression de la β -galactosidase observée en présence d'arginine, par l'intermédiaire de *metY*.

Les fimbriae Efo auraient potentiellement un rôle dans la formation des biofilms, en milieu minimum et à température pièce (Fig. 32B). Mais seulement lorsque l'opéron *fim* est muté, car le simple mutant pour *efo* ne révèle pas de différence avec la souche sauvage. Il est reconnu qu'un « cross-talk » existe entre les différents systèmes fimbriaires présents chez *E. coli* (243, 254, 261). C'est possiblement ce qui arrive avec le fimbria Efo : la seule mutation de l'opéron *fim* peut ne pas suffire pour exposer l'influence des fimbriae Efo dans la formation des biofilms, car d'autres fimbriae prennent le relai avant Efo. À cela s'ajoute qu'il n'est pas suffisant pour augmenter la production de biofilm chez une souche afimbriaire (Fig. 32).

La surexpression de *efoABCD* diminue significativement la nage des bactéries, à travers des géloses semi-solides, par encombrement spatial notamment (Fig. 33). Ce résultat est en adéquation avec les images de microscopie électronique, où aucun flagelle n'a été observé en présence des fimbriae Efo et avec des études précédentes sur les fimbriae Type 1 (253, 254). Un résultat similaire a été obtenu avec les fimbriae CFA/I, chez *Salmonella* (370).

Enfin, des co-infections *in vivo* de souris avec la souche sauvage de ECOR72 et les mutants pour les fimbriae Type 1 et Efo, n'ont montré aucun rôle de Efo dans la virulence de cette souche UPEC (Fig. 34). Ceci peut être dû, comme pour les expériences de biofilms, à la régulation croisée entre les fimbriae et donc un niveau de hiérarchisation plus faible pour Efo.

Perspectives

Grâce aux photographies par microscopie électronique et à l'immunobuvardage de type Western, il a été prouvé que les gènes *efoABCD* sont nécessaires et suffisants pour produire des structures filamenteuses à la surface d'une bactérie afimbriaire. Cependant, ce sont des résultats qui peuvent être considérés comme indirects. Des tests, effectués avec des anticorps secondaires marqués à l'or colloïdal, n'ont pas fonctionné, car l'anticorps primaire (tourné contre une séquence peptidique de la sous-unité majeure du fimbria Efo) ne reconnaît pas la forme native de la protéine. Par conséquent, la création d'un nouvel anticorps primaire serait à considérer, afin de faire un marquage direct des fimbriae en microscopie électronique.

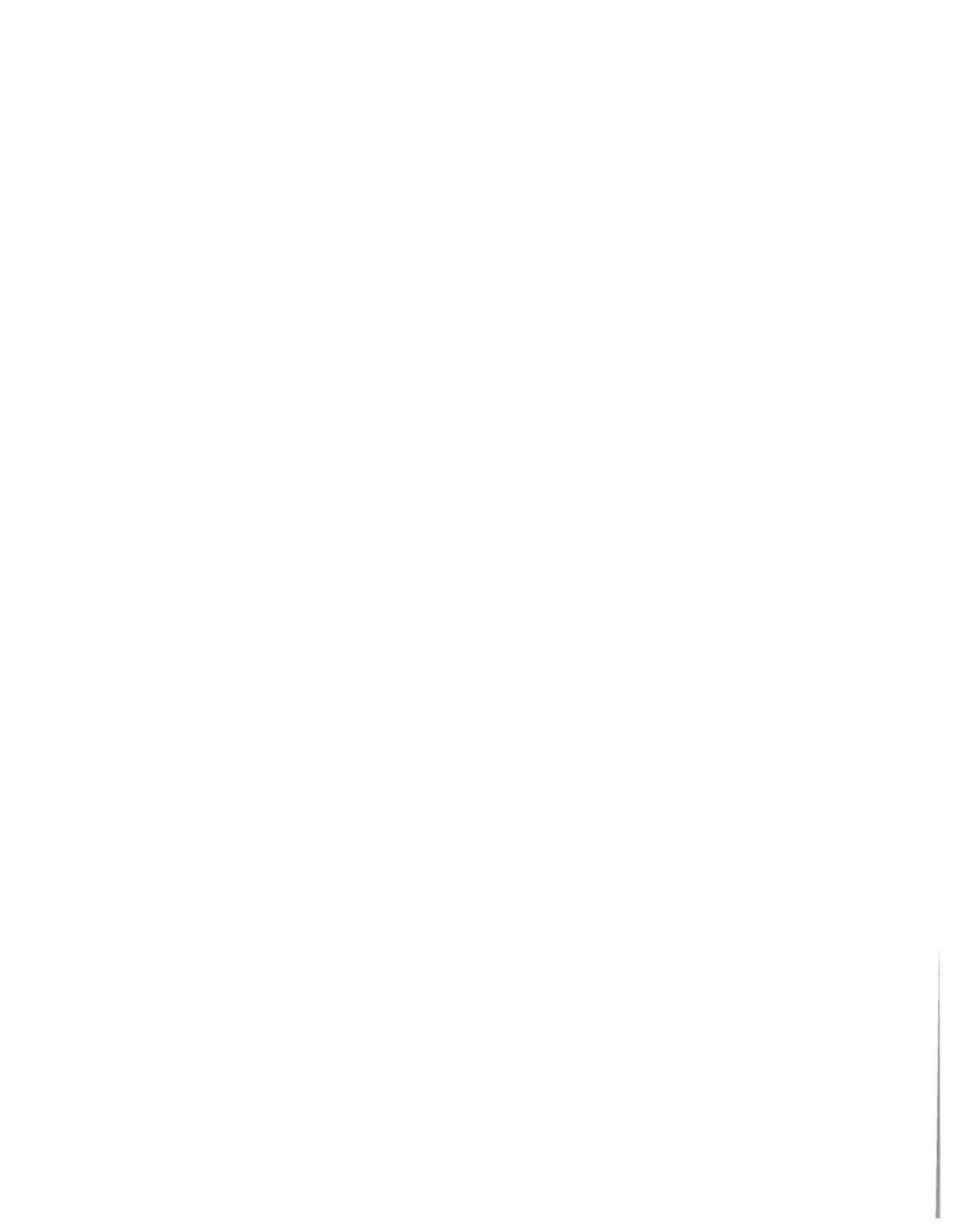
La forte proportion de souches qui possèdent le fimbria Efo porte à croire qu'il joue un rôle biologique. De plus, ce fimbria fait partie de la classe α des fimbriae. Or, tous les fimbriae de cette classe sont des facteurs de colonisation spécifiques à l'humain (200). Malheureusement, cette étude n'a pas permis de déterminer un phénotype pour les fimbriae Efo. Afin de pallier cela, d'autres expérimentations seront nécessaires, comme le criblage de souches d'*Escherichia coli*, pour en détecter une qui produira les fimbriae Efo de façon « naturelle » (ECOR72 ne produit pas le fimbria avec une assez grande intensité pour être détecté par immunobuvardage de type Western). Le pathotype de cette souche hypothétique donnerait une indication quant au rôle de Efo. De même, des tests d'adhérence et d'invasion sur différentes lignées cellulaires permettront d'avoir, potentiellement, un phénotype. En parallèle avec le criblage, la recherche d'un régulateur positif serait intéressante du fait que l'expression des fimbriae CS et CFA/I requiert la présence de la protéine Rns, codée par un gène présent sur un autre plasmide que les gènes de structures de ces fimbriae (371).

Dans le cadre de la régulation de l'expression de Efo, les résultats de β -galactosidase ont permis de montrer un effet de la protéine ArgR, mais d'une façon indirecte. Mais quelle est la relation entre métabolisme de l'arginine et la production des fimbriae ? En décomposant les

mécanismes moléculaires à l'origine de cet effet, un nouveau type de régulation des facteurs de virulence, passant par la protéine ArgR, pourrait être découvert.

Le fait qu'à ce jour seules les souches APEC (et des bactéries de l'espèce *Shigella*) possèdent le fimbria Efo interrompu par une IS est exceptionnel. Une des possibilités mentionnées est l'influence du système immunitaire de l'hôte sur la sélection de ces souches. Ajouter un plasmide codant pour Efo et faire des tests d'infections *in vivo* avec une de ces souches pourrait révéler une diminution de la virulence. Présentement, une seule bactérie de l'espèce *Shigella* possède les quatre gènes du fimbria Efo intacts, soit *Shigella sonnei* 53G. Une investigation plus poussée sur le rôle de Efo chez cette souche donnerait une autre portée quant à l'importance du fimbria Efo dans la virulence bactérienne.

Enfin, la caractérisation plus précise des sous-unités, qui composent le fimbria Efo et la détermination des récepteurs, qu'il va reconnaître, sont des informations primordiales dans l'étude du fimbria Efo.



CONCLUSION

Les facteurs de virulence sont ce qui va différencier les souches pathogènes d'*Escherichia coli*, des souches commensales. Ce sont aussi des cibles de choix pour contrer les maladies infectieuses provoquées par celles-ci. Dans le but de toujours découvrir de nouveaux facteurs de virulence et de comprendre la relation intime qui existe entre les souches pathogènes et l'hôte, un fimbria inédit a été caractérisé dans cette étude. Le fimbria Efo est codé par quatre gènes *efoABCD*, qui sont situés entre les gènes *exuRT* et qui sont sous forme d'opéron. C'est un fimbria de la classe α , ayant la méthode d'assemblage des fimbriae de la voie alternative des « chaperonnes-ushers ». Il a de fortes homologues avec les fimbriae Tcf de *Salmonella* et Cbl de *Burkholderia*. De plus, il est relié phylogénétiquement avec le groupe B1 en particulier, et le sérotype O78 chez les souches APEC. Par les études phylogénétiques, le fimbria Efo posséderait trois allèles différents. Plusieurs types de régulation ont été mis au jour, notamment par l'oxygénation du milieu, la température et la source de carbone. À cela s'ajoute qu'un effet indirect du régulateur du métabolisme de l'arginine, ArgR, a été mis en évidence dans cette étude. Enfin, le fimbria Efo a potentiellement un rôle dans la formation des biofilms, en association avec d'autres fimbriae et dans l'auto-agrégation des bactéries entre elles. L'étude du fimbria Efo n'en est qu'à ces débuts et plusieurs autres expérimentations seront nécessaires pour dévoiler toute la portée que possède le fimbria Efo dans la virulence des souches d'*Escherichia coli*.



RÉFÉRENCES

- 1 **Ullmann A (2011)** Chapter 1.1.2, *Escherichia coli* and the Emergence of Molecular Biology *In* A. Böck, R. Curtiss III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, and D. Ussery. EcoSal—*Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. <http://www.ecosal.org>. (ASM Press, Washington, DC)
- 2 **Bergthorsson U & Ochman H (1998)** Distribution of chromosome length variation in natural isolates of *Escherichia coli*. *Mol Biol Evol* 15(1):6-16.
- 3 **Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B & Shao Y (1997)** The complete genome sequence of *Escherichia coli* K-12. *Science* 277(5331):1453-1462.
- 4 **Dozois CM, Daigle F & Curtiss R (2003)** Identification of pathogen-specific and conserved genes expressed in vivo by an avian pathogenic *Escherichia coli* strain. *Proc Natl Acad Sci USA* 100(1):247-252.
- 5 **Hannan TJ, Totsika M, Mansfield KJ, Moore KH, Schembri MA & Hultgren SJ (2012)** Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection. *FEMS Microbiol Rev* 36(3):616-648.
- 6 **Nielubowicz GR & Mobley HLT (2010)** Host-pathogen interactions in urinary tract infection. *Nat Rev Urol* 7(8):430-441.
- 7 **Ochman H & Selander RK (1984)** Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* 157(2):690-693.
- 8 **Stenutz R, Weintraub A & Widmalm G (2006)** The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS Microbiol Rev* 30(3):382-403.
- 9 **Kaper JB, Nataro JP & Mobley HL (2004)** Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2(2):123-140.
- 10 **Dobrindt U (2005)** (Patho-)Genomics of *Escherichia coli*. *Int J Med Microbiol* 295(6-7):357-371.
- 11 **Johnson JR (1991)** Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev* 4(1):80-128.
- 12 **Tenaillon O, Skurnik D, Picard B & Denamur E (2010)** The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* 8(3):207-217.
- 13 **Maslow JN, Whittam TS, Gilks CF, Wilson RA, Mulligan ME, Adams KS & Arbeit RD (1995)** Clonal relationships among bloodstream isolates of *Escherichia coli*. *Infection and Immunity* 63(7):2409-2417.
- 14 **Herzer PJ, Inouye S, Inouye M & Whittam TS (1990)** Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *Journal of Bacteriology* 172(11):6175-6181.

- 15 **Whittam TS, Ochman H & Selander RK (1983)** Multilocus genetic structure in natural populations of *Escherichia coli*. *Proc Natl Acad Sci USA* 80(6):1751-1755.
- 16 **Manges AR, Johnson JR, Foxman B, O'Bryan TT, Fullerton KE & Riley LW (2001)** Widespread distribution of urinary tract infections caused by a multidrug-resistant *Escherichia coli* clonal group. *N Engl J Med* 345(14):1007-1013.
- 17 **Kauffmann F (1947)** The serology of the coli group. *J Immunol* 57(1):71-100.
- 18 **Harry L. T. Mobley MSDaECH (2009)** Chapter 8.6.1.3, Uropathogenic *Escherichia coli*. In A. Böck, R. Curtiss III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, and D. Ussery. *EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology*. <http://www.ecosal.org>. (ASM Press, Washington, DC)
- 19 **Whitfield C & Roberts IS (1999)** Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Mol Microbiol* 31(5):1307-1319.
- 20 **Nataro JP & Kaper JB (1998)** Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 11(1):142-201.
- 21 **Orskov I, Orskov F, Jann B & Jann K (1977)** Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. *Bacteriol Rev* 41(3):667-710.
- 22 **Schoenhals G & Whitfield C (1993)** Comparative analysis of flagellin sequences from *Escherichia coli* strains possessing serologically distinct flagellar filaments with a shared complex surface pattern. *J Bacteriol* 175(17):5395-5402.
- 23 **Pupo GM, Karaolis DK, Lan R & Reeves PR (1997)** Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and mdh sequence studies. *Infect Immun* 65(7):2685-2692.
- 24 **Whitfield C (2009)** Chapter 4.7.3, Structure and Assembly of *Escherichia coli* Capsules. In A. Böck, R. Curtiss III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, and D. Ussery. *EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology*. <http://www.ecosal.org>. (ASM Press, Washington, DC)
- 25 **Orskov I & Orskov F (1983)** Serology of *Escherichia coli* fimbriae. *Prog Allergy* 33:80-105.
- 26 **David G. Thanassi S-PN, Stephane Shu Kin So and Andreas J. Bäumlner (2007)** Chapter 2.4.2.1, Fimbriae: Classification and Biochemistry. In A. Böck, R. Curtiss III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, and D. Ussery. *EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology*. <http://www.ecosal.org>. (ASM Press, Washington, DC)
- 27 **Robins-Browne RM (1987)** Traditional enteropathogenic *Escherichia coli* of infantile diarrhea. *Rev Infect Dis* 9(1):28-53.
- 28 **Farthing MJG (2004)** Bugs and the gut: an unstable marriage. *Best Pract Res Clin Gastroenterol* 18(2):233-239.
- 29 **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.

- 30 **Taxt A, Aasland R, Sommerfelt H, Nataro J & Puntervoll P (2010)** Heat-stable enterotoxin of enterotoxigenic *Escherichia coli* as a vaccine target. *Infection and Immunity* 78(5):1824-1831.
- 31 **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.
- 32 **Bilge SS, Clausen CR, Lau W & Moseley SL (1989)** Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. *J Bacteriol* 171(8):4281-4289.
- 33 **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.
- 34 **Strober W (2011)** Adherent-invasive *E. coli* in Crohn disease: bacterial "agent provocateur". *J Clin Invest* 121(3):841-844.
- 35 **Simpson KW, Dogan B, Rishniw M, Goldstein RE, Klaessig S, McDonough PL, German AJ, Yates RM, Russell DG, Johnson SE, Berg DE, Harel J, Bruant G, McDonough SP & Schukken YH (2006)** Adherent and invasive *Escherichia coli* is associated with granulomatous colitis in boxer dogs. *Infection and Immunity* 74(8):4778-4792.
- 36 **Darfeuille-Michaud A (2002)** Adherent-invasive *Escherichia coli*: a putative new *E. coli* pathotype associated with Crohn's disease. *Int J Med Microbiol* 292(3-4):185-193.
- 37 **Rolhion N & Darfeuille-Michaud A (2007)** Adherent-invasive *Escherichia coli* in inflammatory bowel disease. *Inflamm Bowel Dis* 13(10):1277-1283.
- 38 **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).
- 39 **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).
- 40 **Russo TA & Johnson JR (2003)** Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes Infect* 5(5):449-456.
- 41 **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.
- 42 **Alteri CJ & Mobley HLT (2012)** *Escherichia coli* physiology and metabolism dictates adaptation to diverse host microenvironments. *Curr Opin Microbiol* 15(1):3-9.
- 43 **Johnson TJ, Wannemuehler Y, Johnson SJ, Stell AL, Doetkott C, Johnson JR, Kim KS, Spanjaard L & Nolan LK (2008)** Comparison of extraintestinal pathogenic

- Escherichia coli* strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens. *Appl Environ Microbiol* 74(22):7043-7050.
- 44 **Kim KS (2006)** Meningitis-Associated *Escherichia coli*. In A. Böck, R. Curtiss III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, and D. Ussery. *EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology*. <http://www.ecosal.org>. (ASM Press, Washington, DC)
- 45 **Unhanand M, Mustafa MM, McCracken GH & Nelson JD (1993)** Gram-negative enteric bacillary meningitis: a twenty-one-year experience. *J Pediatr* 122(1):15-21.
- 46 **Bonacorsi S & Bingen E (2005)** Molecular epidemiology of *Escherichia coli* causing neonatal meningitis. *Int J Med Microbiol* 295(6-7):373-381.
- 47 **Sussman M. (1997)**. *Escherichia Coli: Mechanisms of Virulence*. pp 3-25. Cambridge University Press, Cambridge
- 48 **Glode MP, Sutton A, Moxon ER & Robbins JB (1977)** Pathogenesis of neonatal *Escherichia coli* meningitis: induction of bacteremia and meningitis in infant rats fed *E. coli* K1. *Infection and Immunity* 16(1):75-80.
- 49 **Xie Y, Kim KJ & Kim KS (2004)** Current concepts on *Escherichia coli* K1 translocation of the blood-brain barrier. *FEMS Immunol Med Microbiol* 42(3):271-279.
- 50 **Kim KS, Itabashi H, Gemski P, Sadoff J, Warren RL & Cross AS (1992)** The K1 capsule is the critical determinant in the development of *Escherichia coli* meningitis in the rat. *J Clin Invest* 90(3):897-905.
- 51 **Gay CC (1965)** *Escherichia coli* and neonatal disease of calves. *Bacteriol Rev* 29:75-101.
- 52 **La Ragione RM & Woodward MJ (2002)** Virulence factors of *Escherichia coli* serotypes associated with avian colisepticaemia. *Res Vet Sci* 73(1):27-35.
- 53 **Dho-Moulin M & Fairbrother JM (1999)** Avian pathogenic *Escherichia coli* (APEC). *Vet Res* 30(2-3):299-316.
- 54 **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. *Journal of Bacteriology* 188(18):6449-6459.
- 55 **Caza M, Lépine F & Dozois CM (2011)** Secretion, but not overall synthesis, of catecholate siderophores contributes to virulence of extraintestinal pathogenic *Escherichia coli*. *Mol Microbiol* 80(1):266-282.
- 56 **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.
- 57 **Chérifi A, Contrepois M, Picard B, Goulet P, Orskov I & Orskov F (1994)** Clonal relationships among *Escherichia coli* serogroup O78 isolates from human and animal infections. *J Clin Microbiol* 32(5):1197-1202.

- 58 **Chanteloup NK, Porcheron G, Delaleu B, Germon P, Schouler C, Moulin-Schouleur M & Gilot P (2011)** The extra-intestinal avian pathogenic *Escherichia coli* strain BEN2908 invades avian and human epithelial cells and survives intracellularly. *Vet Microbiol* 147(3-4):435-439.
- 59 **Rodriguez-Siek KE, Giddings CW, Doetkott C, Johnson TJ, Fakhr MK & Nolan LK (2005)** Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis. *Microbiology (Reading, Engl)* 151(Pt 6):2097-2110.
- 60 **Fairbrother JM, Nadeau E & Gyles CL (2005)** *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. *Anim Health Res Rev* 6(1):17-39.
- 61 **Bélanger L, Garenaux A, Harel J, Boulianne M, Nadeau E & Dozois CM (2011)** *Escherichia coli* from animal reservoirs as a potential source of human extraintestinal pathogenic *E. coli*. *FEMS Immunol Med Microbiol* 62(1):1-10.
- 62 **Foxman B, Barlow R, D'Arcy H, Gillespie B & Sobel JD (2000)** Urinary tract infection: self-reported incidence and associated costs. *Ann Epidemiol* 10(8):509-515.
- 63 **Ronald A (2003)** The etiology of urinary tract infection: traditional and emerging pathogens. *Dis Mon* 49(2):71-82.
- 64 **Xie J, Foxman B, Zhang L & Marrs CF (2006)** Molecular epidemiologic identification of *Escherichia coli* genes that are potentially involved in movement of the organism from the intestinal tract to the vagina and bladder. *J Clin Microbiol* 44(7):2434-2441.
- 65 **Ikäheimo R, Siitonen A, Heiskanen T, Kärkkäinen U, Kuosmanen P, Lipponen P & Mäkelä PH (1996)** Recurrence of urinary tract infection in a primary care setting: analysis of a 1-year follow-up of 179 women. *Clin Infect Dis* 22(1):91-99.
- 66 **Foxman B & Frerichs RR (1985)** Epidemiology of urinary tract infection: I. Diaphragm use and sexual intercourse. *Am J Public Health* 75(11):1308-1313.
- 67 **Fihn SD, Boyko EJ, Normand EH, Chen CL, Grafton JR, Hunt M, Yarbro P, Scholes D & Stergachis A (1996)** Association between use of spermicide-coated condoms and *Escherichia coli* urinary tract infection in young women. *Am J Epidemiol* 144(5):512-520.
- 68 **Foxman B (2002)** Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *Am J Med* 113 Suppl 1A:5S-13S.
- 69 **Foxman B (1990)** Recurring urinary tract infection: incidence and risk factors. *Am J Public Health* 80(3):331-333.
- 70 **Scholes D, Hooton TM, Roberts PL, Stapleton AE, Gupta K & Stamm WE (2000)** Risk factors for recurrent urinary tract infection in young women. *J Infect Dis* 182(4):1177-1182.
- 71 **Nicolle LE (2003)** Urinary tract infection: traditional pharmacologic therapies. *Dis Mon* 49(2):111-128.
- 72 **Schaeffer AJ (2001)** What do we know about the urinary tract infection-prone individual? *J Infect Dis* 183 Suppl 1:S66-69.

- 73 **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.
- 74 **Lane MC, Alteri CJ, Smith SN & Mobley HLT (2007)** Expression of flagella is coincident with uropathogenic *Escherichia coli* ascension to the upper urinary tract. *Proc Natl Acad Sci USA* 104(42):16669-16674.
- 75 **Dhakai BK, Kulesus RR & Mulvey MA (2008)** Mechanisms and consequences of bladder cell invasion by uropathogenic *Escherichia coli*. *Eur J Clin Invest* 38 Suppl 2:2-11.
- 76 **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. *Proc Natl Acad Sci USA* 101(5):1333-1338.
- 77 **Mulvey MA, Schilling JD & Hultgren SJ (2001)** Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infection and Immunity* 69(7):4572-4579.
- 78 **McTaggart LA, Rigby RC & Elliott TS (1990)** The pathogenesis of urinary tract infections associated with *Escherichia coli*, *Staphylococcus saprophyticus* and *S. epidermidis*. *J Med Microbiol* 32(2):135-141.
- 79 **Rosen DA, Hooton TM, Stamm WE, Humphrey PA & Hultgren SJ (2007)** Detection of intracellular bacterial communities in human urinary tract infection. *PLoS Med* 4(12):e329.
- 80 **Martinez JJ, Mulvey MA, Schilling JD, Pinkner JS & Hultgren SJ (2000)** Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO J* 19(12):2803-2812.
- 81 **Hunstad DA & Justice SS (2010)** Intracellular lifestyles and immune evasion strategies of uropathogenic *Escherichia coli*. *Annu Rev Microbiol* 64:203-221.
- 82 **Colgan R, Nicolle LE, McGlone A & Hooton TM (2006)** Asymptomatic bacteriuria in adults. *Am Fam Physician* 74(6):985-990.
- 83 **Raz R (2003)** Asymptomatic bacteriuria. Clinical significance and management. *Int J Antimicrob Agents* 22 Suppl 2:45-47.
- 84 **Zdziarski J, Svanborg C, Wullt B, Hacker J & Dobrindt U (2008)** Molecular basis of commensalism in the urinary tract: low virulence or virulence attenuation? *Infection and Immunity* 76(2):695-703.
- 85 **Roos V, Nielsen EM & Klemm P (2006)** Asymptomatic bacteriuria *Escherichia coli* strains: adhesins, growth and competition. *FEMS Microbiol Lett* 262(1):22-30.
- 86 **Roos V, Schembri MA, Ulett GC & Klemm P (2006)** Asymptomatic bacteriuria *Escherichia coli* strain 83972 carries mutations in the *foc* locus and is unable to express F1C fimbriae. *Microbiology (Reading, Engl)* 152(Pt 6):1799-1806.
- 87 **Klemm P, Roos V, Ulett GC, Svanborg C & Schembri MA (2006)** Molecular characterization of the *Escherichia coli* asymptomatic bacteriuria strain 83972: the taming of a pathogen. *Infection and Immunity* 74(1):781-785.

- 88 **Nagy G, Dobrindt U, Schneider G, Khan AS, Hacker J & Emödy L (2002)** Loss of regulatory protein RfaH attenuates virulence of uropathogenic *Escherichia coli*. *Infection and Immunity* 70(8):4406-4413.
- 89 **Salvador E, Wagenlehner F, Köhler C-D, Mellmann A, Hacker J, Svanborg C & Dobrindt U (2012)** Comparison of asymptomatic bacteriuria *Escherichia coli* isolates from healthy individuals versus those from hospital patients shows that long-term bladder colonization selects for attenuated virulence phenotypes. *Infection and Immunity* 80(2):668-678.
- 90 **Zhang Y & Lin K (2012)** A phylogenomic analysis of *Escherichia coli* / *Shigella* group: implications of genomic features associated with pathogenicity and ecological adaptation. *BMC Evol Biol* 12:174.
- 91 **Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MCJ, Ochman H & Achtman M (2006)** Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* 60(5):1136-1151.
- 92 **Zhang L, Foxman B & Marrs C (2002)** Both urinary and rectal *Escherichia coli* isolates are dominated by strains of phylogenetic group B2. *J Clin Microbiol* 40(11):3951-3955.
- 93 **Johnson JR, O'Bryan TT, Kuskowski M & Maslow JN (2001)** Ongoing horizontal and vertical transmission of virulence genes and *papA* alleles among *Escherichia coli* blood isolates from patients with diverse-source bacteremia. *Infection and Immunity* 69(9):5363-5374.
- 94 **Johnson JR, Delavari P, Kuskowski M & Stell AL (2001)** Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. *J Infect Dis* 183(1):78-88.
- 95 **Johnson JR & Stell AL (2000)** Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J Infect Dis* 181(1):261-272.
- 96 **Picard B, Garcia JS, Gouriou S, Duriez P, Brahim N, Bingen E, Elion J & Denamur E (1999)** The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* 67(2):546-553.
- 97 **Lecointre G, Rachdi L, Darlu P & Denamur E (1998)** *Escherichia coli* molecular phylogeny using the incongruence length difference test. *Mol Biol Evol* 15(12):1685-1695.
- 98 **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.
- 99 **Bingen E, Picard B, Brahim N, Mathy S, Desjardins P, Elion J & Denamur E (1998)** Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *J Infect Dis* 177(3):642-650.

- 100 **Selander RK, Caugant DA, Ochman H, Musser JM, Gilmour MN & Whittam TS (1986)** Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Appl Environ Microbiol* 51(5):873-884.
- 101 **Carlos C, Pires MM, Stoppe NC, Hachich EM, Sato MIZ, Gomes TAT, Amaral LA & Ottoboni LMM (2010)** *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. *BMC Microbiol* 10:161.
- 102 **Baldy-Chudzik K, Mackiewicz P & Stosik M (2008)** Phylogenetic background, virulence gene profiles, and genomic diversity in commensal *Escherichia coli* isolated from ten mammal species living in one zoo. *Vet Microbiol* 131(1-2):173-184.
- 103 **Gordon DM & Cowling A (2003)** The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology (Reading, Engl)* 149(Pt 12):3575-3586.
- 104 **Walk ST, Alm EW, Calhoun LM, Mladonicky JM & Whittam TS (2007)** Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environ Microbiol* 9(9):2274-2288.
- 105 **Clermont O, Bonacorsi S & Bingen E (2000)** Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* 66(10):4555-4558.
- 106 **Lee J, Hiibel SR, Reardon KF & Wood TK (2010)** Identification of stress-related proteins in *Escherichia coli* using the pollutant cis-dichloroethylene. *J Appl Microbiol* 108(6):2088-2102.
- 107 **Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M & Spratt BG (1998)** Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci USA* 95(6):3140-3145.
- 108 **Noller AC, McEllistrem MC, Stine OC, Morris JG, Boxrud DJ, Dixon B & Harrison LH (2003)** Multilocus sequence typing reveals a lack of diversity among *Escherichia coli* O157:H7 isolates that are distinct by pulsed-field gel electrophoresis. *J Clin Microbiol* 41(2):675-679.
- 109 **Robertson GA, Thiruvenkataswamy V, Shilling H, Price EP, Huygens F, Henskens FA & Giffard PM (2004)** Identification and interrogation of highly informative single nucleotide polymorphism sets defined by bacterial multilocus sequence typing databases. *J Med Microbiol* 53(Pt 1):35-45.
- 110 **Weissman SJ, Moseley SL, Dykhuizen DE & Sokurenko EV (2003)** Enterobacterial adhesins and the case for studying SNPs in bacteria. *Trends Microbiol* 11(3):115-117.
- 111 **Moorhead SM, Dykes GA & Cursons RT (2003)** An SNP-based PCR assay to differentiate between *Listeria monocytogenes* lineages derived from phylogenetic analysis of the *sigB* gene. *J Microbiol Methods* 55(2):425-432.
- 112 **Waters VL & Crosa JH (1991)** Colicin V virulence plasmids. *Microbiol Rev* 55(3):437-450.

- 113 **Clancy J & Savage DC (1981)** Another Colicin V phenotype: in vitro adhesion of *Escherichia coli* to mouse intestinal epithelium. *Infection and Immunity* 32(1):343-352.
- 114 **Braun V (1981)** *Escherichia coli* cells containing the plasmid ColV produce the iron ionophore aerobactin. *FEMS Microbiol Lett* 11(4):225-228.
- 115 **Falkow S (1988)** Molecular Koch's postulates applied to microbial pathogenicity. *Rev Infect Dis* 10 Suppl 2:S274-276.
- 116 **Lawrence JG & Ochman H (1998)** Molecular archaeology of the *Escherichia coli* genome. *Proc Natl Acad Sci USA* 95(16):9413-9417.
- 117 **Lawrence JG & Ochman H (1997)** Amelioration of bacterial genomes: rates of change and exchange. *J Mol Evol* 44(4):383-397.
- 118 **Lan R & Reeves PR (1996)** Gene transfer is a major factor in bacterial evolution. *Mol Biol Evol* 13(1):47-55.
- 119 **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.
- 120 **Dubey GP & Ben-Yehuda S (2011)** Intercellular nanotubes mediate bacterial communication. *Cell* 144(4):590-600.
- 121 **Hacker J & Kaper JB (2000)** Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* 54:641-679.
- 122 **Dozois CM & Curtiss R (1999)** Pathogenic diversity of *Escherichia coli* and the emergence of 'exotic' islands in the gene stream. *Vet Res* 30(2-3):157-179.
- 123 **Siguier P, Filée J & Chandler M (2006)** Insertion sequences in prokaryotic genomes. *Curr Opin Microbiol* 9(5):526-531.
- 124 **Mahillon J, Léonard C & Chandler M (1999)** IS elements as constituents of bacterial genomes. *Res Microbiol* 150(9-10):675-687.
- 125 **Köster W (2001)** ABC transporter-mediated uptake of iron, siderophores, heme and vitamin B12. *Res Microbiol* 152(3-4):291-301.
- 126 **Garénaux 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.
- 127 **Wiles TJ, Kulesus RR & Mulvey MA (2008)** Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Exp Mol Pathol* 85(1):11-19.
- 128 **Andrews SC, Robinson AK & Rodríguez-Quñones F (2003)** Bacterial iron homeostasis. *FEMS Microbiology Reviews* 27(2-3):215-237.
- 129 **Faraldo-Gómez JD & Sansom MSP (2003)** Acquisition of siderophores in gram-negative bacteria. *Nat Rev Mol Cell Biol* 4(2):105-116.
- 130 **O'Brien IG, Cox GB & Gibson F (1970)** Biologically active compounds containing 2,3-dihydroxybenzoic acid and serine formed by *Escherichia coli*. *Biochim Biophys Acta* 201(3):453-460.

- 131 **Harris WR, Carrano CJ, Cooper SR, Sofen SR, Avdeef AE, McArdle JV & Raymond KN (1979)** Coordination chemistry of microbial iron transport compounds. 19. Stability constants and electrochemical behavior of ferric enterobactin and model complexes. *Journal of the American Chemical Society* 101(20):6097-6104.
- 132 **Garcia EC, Brumbaugh AR & Mobley HLT (2011)** Redundancy and specificity of *Escherichia coli* iron acquisition systems during urinary tract infection. *Infection and Immunity* 79(3):1225-1235.
- 133 **Goetz DH, Holmes MA, Borregaard N, Bluhm ME, Raymond KN & Strong RK (2002)** The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition. *Mol Cell* 10(5):1033-1043.
- 134 **Hantke K, Nicholson G, Rabsch W & Winkelmann G (2003)** Salmochelins, siderophores of *Salmonella enterica* and uropathogenic *Escherichia coli* strains, are recognized by the outer membrane receptor IroN. *Proc Natl Acad Sci USA* 100(7):3677-3682.
- 135 **Smith KD (2007)** Iron metabolism at the host pathogen interface: lipocalin 2 and the pathogen-associated *iroA* gene cluster. *Int J Biochem Cell Biol* 39(10):1776-1780.
- 136 **Fischbach MA, Lin H, Zhou L, Yu Y, Abergel RJ, Liu DR, Raymond KN, Wanner BL, Strong RK, Walsh CT, Aderem A & Smith KD (2006)** The pathogen-associated *iroA* gene cluster mediates bacterial evasion of lipocalin 2. *Proc Natl Acad Sci USA* 103(44):16502-16507.
- 137 **Lin H, Fischbach MA, Liu DR & Walsh CT (2005)** In vitro characterization of salmochelin and enterobactin trilactone hydrolases IroD, IroE, and Fes. *J Am Chem Soc* 127(31):11075-11084.
- 138 **Bister B, Bischoff D, Nicholson GJ, Valdebenito M, Schneider K, Winkelmann G, Hantke K & Süssmuth RD (2004)** The structure of salmochelins: C-glucosylated enterobactins of *Salmonella enterica*. *Biomaterials* 17(4):471-481.
- 139 **Nègre VL, Bonacorsi S, Schubert S, Bidet P, Nassif X & Bingen E (2004)** The siderophore receptor IroN, but not the high-pathogenicity island or the hemin receptor ChuA, contributes to the bacteremic step of *Escherichia coli* neonatal meningitis. *Infection and Immunity* 72(2):1216-1220.
- 140 **Bonacorsi S, Clermont O, Houdouin V, Cordevant C, Brahimi N, Marecat A, Tinsley C, Nassif X, Lange M & Bingen E (2003)** Molecular analysis and experimental virulence of French and North American *Escherichia coli* neonatal meningitis isolates: identification of a new virulent clone. *J Infect Dis* 187(12):1895-1906.
- 141 **Russo TA, McFadden CD, Carlino-MacDonald UB, Beanan JM, Barnard TJ & Johnson JR (2002)** IroN functions as a siderophore receptor and is a urovirulence factor in an extraintestinal pathogenic isolate of *Escherichia coli*. *Infection and Immunity* 70(12):7156-7160.
- 142 **Johnson JR, Oswald E, O'Bryan TT, Kuskowski MA & Spanjaard L (2002)** Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in the Netherlands. *J Infect Dis* 185(6):774-784.

- 143 **Williams PH & Warner PJ (1980)** ColV plasmid-mediated, colicin V-independent iron uptake system of invasive strains of *Escherichia coli*. *Infection and Immunity* 29(2):411-416.
- 144 **Linggood MA, Roberts M, Ford S, Parry SH & Williams PH (1987)** Incidence of the aerobactin iron uptake system among *Escherichia coli* isolates from infections of farm animals. *J Gen Microbiol* 133(4):835-842.
- 145 **Valvano MA & Crosa JH (1984)** Aerobactin iron transport genes commonly encoded by certain ColV plasmids occur in the chromosome of a human invasive strain of *Escherichia coli* K1. *Infection and Immunity* 46(1):159-167.
- 146 **Bagg A & Neilands JB (1987)** Molecular mechanism of regulation of siderophore-mediated iron assimilation. *Microbiol Rev* 51(4):509-518.
- 147 **Montgomerie JZ, Bindereif A, Neilands JB, Kalmanson GM & Guze LB (1984)** Association of hydroxamate siderophore (aerobactin) with *Escherichia coli* isolated from patients with bacteremia. *Infection and Immunity* 46(3):835-838.
- 148 **Lafont JP, Dho M, D'Hauteville HM, Bree A & Sansonetti PJ (1987)** Presence and expression of aerobactin genes in virulent avian strains of *Escherichia coli*. *Infection and Immunity* 55(1):193-197.
- 149 **Der Vartanian M, Jaffeux B, Contrepois M, Chavarot M, Girardeau JP, Bertin Y & Martin C (1992)** Role of aerobactin in systemic spread of an opportunistic strain of *Escherichia coli* from the intestinal tract of gnotobiotic lambs. *Infection and Immunity* 60(7):2800-2807.
- 150 **Torres AG, Redford P, Welch RA & Payne SM (2001)** TonB-dependent systems of uropathogenic *Escherichia coli*: aerobactin and heme transport and TonB are required for virulence in the mouse. *Infection and Immunity* 69(10):6179-6185.
- 151 **Heesemann J, Hantke K, Vocke T, Saken E, Rakin A, Stojiljkovic I & Berner R (1993)** Virulence of *Yersinia enterocolitica* is closely associated with siderophore production, expression of an iron-repressible outer membrane polypeptide of 65,000 Da and pesticin sensitivity. *Mol Microbiol* 8(2):397-408.
- 152 **Miller MC, Fetherston JD, Pickett CL, Bobrov AG, Weaver RH, DeMoll E & Perry RD (2010)** Reduced synthesis of the Ybt siderophore or production of aberrant Ybt-like molecules activates transcription of yersiniabactin genes in *Yersinia pestis*. *Microbiology (Reading, Engl)* 156(Pt 7):2226-2238.
- 153 **Paauw A, Leverstein-van Hall MA, van Kessel KPM, Verhoef J & Fluit AC (2009)** Yersiniabactin reduces the respiratory oxidative stress response of innate immune cells. *PLoS ONE* 4(12):e8240.
- 154 **Hancock V, Ferrières L & Klemm P (2008)** The ferric yersiniabactin uptake receptor FyuA is required for efficient biofilm formation by urinary tract infectious *Escherichia coli* in human urine. *Microbiology (Reading, Engl)* 154(Pt 1):167-175.
- 155 **Macnab RM (2003)** How bacteria assemble flagella. *Annu Rev Microbiol* 57:77-100.

- 156 **La Ragione RM, Sayers AR & Woodward MJ (2000)** The role of fimbriae and flagella in the colonization, invasion and persistence of *Escherichia coli* O78:K80 in the day-old-chick model. *Epidemiol Infect* 124(3):351-363.
- 157 **La Ragione RM, Cooley WA & 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. *J Med Microbiol* 49(4):327-338.
- 158 **Freter R, O'Brien PC & Macsai MS (1981)** Role of chemotaxis in the association of motile bacteria with intestinal mucosa: in vivo studies. *Infection and Immunity* 34(1):234-240.
- 159 **Freter R, Allweiss B, O'Brien PC, Halstead SA & Macsai MS (1981)** Role of chemotaxis in the association of motile bacteria with intestinal mucosa: in vitro studies. *Infection and Immunity* 34(1):241-249.
- 160 **Parthasarathy G, Yao Y & Kim KS (2007)** Flagella promote *Escherichia coli* K1 association with and invasion of human brain microvascular endothelial cells. *Infection and Immunity* 75(6):2937-2945.
- 161 **Sampaio SCF, Gomes TAT, Pichon C, du Merle L, Guadagnini S, Abe CM, Sampaio JLM & Le Bouguéne C (2009)** The flagella of an atypical enteropathogenic *Escherichia coli* strain are required for efficient interaction with and stimulation of interleukin-8 production by enterocytes in vitro. *Infection and Immunity* 77(10):4406-4413.
- 162 **Wright KJ, Seed PC & Hultgren SJ (2005)** Uropathogenic *Escherichia coli* flagella aid in efficient urinary tract colonization. *Infection and Immunity* 73(11):7657-7668.
- 163 **Lane MC, Lockett V, Monterosso G, Lamphier D, Weinert J, Hebel JR, Johnson DE & Mobley HLT (2005)** Role of motility in the colonization of uropathogenic *Escherichia coli* in the urinary tract. *Infection and Immunity* 73(11):7644-7656.
- 164 **Beutin L (1991)** The different hemolysins of *Escherichia coli*. *Med Microbiol Immunol* 180(4):167-182.
- 165 **Reingold J, Starr N, Maurer J & Lee MD (1999)** Identification of a new *Escherichia coli* She haemolysin homolog in avian *E. coli*. *Veterinary Microbiology* 66(2):125-134.
- 166 **Blanco JE, Blanco M, Mora A & Blanco J (1997)** Production of toxins (enterotoxins, verotoxins, and necrotoxins) and colicins by *Escherichia coli* strains isolated from septicemic and healthy chickens: relationship with in vivo pathogenicity. *J Clin Microbiol* 35(11):2953-2957.
- 167 **Bhakdi S, Mackman N, Menestrina G, Gray L, Hugo F, Seeger W & Holland IB (1988)** The hemolysin of *Escherichia coli*. *Eur J Epidemiol* 4(2):135-143.
- 168 **Boehm DF, Welch RA & Snyder IS (1990)** Calcium is required for binding of *Escherichia coli* hemolysin (HlyA) to erythrocyte membranes. *Infection and Immunity* 58(6):1951-1958.
- 169 **Short EC & Kurtz HJ (1971)** Properties of the Hemolytic Activities of *Escherichia coli*. *Infection and Immunity* 3(5):678-687.

- 170 **Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC & Ala'Aldeen D (2004)** Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 68(4):692-744.
- 171 **Bhakdi S, Mackman N, Nicaud JM & Holland IB (1986)** *Escherichia coli* hemolysin may damage target cell membranes by generating transmembrane pores. *Infection and Immunity* 52(1):63-69.
- 172 **Smith YC, Rasmussen SB, Grande KK, Conran RM & O'Brien AD (2008)** Hemolysin of uropathogenic *Escherichia coli* evokes extensive shedding of the uroepithelium and hemorrhage in bladder tissue within the first 24 hours after intraurethral inoculation of mice. *Infection and Immunity* 76(7):2978-2990.
- 173 **O'Hanley P, Lalonde G & Ji G (1991)** Alpha-hemolysin contributes to the pathogenicity of piliated digalactoside-binding *Escherichia coli* in the kidney: efficacy of an alpha-hemolysin vaccine in preventing renal injury in the BALB/c mouse model of pyelonephritis. *Infection and Immunity* 59(3):1153-1161.
- 174 **Caprioli A, Falbo V, Roda LG, Ruggeri FM & Zona C (1983)** Partial purification and characterization of an *Escherichia coli* toxic factor that induces morphological cell alterations. *Infection and Immunity* 39(3):1300-1306.
- 175 **Davis JM, Carvalho HM, Rasmussen SB & O'Brien AD (2006)** Cytotoxic necrotizing factor type 1 delivered by outer membrane vesicles of uropathogenic *Escherichia coli* attenuates polymorphonuclear leukocyte antimicrobial activity and chemotaxis. *Infection and Immunity* 74(8):4401-4408.
- 176 **Lemonnier M, Landraud L & Lemichez E (2007)** Rho GTPase-activating bacterial toxins: from bacterial virulence regulation to eukaryotic cell biology. *FEMS Microbiology Reviews* 31(5):515-534.
- 177 **Falbo V, Pace T, Picci L, Pizzi E & Caprioli A (1993)** Isolation and nucleotide sequence of the gene encoding cytotoxic necrotizing factor 1 of *Escherichia coli*. *Infection and Immunity* 61(11):4909-4914.
- 178 **Landraud L, Gauthier M, Fosse T & Boquet P (2000)** Frequency of *Escherichia coli* strains producing the cytotoxic necrotizing factor (CNF1) in nosocomial urinary tract infections. *Lett Appl Microbiol* 30(3):213-216.
- 179 **Doye A, Mettouchi A, Bossis G, Clément R, Buisson-Touati C, Flatau G, Gagnoux L, Piechaczyk M, Boquet P & Lemichez E (2002)** CNF1 exploits the ubiquitin-proteasome machinery to restrict Rho GTPase activation for bacterial host cell invasion. *Cell* 111(4):553-564.
- 180 **Davis JM, Rasmussen SB & O'Brien AD (2005)** Cytotoxic necrotizing factor type 1 production by uropathogenic *Escherichia coli* modulates polymorphonuclear leukocyte function. *Infection and Immunity* 73(9):5301-5310.
- 181 **Guyer DM, Henderson IR, Nataro JP & Mobley HL (2000)** Identification of sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. *Mol Microbiol* 38(1):53-66.

- 182 **Maroncle NM, Sivick KE, Brady R, Stokes F-E & Mobley HLT (2006)** Protease activity, secretion, cell entry, cytotoxicity, and cellular targets of secreted autotransporter toxin of uropathogenic *Escherichia coli*. *Infection and Immunity* 74(11):6124-6134.
- 183 **Guyer DM, Radulovic S, Jones F-E & Mobley HLT (2002)** Sat, the secreted autotransporter toxin of uropathogenic *Escherichia coli*, is a vacuolating cytotoxin for bladder and kidney epithelial cells. *Infection and Immunity* 70(8):4539-4546.
- 184 **Parreira VR & Gyles CL (2003)** A novel pathogenicity island integrated adjacent to the *thrW* tRNA gene of avian pathogenic *Escherichia coli* encodes a vacuolating autotransporter toxin. *Infection and Immunity* 71(9):5087-5096.
- 185 **Restieri C, Garriss G, Locas M-C & Dozois CM (2007)** Autotransporter-encoding sequences are phylogenetically distributed among *Escherichia coli* clinical isolates and reference strains. *Appl Environ Microbiol* 73(5):1553-1562.
- 186 **Tarr PI, Bilge SS, Vary JC, Jelacic S, Habeeb RL, Ward TR, Baylor MR & Besser TE (2000)** Iha: a novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. *Infection and Immunity* 68(3):1400-1407.
- 187 **Bauer RJ, Zhang L, Foxman B, Siitonen A, Jantunen ME, Saxen H & Marrs CF (2002)** Molecular epidemiology of 3 putative virulence genes for *Escherichia coli* urinary tract infection-*usp*, *iha*, and *iron*(*E. coli*). *J Infect Dis* 185(10):1521-1524.
- 188 **Johnson JR, Russo TA, Tarr PI, Carlino U, Bilge SS, Vary JC & Stell AL (2000)** Molecular epidemiological and phylogenetic associations of two novel putative virulence genes, *iha* and *iron*(*E. coli*), among *Escherichia coli* isolates from patients with urosepsis. *Infection and Immunity* 68(5):3040-3047.
- 189 **Léveillé S, Caza M, Johnson JR, Clabots C, Sabri M & Dozois CM (2006)** Iha from an *Escherichia coli* urinary tract infection outbreak clonal group A strain is expressed in vivo in the mouse urinary tract and functions as a catecholate siderophore receptor. *Infection and Immunity* 74(6):3427-3436.
- 190 **Houwink AL & Van Iterson W (1950)** Electron microscopical observations on bacterial cytology; a study on flagellation. *Biochim Biophys Acta* 5(1):10-44.
- 191 **Duguid JP, Smith IW, Dempster G & Edmunds PN (1955)** Non-flagellar filamentous appendages (fimbriae) and haemagglutinating activity in *Bacterium coli*. *J Pathol Bacteriol* 70(2):335-348.
- 192 **Ottow JC (1975)** Ecology, physiology, and genetics of fimbriae and pili. *Annu Rev Microbiol* 29:79-108.
- 193 **Korea C-G, Ghigo J-M & Beloin C (2011)** The sweet connection: Solving the riddle of multiple sugar-binding fimbrial adhesins in *Escherichia coli* : Multiple *E. coli* fimbriae form a versatile arsenal of sugar-binding lectins potentially involved in surface-colonisation and tissue tropism. *Bioessays* 33(4):300-311.
- 194 **Le Bouguéneq C (2005)** Adhesins and invasins of pathogenic *Escherichia coli*. *International Journal of Medical Microbiology* 295(6-7):471-478.

- 195 **Soto GE & Hultgren SJ (1999)** Bacterial adhesins: common themes and variations in architecture and assembly. *Journal of Bacteriology* 181(4):1059-1071.
- 196 **Jones CH, Pinkner JS, Roth R, Heuser J, Nicholes AV, Abraham SN & Hultgren SJ (1995)** FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. *Proc Natl Acad Sci USA* 92(6):2081-2085.
- 197 **Henderson IR, Owen P & Nataro JP (1999)** Molecular switches--the ON and OFF of bacterial phase variation. *Mol Microbiol* 33(5):919-932.
- 198 **Nishiyama M, Horst R, Eidam O, Herrmann T, Ignatov O, Vetsch M, Bettendorff P, Jelesarov I, Grütter MG, Wüthrich K, Glockshuber R & Capitani G (2005)** Structural basis of chaperone-subunit complex recognition by the type 1 pilus assembly platform FimD. *EMBO J* 24(12):2075-2086.
- 199 **Duguid JP, Anderson ES & Campbell I (1966)** Fimbriae and adhesive properties in *Salmonellae*. *J Pathol Bacteriol* 92(1):107-138.
- 200 **Nuccio S-P & Bäumler AJ (2007)** Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol Mol Biol Rev* 71(4):551-575.
- 201 **Orndorff PE & Falkow S (1984)** Organization and expression of genes responsible for type 1 piliation in *Escherichia coli*. *Journal of Bacteriology* 159(2):736-744.
- 202 **Mulvey MA (2002)** Adhesion and entry of uropathogenic *Escherichia coli*. *Cell Microbiol* 4(5):257-271.
- 203 **Jones CH, Pinkner JS, Nicholes AV, Slonim LN, Abraham SN & Hultgren SJ (1993)** FimC is a periplasmic PapD-like chaperone that directs assembly of type 1 pili in bacteria. *Proc Natl Acad Sci USA* 90(18):8397-8401.
- 204 **Russell PW & Orndorff PE (1992)** Lesions in two *Escherichia coli* type 1 pilus genes alter pilus number and length without affecting receptor binding. *J Bacteriol* 174(18):5923-5935.
- 205 **Krogfelt KA, Bergmans H & Klemm P (1990)** Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* type 1 fimbriae. *Infection and Immunity* 58(6):1995-1998.
- 206 **Klemm P & Christiansen G (1990)** The *fimD* gene required for cell surface localization of *Escherichia coli* type 1 fimbriae. *Mol Gen Genet* 220(2):334-338.
- 207 **Klemm P & Christiansen G (1987)** Three *fim* genes required for the regulation of length and mediation of adhesion of *Escherichia coli* type 1 fimbriae. *Mol Gen Genet* 208(3):439-445.
- 208 **Orndorff PE & Falkow S (1985)** Nucleotide sequence of *pilA*, the gene encoding the structural component of type 1 pili in *Escherichia coli*. *J Bacteriol* 162(1):454-457.
- 209 **Klemm P (1984)** The *fimA* gene encoding the type-1 fimbrial subunit of *Escherichia coli*. Nucleotide sequence and primary structure of the protein. *Eur J Biochem* 143(2):395-399.

- 210 **Zeiner SA, Dwyer BE & Clegg S (2012)** FimA, FimF, and FimH are necessary for assembly of type 1 fimbriae on *Salmonella enterica* serovar Typhimurium. *Infection and Immunity* 80(9):3289-3296.
- 211 **Valenski ML, Harris SL, Spears PA, Horton JR & Orndorff PE (2003)** The Product of the *fimI* gene is necessary for *Escherichia coli* type 1 pilus biosynthesis. *J Bacteriol* 185(16):5007-5011.
- 212 **Klemm P (1986)** Two regulatory fim genes, *fimB* and *fimE*, control the phase variation of type 1 fimbriae in *Escherichia coli*. *EMBO J* 5(6):1389-1393.
- 213 **Klemm P, Jørgensen BJ, Van Die I, de Ree H & Bergmans H (1985)** The *fim* genes responsible for synthesis of type 1 fimbriae in *Escherichia coli*, cloning and genetic organization. *Mol Gen Genet* 199(3):410-414.
- 214 **Orndorff PE & Falkow S (1984)** Identification and characterization of a gene product that regulates type 1 piliation in *Escherichia coli*. *J Bacteriol* 160(1):61-66.
- 215 **Lowe MA, Holt SC & Eisenstein BI (1987)** Immunoelectron microscopic analysis of elongation of type 1 fimbriae in *Escherichia coli*. *Journal of Bacteriology* 169(1):157-163.
- 216 **Knight SD, Berglund J & Choudhury D (2000)** Bacterial adhesins: structural studies reveal chaperone function and pilus biogenesis. *Curr Opin Chem Biol* 4(6):653-660.
- 217 **Zhou G, Mo WJ, Sebbel P, Min G, Neubert TA, Glockshuber R, Wu XR, Sun TT & Kong XP (2001)** Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from in vitro FimH binding. *J Cell Sci* 114(Pt 22):4095-4103.
- 218 **Schilling JD, Mulvey MA & Hultgren SJ (2001)** Structure and function of *Escherichia coli* type 1 pili: new insight into the pathogenesis of urinary tract infections. *J Infect Dis* 183 Suppl 1:S36-40.
- 219 **Saulino ET, Bullitt E & Hultgren SJ (2000)** Snapshots of usher-mediated protein secretion and ordered pilus assembly. *Proc Natl Acad Sci USA* 97(16):9240-9245.
- 220 **Snyder JA, Haugen BJ, Buckles EL, Lockett CV, Johnson DE, Donnenberg MS, Welch RA & Mobley HLT (2004)** Transcriptome of uropathogenic *Escherichia coli* during urinary tract infection. *Infection and Immunity* 72(11):6373-6381.
- 221 **Bahrani-Mougeot FK, Buckles EL, Lockett CV, Hebel JR, Johnson DE, Tang CM & Donnenberg MS (2002)** Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic *Escherichia coli* virulence determinants in the murine urinary tract. *Mol Microbiol* 45(4):1079-1093.
- 222 **Connell I, Agace W, Klemm P, Schembri M, Mårild S & Svanborg C (1996)** Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. *Proc Natl Acad Sci USA* 93(18):9827-9832.
- 223 **Leffler H & Svanborg-Edén C (1981)** Glycolipid receptors for uropathogenic *Escherichia coli* on human erythrocytes and uroepithelial cells. *Infection and Immunity* 34(3):920-929.

- 224 **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. *Infection and Immunity* 33(3):933-938.
- 225 **Brauner A, Leissner M, Wretling B, Julander I, Svenson SB & Källenius G (1985)** Occurrence of P-fimbriated *Escherichia coli* in patients with bacteremia. *Eur J Clin Microbiol* 4(6):566-569.
- 226 **Lane MC & Mobley HLT (2007)** Role of P-fimbrial-mediated adherence in pyelonephritis and persistence of uropathogenic *Escherichia coli* (UPEC) in the mammalian kidney. *Kidney Int* 72(1):19-25.
- 227 **Thanassi DG, Saulino ET, Lombardo MJ, Roth R, Heuser J & Hultgren SJ (1998)** The PapC usher forms an oligomeric channel: implications for pilus biogenesis across the outer membrane. *Proc Natl Acad Sci USA* 95(6):3146-3151.
- 228 **Soto GE, Dodson KW, Ogg D, Liu C, Heuser J, Knight S, Kihlberg J, Jones CH & Hultgren SJ (1998)** Periplasmic chaperone recognition motif of subunits mediates quaternary interactions in the pilus. *EMBO J* 17(21):6155-6167.
- 229 **Jones CH, Danese PN, Pinkner JS, Silhavy TJ & Hultgren SJ (1997)** The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems. *EMBO J* 16(21):6394-6406.
- 230 **Lindberg F, Tennent JM, Hultgren SJ, Lund B & Normark S (1989)** PapD, a periplasmic transport protein in P-pilus biogenesis. *J Bacteriol* 171(11):6052-6058.
- 231 **Lund B, Lindberg F, Marklund BI & Normark S (1987)** The PapG protein is the alpha-D-galactopyranosyl-(1---4)-beta-D-galactopyranose-binding adhesin of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* 84(16):5898-5902.
- 232 **Lindberg FP, Lund B & Normark S (1984)** Genes of pyelonephritogenic *E. coli* required for digalactoside-specific agglutination of human cells. *EMBO J* 3(5):1167-1173.
- 233 **Fernández LA & Berenguer J (2000)** Secretion and assembly of regular surface structures in Gram-negative bacteria. *FEMS Microbiology Reviews* 24(1):21-44.
- 234 **Strömberg 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(6):2001-2010.
- 235 **Leffler H & Edén CS (1980)** Chemical identification of a glycosphingolipid receptor for *Escherichia coli* attaching to human urinary tract epithelial cells and agglutinating human erythrocytes. *FEMS Microbiol Lett* 8(3):127-134.
- 236 **Wullt B, Bergsten G, Samuelsson M, Gebretsadik N, Hull R & Svanborg C (2001)** The role of P fimbriae for colonization and host response induction in the human urinary tract. *J Infect Dis* 183 Suppl 1:S43-46.
- 237 **Welch RA, Burland V, Plunkett G, Redford P, Roesch P, Rasko D, Buckles EL, Liou S-R, Boutin A, Hackett J, Stroud D, Mayhew GF, Rose DJ, Zhou S, Schwartz DC, Perna NT, Mobley HLT, Donnenberg MS & Blattner FR (2002)** Extensive

- mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* 99(26):17020-17024.
- 238 **Gally DL, Leathart J & Blomfield IC (1996)** Interaction of FimB and FimE with the fim switch that controls the phase variation of type 1 fimbriae in *Escherichia coli* K-12. *Mol Microbiol* 21(4):725-738.
- 239 **Abraham JM, Freitag CS, Clements JR & Eisenstein BI (1985)** An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Proc Natl Acad Sci USA* 82(17):5724-5727.
- 240 **McClain MS, Blomfield IC & Eisenstein BI (1991)** Roles of *fimB* and *fimE* in site-specific DNA inversion associated with phase variation of type 1 fimbriae in *Escherichia coli*. *J Bacteriol* 173(17):5308-5314.
- 241 **Orndorff PE, Spears PA, Schauer D & Falkow S (1985)** Two modes of control of *pilA*, the gene encoding type 1 pilin in *Escherichia coli*. *Journal of Bacteriology* 164(1):321-330.
- 242 **Kulasekara HD & Blomfield IC (1999)** The molecular basis for the specificity of *fimE* in the phase variation of type 1 fimbriae of *Escherichia coli* K-12. *Mol Microbiol* 31(4):1171-1181.
- 243 **van der Woude MW & Bäumlér AJ (2004)** Phase and antigenic variation in bacteria. *Clin Microbiol Rev* 17(3):581-611, table of contents.
- 244 **Bryan A, Roesch P, Davis L, Moritz R, Pellett S & Welch RA (2006)** Regulation of type 1 fimbriae by unlinked FimB- and FimE-like recombinases in uropathogenic *Escherichia coli* strain CFT073. *Infection and Immunity* 74(2):1072-1083.
- 245 **McClain MS, Blomfield IC, Eberhardt KJ & Eisenstein BI (1993)** Inversion-independent phase variation of type 1 fimbriae in *Escherichia coli*. *J Bacteriol* 175(14):4335-4344.
- 246 **Blomfield IC, Kulasekara DH & Eisenstein BI (1997)** Integration host factor stimulates both FimB- and FimE-mediated site-specific DNA inversion that controls phase variation of type 1 fimbriae expression in *Escherichia coli*. *Mol Microbiol* 23(4):705-717.
- 247 **Gally DL, Bogan JA, Eisenstein BI & Blomfield IC (1993)** Environmental regulation of the fim switch controlling type 1 fimbrial phase variation in *Escherichia coli* K-12: effects of temperature and media. *J Bacteriol* 175(19):6186-6193.
- 248 **Schwan WR, Lee JL, Lenard FA, Matthews BT & Beck MT (2002)** Osmolarity and pH growth conditions regulate *fim* gene transcription and type 1 pilus expression in uropathogenic *Escherichia coli*. *Infection and Immunity* 70(3):1391-1402.
- 249 **Dorman CJ & Ní Bhriain N (1992)** Thermal regulation of *fimA*, the *Escherichia coli* gene coding for the type 1 fimbrial subunit protein. *FEMS Microbiol Lett* 99(2-3):125-130.
- 250 **Old DC & Duguid JP (1970)** Selective outgrowth of fimbriate bacteria in static liquid medium. *J Bacteriol* 103(2):447-456.

- 251 **Crépin S, Houle S, Charbonneau M-È, Mourez M, Harel J & Dozois CM (2012)** Decreased expression of type 1 fimbriae by a pst mutant of uropathogenic *Escherichia coli* reduces urinary tract infection. *Infection and Immunity* 80(8):2802-2815.
- 252 **Aberg A, Shingler V & Balsalobre C (2006)** (p)ppGpp regulates type 1 fimbriation of *Escherichia coli* by modulating the expression of the site-specific recombinase FimB. *Mol Microbiol* 60(6):1520-1533.
- 253 **Lane MC, Simms AN & Mobley HLT (2007)** Complex interplay between type 1 fimbrial expression and flagellum-mediated motility of uropathogenic *Escherichia coli*. *J Bacteriol* 189(15):5523-5533.
- 254 **Holden NJ & Gally DL (2004)** Switches, cross-talk and memory in *Escherichia coli* adherence. *J Med Microbiol* 53(Pt 7):585-593.
- 255 **Blyn LB, Braaten BA, White-Ziegler CA, Rolfson DH & Low DA (1989)** Phase-variation of pyelonephritis-associated pili in *Escherichia coli*: evidence for transcriptional regulation. *EMBO J* 8(2):613-620.
- 256 **Nou X, Braaten B, Kaltenbach L & Low DA (1995)** Differential binding of Lrp to two sets of pap DNA binding sites mediated by Pap I regulates Pap phase variation in *Escherichia coli*. *EMBO J* 14(23):5785-5797.
- 257 **Kaltenbach LS, Braaten BA & Low DA (1995)** Specific binding of PapI to Lrp-pap DNA complexes. *Journal of Bacteriology* 177(22):6449-6455.
- 258 **Nou X, Skinner B, Braaten B, Blyn L, Hirsch D & Low D (1993)** Regulation of pyelonephritis-associated pili phase-variation in *Escherichia coli*: binding of the PapI and the Lrp regulatory proteins is controlled by DNA methylation. *Mol Microbiol* 7(4):545-553.
- 259 **Forsman K, Sondén B, Göransson M & Uhlin BE (1992)** Antirepression function in *Escherichia coli* for the cAMP-cAMP receptor protein transcriptional activator. *Proc Natl Acad Sci USA* 89(20):9880-9884.
- 260 **Göransson M, Forsman P, Nilsson P & Uhlin BE (1989)** Upstream activating sequences that are shared by two divergently transcribed operons mediate cAMP-CRP regulation of pilus-adhesin in *Escherichia coli*. *Mol Microbiol* 3(11):1557-1565.
- 261 **Snyder JA, Lloyd AL, Lockett CV, Johnson DE & Mobley HLT (2006)** Role of phase variation of type 1 fimbriae in a uropathogenic *Escherichia coli* cystitis isolate during urinary tract infection. *Infection and Immunity* 74(2):1387-1393.
- 262 **Sakellaris H, Balding DP & Scott JR (1996)** Assembly proteins of CS1 pili of enterotoxigenic *Escherichia coli*. *Molecular Microbiology* 21(3):529-541.
- 263 **Sajjan US, Xie H, Lefebvre MD, Valvano MA & Forstner JF (2003)** Identification and molecular analysis of cable pilus biosynthesis genes in *Burkholderia cepacia*. *Microbiology (Reading, Engl)* 149(Pt 4):961-971.
- 264 **Folkesson A, Advani A, Sukupolvi S, Pfeifer JD, Normark S & Löfdahl S (1999)** Multiple insertions of fimbrial operons correlate with the evolution of *Salmonella* serovars responsible for human disease. *Mol Microbiol* 33(3):612-622.

- 265 **Berger CN, Billker O, Meyer TF, Servin AL & Kansau I (2004)** Differential recognition of members of the carcinoembryonic antigen family by Afa/Dr adhesins of diffusely adhering *Escherichia coli* (Afa/Dr DAEC). *Mol Microbiol* 52(4):963-983.
- 266 **Nowicki B, Hart A, Coyne KE, Lublin DM & Nowicki S (1993)** Short consensus repeat-3 domain of recombinant decay-accelerating factor is recognized by *Escherichia coli* recombinant Dr adhesin in a model of a cell-cell interaction. *J Exp Med* 178(6):2115-2121.
- 267 **Nowicki B, Labigne A, Moseley S, Hull R, Hull S & Moulds J (1990)** The Dr hemagglutinin, afimbrial adhesins AFA-I and AFA-III, and F1845 fimbriae of uropathogenic and diarrhea-associated *Escherichia coli* belong to a family of hemagglutinins with Dr receptor recognition. *Infection and Immunity* 58(1):279-281.
- 268 **Medof ME, Walter EI, Rutgers JL, Knowles DM & Nussenzweig V (1987)** Identification of the complement decay-accelerating factor (DAF) on epithelium and glandular cells and in body fluids. *J Exp Med* 165(3):848-864.
- 269 **Goluszko P, Popov V, Selvarangan R, Nowicki S, Pham T & Nowicki BJ (1997)** Dr fimbriae operon of uropathogenic *Escherichia coli* mediate microtubule-dependent invasion to the HeLa epithelial cell line. *J Infect Dis* 176(1):158-167.
- 270 **Khan AS, Kniep B, Oelschlaeger TA, Van Die I, Korhonen T & Hacker J (2000)** Receptor structure for F1C fimbriae of uropathogenic *Escherichia coli*. *Infection and Immunity* 68(6):3541-3547.
- 271 **Bäckhed F, Alsén B, Roche N, Angström J, von Euler A, Breimer ME, Westerlund-Wikström B, Teneberg S & Richter-Dahlfors A (2002)** Identification of target tissue glycosphingolipid receptors for uropathogenic, F1C-fimbriated *Escherichia coli* and its role in mucosal inflammation. *J Biol Chem* 277(20):18198-18205.
- 272 **Klemm P, Jørgensen BJ, Kreft B & Christiansen G (1995)** The export systems of type 1 and F1C fimbriae are interchangeable but work in parental pairs. *J Bacteriol* 177(3):621-627.
- 273 **Klemm P, Christiansen G, Kreft B, Marre R & Bergmans H (1994)** Reciprocal exchange of minor components of type 1 and F1C fimbriae results in hybrid organelles with changed receptor specificities. *J Bacteriol* 176(8):2227-2234.
- 274 **Reitzer L (2003)** Nitrogen assimilation and global regulation in *Escherichia coli*. *Annu Rev Microbiol* 57:155-176.
- 275 **Llácer JL, Fita I & Rubio V (2008)** Arginine and nitrogen storage. *Curr Opin Struct Biol* 18(6):673-681.
- 276 **Schneider BL, Kiupakis AK & Reitzer LJ (1998)** Arginine catabolism and the arginine succinyltransferase pathway in *Escherichia coli*. *Journal of Bacteriology* 180(16):4278-4286.
- 277 **Houtsmuller UM & Van Deenen L (1963)** Identification of a bacterial phospholipid as an O-ornithine ester of phosphatidyl glycerol. *Biochim Biophys Acta* 70:211-213.

- 278 **Lin J, Lee IS, Frey J, Slonczewski JL & Foster JW (1995)** Comparative analysis of extreme acid survival in *Salmonella typhimurium*, *Shigella flexneri*, and *Escherichia coli*. *Journal of Bacteriology* 177(14):4097-4104.
- 279 **Chaddock AM, Mant A, Karnauchov I, Brink S, Herrmann RG, Klösgen RB & Robinson C (1995)** A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the delta pH-dependent thylakoidal protein translocase. *EMBO J* 14(12):2715-2722.
- 280 **Maas WK (1964)** Studies on the mechanism of repression of arginine biosynthesis in *Escherichia coli*. II. Dominance of repressibility in diploids. *J Mol Biol* 8:365-370.
- 281 **Demerec M, Adelberg EA, Clark AJ & Hartman PE (1966)** A proposal for a uniform nomenclature in bacterial genetics. *Genetics* 54(1):61-76.
- 282 **Maas WK, Maas R, Wiame JM & Glansdorff N (1964)** Studies on the mechanism of repression of arginine biosynthesis in *Escherichia coli*. I. Dominance of repressibility in zygotes. *J Mol Biol* 8:359-364.
- 283 **Maas WK (1961)** Studies on repression of arginine biosynthesis in *Escherichia coli*. *Cold Spring Harb Symp Quant Biol* 26:183-191.
- 284 **Gorini L, Gundersen W & Burger M (1961)** Genetics of regulation of enzyme synthesis in the arginine biosynthetic pathway of *Escherichia coli*. *Cold Spring Harb Symp Quant Biol* 26:173-182.
- 285 **Cunin R, Eckhardt T, Piette J, Boyen A, Piérard A & Glansdorff N (1983)** Molecular basis for modulated regulation of gene expression in the arginine regulon of *Escherichia coli* K-12. *Nucleic Acids Res* 11(15):5007-5019.
- 286 **Van Duyne GD, Ghosh G, Maas WK & Sigler PB (1996)** Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J Mol Biol* 256(2):377-391.
- 287 **Caldara M, Charlier D & Cunin R (2006)** The arginine regulon of *Escherichia coli*: whole-system transcriptome analysis discovers new genes and provides an integrated view of arginine regulation. *Microbiology (Reading, Engl)* 152(Pt 11):3343-3354.
- 288 **Tian G, Lim D, Carey J & Maas WK (1992)** Binding of the arginine repressor of *Escherichia coli* K12 to its operator sites. *J Mol Biol* 226(2):387-397.
- 289 **Charlier D, Roovers M, Van Vliet F, Boyen A, Cunin R, Nakamura Y, Glansdorff N & Piérard A (1992)** Arginine regulon of *Escherichia coli* K-12. A study of repressor-operator interactions and of in vitro binding affinities versus in vivo repression. *J Mol Biol* 226(2):367-386.
- 290 **Rosen BP (1971)** Basic amino acid transport in *Escherichia coli*. *J Biol Chem* 246(11):3653-3662.
- 291 **Celis TF, Rosenfeld HJ & Maas WK (1973)** Mutant of *Escherichia coli* K-12 defective in the transport of basic amino acids. *Journal of Bacteriology* 116(2):619-626.

- 292 **Celis RT (1981)** Chain-terminating mutants affecting a periplasmic binding protein involved in the active transport of arginine and ornithine in *Escherichia coli*. *J Biol Chem* 256(2):773-779.
- 293 **Celis TF (1977)** Properties of an *Escherichia coli* K-12 mutant defective in the transport of arginine and ornithine. *Journal of Bacteriology* 130(3):1234-1243.
- 294 **Rosen BP (1973)** Basic amino acid transport in *Escherichia coli*. II. Purification and properties of an arginine-specific binding protein. *J Biol Chem* 248(4):1211-1218.
- 295 **Wissenbach U, Six S, Bongaerts J, Ternes D, Steinwachs S & Uden G (1995)** A third periplasmic transport system for L-arginine in *Escherichia coli*: molecular characterization of the *artPIQMJ* genes, arginine binding and transport. *Mol Microbiol* 17(4):675-686.
- 296 **Celis TF (1977)** Independent regulation of transport and biosynthesis of arginine in *Escherichia coli* K-12. *Journal of Bacteriology* 130(3):1244-1252.
- 297 **Higgins CF (1992)** ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* 8:67-113.
- 298 **Celis RT, Leadlay PF, Roy I & Hansen A (1998)** Phosphorylation of the periplasmic binding protein in two transport systems for arginine incorporation in *Escherichia coli* K-12 is unrelated to the function of the transport system. *Journal of Bacteriology* 180(18):4828-4833.
- 299 **Celis RT (1990)** Mutant of *Escherichia coli* K-12 with defective phosphorylation of two periplasmic transport proteins. *J Biol Chem* 265(3):1787-1793.
- 300 **Hobson AC, Weatherwax R & Ames GF (1984)** ATP-binding sites in the membrane components of histidine permease, a periplasmic transport system. *Proc Natl Acad Sci USA* 81(23):7333-7337.
- 301 **Celis RT (1984)** Phosphorylation in vivo and in vitro of the arginine-ornithine periplasmic transport protein of *Escherichia coli*. *Eur J Biochem* 145(2):403-411.
- 302 **Caldara M, Minh PNL, Bostoen S, Massant J & Charlier D (2007)** ArgR-dependent repression of arginine and histidine transport genes in *Escherichia coli* K-12. *J Mol Biol* 373(2):251-267.
- 303 **Higgins CF, Haag PD, Nikaido K, Ardeshir F, Garcia G & Ames GF (1982)** Complete nucleotide sequence and identification of membrane components of the histidine transport operon of *S. typhimurium*. *Nature* 298(5876):723-727.
- 304 **Glansdorff N (1996)** Chapter 3.6.1.10, Biosynthesis of Arginine and Polyamines. In A. Böck, R. Curtiss III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, and D. Ussery. *EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology*. <http://www.ecosal.org>. (ASM Press, Washington, DC)
- 305 **Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, Peter BJ, Bender RA & Kustu S (2000)** Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation. *Proc Natl Acad Sci USA* 97(26):14674-14679.

- 306 **Ames GF, Nikaido K, Wang IX, Liu PQ, Liu CE & Hu C (2001)** Purification and characterization of the membrane-bound complex of an ABC transporter, the histidine permease. *J Bioenerg Biomembr* 33(2):79-92.
- 307 **Kerppola RE, Shyamala VK, Klebba P & Ames GF (1991)** The membrane-bound proteins of periplasmic permeases form a complex. Identification of the histidine permease HisQMP complex. *J Biol Chem* 266(15):9857-9865.
- 308 **Kustu SG & Ames GF (1973)** The HisP protein, a known histidine transport component in *Salmonella typhimurium*, is also an arginine transport component. *Journal of Bacteriology* 116(1):107-113.
- 309 **Higgins CF & Ames GF (1981)** Two periplasmic transport proteins which interact with a common membrane receptor show extensive homology: complete nucleotide sequences. *Proc Natl Acad Sci USA* 78(10):6038-6042.
- 310 **Tam R & Saier MH (1993)** Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol Rev* 57(2):320-346.
- 311 **Nonet ML, Marvel CC & Tolan DR (1987)** The *hisT-purF* region of the *Escherichia coli* K-12 chromosome. Identification of additional genes of the *hisT* and *purF* operons. *J Biol Chem* 262(25):12209-12217.
- 312 **Ames GF & Nikaido K (1978)** Identification of a membrane protein as a histidine transport component in *Salmonella typhimurium*. *Proc Natl Acad Sci USA* 75(11):5447-5451.
- 313 **Ames GF & Lever JE (1972)** The histidine-binding protein J is a component of histidine transport. Identification of its structural gene, *hisJ*. *J Biol Chem* 247(13):4309-4316.
- 314 **Ames GF, Nikaido K, Hobson A & Malcolm B (1985)** Overproduction of the membrane-bound components of the histidine permease from *Salmonella typhimurium*: identification of the M protein. *Biochimie* 67(1):149-154.
- 315 **Ikeda TP, Shauger AE & Kustu S (1996)** *Salmonella typhimurium* apparently perceives external nitrogen limitation as internal glutamine limitation. *J Mol Biol* 259(4):589-607.
- 316 **Kustu SG, McFarland NC, Hui SP, Esmon B & Ames GF (1979)** Nitrogen control of *Salmonella typhimurium*: co-regulation of synthesis of glutamine synthetase and amino acid transport systems. *Journal of Bacteriology* 138(1):218-234.
- 317 **Celis RT (1982)** Mapping of two loci affecting the synthesis and structure of a periplasmic protein involved in arginine and ornithine transport in *Escherichia coli* K-12. *Journal of Bacteriology* 151(3):1314-1319.
- 318 **Wissenbach U, Keck B & Uden G (1993)** Physical map location of the new *artPIQMJ* genes of *Escherichia coli*, encoding a periplasmic arginine transport system. *Journal of Bacteriology* 175(11):3687-3688.
- 319 **Reitzer L & Schneider BL (2001)** Metabolic context and possible physiological themes of sigma(54)-dependent genes in *Escherichia coli*. *Microbiol Mol Biol Rev* 65(3):422-444, table of contents.

- 320 **Russo TA, Carlino UB, Mong A & Jodush ST (1999)** Identification of genes in an extraintestinal isolate of *Escherichia coli* with increased expression after exposure to human urine. *Infect Immun* 67(10):5306-5314.
- 321 **Paul L, Mishra PK, Blumenthal RM & Matthews RG (2007)** Integration of regulatory signals through involvement of multiple global regulators: control of the *Escherichia coli* *gltBDF* operon by Lrp, IHF, Crp, and ArgR. *BMC Microbiol* 7:2.
- 322 **Cunin R, Glansdorff N, Piérard A & Stalon V (1986)** Biosynthesis and metabolism of arginine in bacteria. *Microbiol Rev* 50(3):314-352.
- 323 **Kiupakis AK & Reitzer L (2002)** ArgR-independent induction and ArgR-dependent superinduction of the *astCADBE* operon in *Escherichia coli*. *Journal of Bacteriology* 184(11):2940-2950.
- 324 **Maas WK (1994)** The arginine repressor of *Escherichia coli*. *Microbiol Rev* 58(4):631-640.
- 325 **Lim DB, Oppenheim JD, Eckhardt T & Maas WK (1987)** Nucleotide sequence of the *argR* gene of *Escherichia coli* K-12 and isolation of its product, the arginine repressor. *Proc Natl Acad Sci USA* 84(19):6697-6701.
- 326 **Szwajkajzer D, Dai L, Fukayama JW, Abramczyk B, Fairman R & Carey J (2001)** Quantitative analysis of DNA binding by the *Escherichia coli* arginine repressor. *J Mol Biol* 312(5):949-962.
- 327 **Tian G, Lim D, Oppenheim JD & Maas WK (1994)** Explanation for different types of regulation of arginine biosynthesis in *Escherichia coli* B and *Escherichia coli* K12 caused by a difference between their arginine repressors. *J Mol Biol* 235(1):221-230.
- 328 **Lissens W, Cunin R, Kelker N, Glansdorff N & Piérard A (1980)** In vitro synthesis of *Escherichia coli* carbamoylphosphate synthase: evidence for participation of the arginine repressor in cumulative repression. *Journal of Bacteriology* 141(1):58-66.
- 329 **Makarova KS, Mironov AA & Gelfand MS (2001)** Conservation of the binding site for the arginine repressor in all bacterial lineages. *Genome Biol* 2(4):RESEARCH0013.
- 330 **Mironov AA, Koonin EV, Roytberg MA & Gelfand MS (1999)** Computer analysis of transcription regulatory patterns in completely sequenced bacterial genomes. *Nucleic Acids Res* 27(14):2981-2989.
- 331 **Stirling CJ, Szatmari G, Stewart G, Smith MC & Sherratt DJ (1988)** The arginine repressor is essential for plasmid-stabilizing site-specific recombination at the ColE1 *cer* locus. *EMBO J* 7(13):4389-4395.
- 332 **Wang H, Glansdorff N & Charlier D (1998)** The arginine repressor of *Escherichia coli* K-12 makes direct contacts to minor and major groove determinants of the operators. *J Mol Biol* 277(4):805-824.
- 333 **Maas WK (1956)** Regulation of arginine biosynthesis in *Escherichia coli*. *Biol. Bull.* 111(2):319.

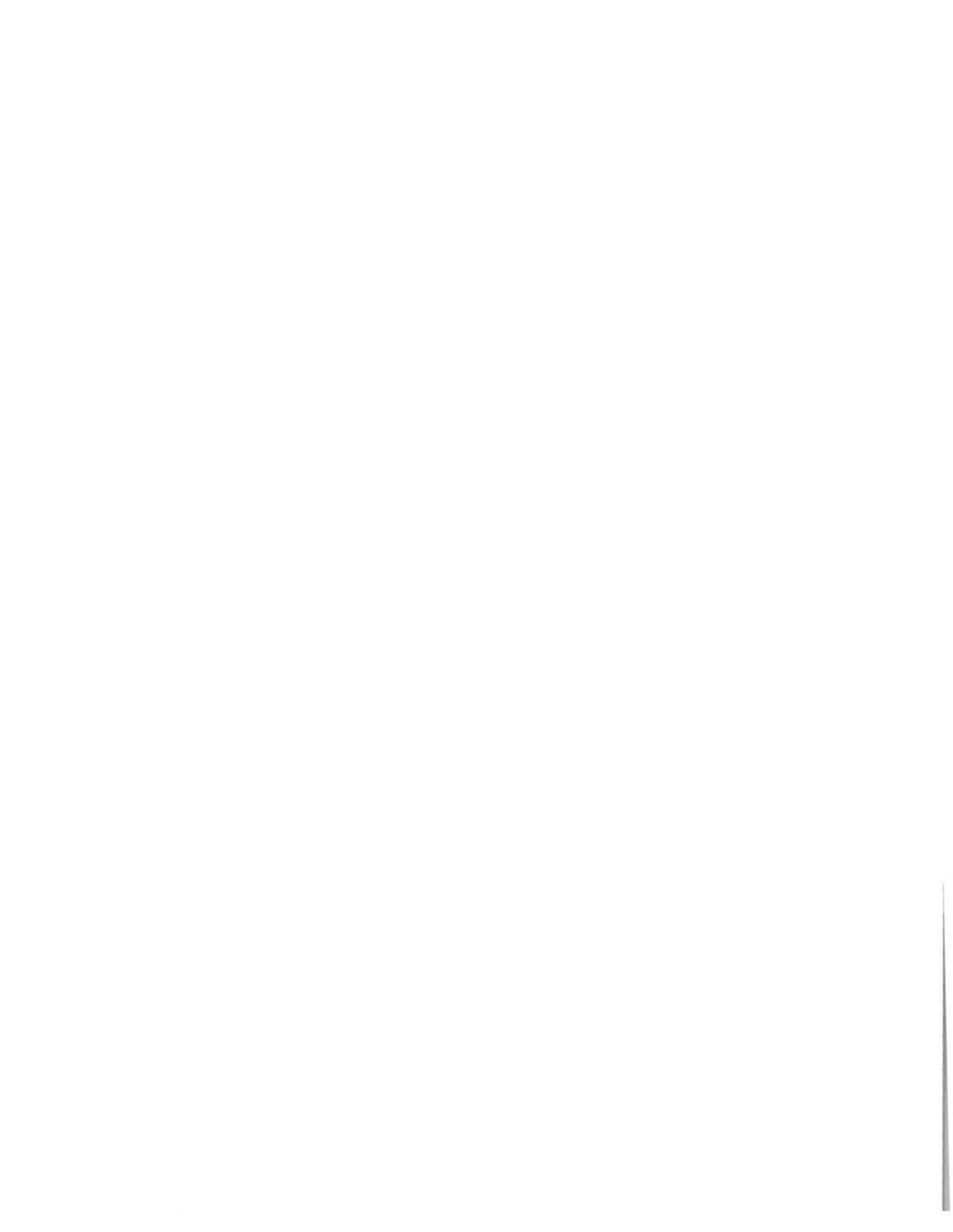
- 334 **Jacoby GA & Gorini L (1969)** A unitary account of the repression mechanism of arginine biosynthesis in *Escherichia coli*. I. The genetic evidence. *J Mol Biol* 39(1):73-87.
- 335 **Krin E, Laurent-Winter C, Bertin PN, Danchin A & Kolb A (2003)** Transcription regulation coupling of the divergent *argG* and *metY* promoters in *Escherichia coli* K-12. *Journal of Bacteriology* 185(10):3139-3146.
- 336 **Guhathakurta A, Viney I & Summers D (1996)** Accessory proteins impose site selectivity during ColE1 dimer resolution. *Mol Microbiol* 20(3):613-620.
- 337 **Guhathakurta A & Summers D (1995)** Involvement of ArgR and PepA in the pairing of ColE1 dimer resolution sites. *Microbiology (Reading, Engl)* 141 (Pt 5):1163-1171.
- 338 **Summers D, Yaish S, Archer J & Sherratt D (1985)** Multimer resolution systems of ColE1 and ColK: localisation of the crossover site. *Mol Gen Genet* 201(2):334-338.
- 339 **Paul S & Summers D (2004)** ArgR and PepA, accessory proteins for XerCD-mediated resolution of ColE1 dimers, are also required for stable maintenance of the P1 prophage. *Plasmid* 52(1):63-68.
- 340 **Schaumburg CS & Tan M (2006)** Arginine-dependent gene regulation via the ArgR repressor is species specific in chlamydia. *Journal of Bacteriology* 188(3):919-927.
- 341 **Lymberopoulos M (2004)** *Identification, caractérisation et distribution phylogénétique du fimbriae IR chez Escherichia coli*. M.Sc en microbiologie appliquée (Université du Québec. INRS-Institut Armand-Frappier, Laval). 109 p
- 342 **Provence DL & Curtiss R (1992)** Role of *crl* in avian pathogenic *Escherichia coli*: a knockout mutation of *crl* does not affect hemagglutination activity, fibronectin binding, or Curli production. *Infection and Immunity* 60(11):4460-4467.
- 343 **Haldimann A & Wanner BL (2001)** Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *Journal of Bacteriology* 183(21):6384-6393.
- 344 **Dozois CM, Dho-Moulin M, Brée A, Fairbrother JM, Desautels C & Curtiss 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.
- 345 **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. *Journal of Bacteriology* 175(9):2770-2778.
- 346 **Ewing WH, Davis BR, Fife MA & Lessel EF (1973)** Biochemical Characterization of *Serratia liquefaciens* (Grimes and Hennerty) Bascomb et al. (Formerly *Enterobacter liquefaciens*) and *Serratia rubidaea* (Stapp) comb. nov. and Designation of Type and Neotype Strains. *International Journal of Systematic Bacteriology* 23(3):217-225.
- 347 **Séguin J (2009)** *Inhibition de l'adhérence des Escherichia coli pathogènes extraintestinaux par l'utilisation de composées mannosylé*. Québec).

- 348 **Datsenko KA & Wanner BL (2000)** One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97(12):6640-6645.
- 349 **Crépin S, Harel J & Dozois CM (2012)** Chromosomal complementation using tn7 transposon vectors in enterobacteriaceae. *Appl Environ Microbiol* 78(17):6001-6008.
- 350 **Simons RW, Houman F & Kleckner N (1987)** Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene* 53(1):85-96.
- 351 **Sambrook J & Russell DDW. (2001).** *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press,
- 352 **Le Bouguenec C, Archambaud M & Labigne A (1992)** Rapid and specific detection of the *pap*, *afa*, and *sfa* adhesin-encoding operons in uropathogenic *Escherichia coli* strains by polymerase chain reaction. *J Clin Microbiol* 30(5):1189-1193.
- 353 **Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C & Bryant SH (2011)** CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* 39(Database issue):D225-229.
- 354 **Tamura K, Peterson D, Peterson N, Stecher G, Nei M & Kumar S (2011)** MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731-2739.
- 355 **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ & Higgins DG (2007)** Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-2948.
- 356 **Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J & Lopez R (2010)** A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Research* 38(suppl 2):695-699.
- 357 **Miller JH. (1992).** *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia Coli and Related Bacteria and '99 Source Book*. Cold Spring Harbor Laboratory Press,
- 358 **Holmgren A & Branden CI (1989)** Crystal structure of chaperone protein PapD reveals an immunoglobulin fold. *Nature* 342(6247):248-251.
- 359 **Hultgren SJ, Normark S & Abraham SN (1991)** Chaperone-assisted assembly and molecular architecture of adhesive pili. *Annu Rev Microbiol* 45:383-415.
- 360 **Sakellaris H & Scott JR (1998)** New tools in an old trade: CS1 pilus morphogenesis. *Mol Microbiol* 30(4):681-687.
- 361 **Jacob-Dubuisson F, Kuehn M & Hultgren SJ (1993)** A novel secretion apparatus for the assembly of adhesive bacterial pili. *Trends Microbiol* 1(2):50-55.
- 362 **Sajjan US, Xie H, Lefebvre MD, Valvano MA & Forstner JF (2003)** Identification and molecular analysis of cable pilus biosynthesis genes in *Burkholderia cepacia*. *Microbiology* 149(Pt 4):961-971.

- 363 **Hamrick TS, Harris SL, Spears PA, Havell EA, Horton JR, Russell PW & Orndorff PE (2000)** Genetic characterization of *Escherichia coli* type 1 pilus adhesin mutants and identification of a novel binding phenotype. *Journal of Bacteriology* 182(14):4012-4021.
- 364 **Jaureguy F, Landraud L, Passet V, Diancourt L, Frapy E, Guigon G, Carbonnelle E, Lortholary O, Clermont O, Denamur E, Picard B, Nassif X & Brisse S (2008)** Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics* 9:560.
- 365 **Hugovieux-Cotte-Pattat N & Robert-Baudouy J (1982)** Regulation and transcription direction of *exuR*, a self-regulated repressor in *Escherichia coli* K-12. *J Mol Biol* 156(1):221-228.
- 366 **Rodionov DA, Mironov AA, Rakhmaninova AB & Gelfand MS (2000)** Transcriptional regulation of transport and utilization systems for hexuronides, hexuronates and hexonates in gamma purple bacteria. *Mol Microbiol* 38(4):673-683.
- 367 **McPeake SJW, Smyth JA & Ball HJ (2005)** Characterisation of avian pathogenic *Escherichia coli* (APEC) associated with colisepticaemia compared to faecal isolates from healthy birds. *Veterinary Microbiology* 110(3-4):245-253.
- 368 **Malpica R, Sandoval GRP, Rodríguez C, Franco B & Georgellis D (2006)** Signaling by the arc two-component system provides a link between the redox state of the quinone pool and gene expression. *Antioxid Redox Signal* 8(5-6):781-795.
- 369 **Gunsalus RP & Park SJ (1994)** Aerobic-anaerobic gene regulation in *Escherichia coli*: control by the ArcAB and Fnr regulons. *Res Microbiol* 145(5-6):437-450.
- 370 **Cao L, Suo Z, Lim T, Jun S, Deliorman M, Riccardi C, Kellerman L, Avci R & Yang X (2012)** Role of overexpressed CFA/I fimbriae in bacterial swimming. *Phys Biol* 9(3):036005.
- 371 **Caron J, Coffield LM & Scott JR (1989)** A plasmid-encoded regulatory gene, *rns*, required for expression of the CS1 and CS2 adhesins of enterotoxigenic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 86(3):963-967.
- 372 **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ (1997)** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25(17):3389-3402.
- 373 **Rojo-Bezares B, Martin C, Lopez M, Torres C & Saenz Y (2012)** First detection of *bla_{IMI-2}* gene in a clinical *Escherichia coli* strain. *Antimicrob Agents Chemother* 56(2):1146-1147.
- 374 **Chouikha I, Germon P, Bree A, Gilot P, Moulin-Schouleur M & Schouler C (2006)** A selC-associated genomic island of the extraintestinal avian pathogenic *Escherichia coli* strain BEN2908 is involved in carbohydrate uptake and virulence. *J Bacteriol* 188(3):977-987.
- 375 **Giovanardi D, Campagnari E, Ruffoni LS, Pesente P, Ortali G & Furlattini V (2005)** Avian pathogenic *Escherichia coli* transmission from broiler breeders to their progeny in an integrated poultry production chain. *Avian Pathol* 34(4):313-318.

376 **Petersen TN, Brunak S, von Heijne G & Nielsen H (2011)** SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8(10):785-786.

ANNEXES



ANNEXE I

Brin	+	+	+
Identité en pb (%)	1329/1331 (99 %)	1098/1321 (83 %)	1011/1321 (77 %)
Origine	<i>Escherichia coli</i>	<i>Escherichia coli</i>	<i>Escherichia coli</i>
Groupe de l'IS	IS2	IS2	IS2
Famille de l'IS	IS3	IS3	IS3
Nom des IS produisant un alignement significatif (longueur pb)	IS2 (1331)	ISEc27 (1321)	ISEc36 (1321)

Tableau 5 : Résultat de l'analyse *in silico* de l'ORF *efoC_{χ7122}* par le logiciel en ligne IS Finder (372)

Résultat obtenu à partir du logiciel ISFinder se trouvant à l'adresse : <https://www-is.biotoul.fr/>

La séquence fournie, pour faire la recherche de similitudes, est les deux parties de l'ORF de *efoC* de la souche χ 7122, avec l'IS potentielle (1332 pb). Une mise à jour de l'analyse a été effectuée le 23 janvier 2013. Les paramètres utilisés sont ceux par défaut.

Seuls les trois premiers résultats, les plus significatifs, sont représentés dans ce tableau. Les ISEc27 et ISEc36 sont des dérivés de l'IS2 (373, 374). L'IS2 est celui qui interrompt l'ORF *efoC_{χ7122}*.

ANNEXE II

<u>APEC-1</u> 56 ; TK12 ; 137 ; TK1 ; 3	<u>ENV-1</u> TK105 ; TK119 ; TK99 ; TK78 ; TK76	<u>UHUM-1</u> H15 ; PM6 ; 2P6 ; V6 ; V24 ; H19	<u>ECOR-A1</u> 3 ; 6 ; 8 ; 11 ; 22
<u>APEC-2</u> 151 ; 144 ; 138 ; 140 ; 71	<u>ENV-2</u> C80 ; C97 ; C113 ; C107 ; C75	<u>UHUM-2</u> PM8 ; V31 ; PM4 ; H26 ; V19 ; V22 ; CFT073	<u>ECOR-A2</u> 14 ; 15 ; 17 ; 19 ; 21
<u>APEC-3</u> 40 ; 139 ; 154 ; 33 ; 27	<u>ENV-3</u> C128 ; C29 ; C114 ; C65 ; C125	<u>UHUM-3</u> K5 ; V21 ; 2H19 ; V29 ; PM7	<u>ECOR-B1</u> 27 ; 33 ; 45 ; 67 ; 71
<u>APEC-4</u> 30 ; 163 ; 49 ; 165 ; 14	<u>ENV-4</u> C111 ; C108 ; C92 ; C109 ; C82	<u>UHUM-4</u> U6 ; H27 ; U1 ; V28 ; PM9 ; V9	<u>ECOR-B2</u> 51 ; 56 ; 60 ; 62 ; 64
<u>APEC-5</u> 69 ; 60 ; 9 ; 61 ; TK2	<u>ENV-5</u> TK101 ; TK112 ; C115 ; C121 ; TK106 ; C67	<u>UHUM-5</u> V26 ; V32 ; 2P9 ; H17 ; CP9	<u>ECOR-D</u> 35 ; 38 ; 40 ; 48 ; 50
<u>APEC-6</u> MT513 ; MT512 ; MT78 ; MT515 ; MT181 ; MT458	<u>ENV-6</u> C120 ; C110 ; C64 ; C63 ; C126 ; C129		<u>ECOR-non</u> 31 ; 37 ; 42 ; 43

Tableau 6 : Liste des pools de souches utilisés et le résultat de la PCR multiplex avec les amorces Ecs4-F/R
Le nom des pools est écrit en gras. APEC : souches pathogènes aviaires ; ENV : souches de l'environnement des oiseaux ; UHUM : souches uropathogènes humaines ; ECOR : souches de la collection ECOR (pool déterminé par les groupes phylogénétiques (14)). Les groupes soulignés sont positifs pour la PCR multiplex avec les amorces Esc4-F et Esc4-R, avec une mise à jour sur les pools positifs de la collection ECOR. Pour plus d'information, voir le mémoire de maîtrise de M. Lymberopoulos (2004) (341).

ANNEXE III

Tableau 7 : Liste des souches criblées pour la présence de *efo*

Les souches de la collection ECOR proviennent de la référence (7). Les souches PPEC proviennent du laboratoire de J. Fairbrother (ECL, Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, Canada). Les souches APEC proviennent des références (344) et (375). Elles ont été fournies par les laboratoires de M. Moulin-Schouleur (INRA, Tours, France) et de L. Sperati Ruffoni (Laboratorio Tre Valli, Verona, Italy). Le groupe phylogénétique de la collection ECOR a été déterminé par MLST sauf pour les souches ECOR5/6/9/13/64 par MLEE. Le groupe des souches PPEC et APEC a été déterminé par la méthode Clermont. Le signe « + » signifie que la souche est positive pour *efo*. Le signe « - » signifie que la souche est négative. Le signe « \diamond » signifie que la souche est positive pour *efo*, mais avec une IS.

Souches	Groupe phylogénétique	Positive pour <i>efo</i>	Souches	Groupe phylogénétique	Positive pour <i>efo</i>
Collection					
ECOR					
ECOR1	A		ECOR37	E	
ECOR2	A		ECOR38	D	+
ECOR3	A		ECOR39	D	+
ECOR4	A	+	ECOR40	D	+
ECOR5	A		ECOR41	D	+
ECOR6	A		ECOR42	E	
ECOR7	A	+	ECOR43	A	
ECOR8	A		ECOR44	D	
ECOR9	A		ECOR45	B1	+
ECOR10	A		ECOR46	D	
ECOR11	A		ECOR47	D	
ECOR12	A		ECOR48	D	
ECOR13	A		ECOR49	D	
ECOR14	A		ECOR50	D	
ECOR15	A	+	ECOR51	B2	
ECOR16	A	+	ECOR52	B2	
ECOR17	A	+	ECOR53	B2	
ECOR18	A		ECOR54	B2	
ECOR19	A		ECOR55	B2	
ECOR20	A		ECOR56	B2	
ECOR21	A		ECOR57	B2	
ECOR22	A	+	ECOR58	B1	+
ECOR23	A		ECOR59	B2	
ECOR24	A	+	ECOR60	B2	
ECOR25	A		ECOR61	B2	
ECOR26	B1	+	ECOR62	B2	
ECOR27	B1	+	ECOR63	B2	+
ECOR28	B1	+	ECOR64	B2	+
ECOR29	B1	+	ECOR65	B2	

ECOR30	B1	+	ECOR66	B2	
ECOR31	E		ECOR67	B1	+
ECOR32	B1	+	ECOR68	B1	+
ECOR33	B1	+	ECOR69	B1	+
ECOR34	B1	+	ECOR70	B1	+
ECOR35	D	+	ECOR71	B1	+
ECOR36	D	+	ECOR72	B1	+
Souches	Groupe phylogénétique	Positive pour <i>efo</i>	Souches	Groupe phylogénétique	Positive pour <i>efo</i>
PPEC					
D149	A		P01-B005-1	B1	+
D58	A	+	P01-B629B-2	D	
ETEC315	A	+	P01-C005B-1	D	
ETEC323	A	+	P01-D634-1	A	
ETEC329	A		P01-D827A1	A	
ETEC380	A		P01-E845B-3	A	
ETEC388	A		P01-F103A2	A	
ETEC395	A		P01-F691B1	A	
ETEC94	A		P01-F711B2	A	+
P01-0604-3-2	A		P01-H207C-1	A	+
P01-7086-1	A	+	P81-4420A	A	+
P01-7291H2	D		P90-1513	B1	+
P01-8963B1	A		P97-1463A1	A	
P01-8963B2	A		P97-2177	B1	+
APECO1					
BEN 0071	D		BEN 2426	B2	
BEN 0075	A		BEN 2427	B2	
BEN 0120	B2		BEN 2444	A	◇
BEN 0121	B2		BEN 2446	B2	
BEN 0124	B2		BEN 2563	B2	
BEN 0130	B2		BEN 2596	B2	
BEN 0131	D		BEN 2638	A	
BEN 0135	B2				
BEN 0140	B2				
BEN 0156	B2				
BEN 1714	B2				
BEN 1718	B2				
BEN 1722	B2				
BEN 1731	B2				
BEN 1734	B2				
BEN 1778	B2				
BEN 1780	B2				
BEN 2423	D	+			
BEN 2425	B2	◇			

Souches	Groupe phylogénétique	Positive pour <i>efo</i>	Souches	Groupe phylogénétique	Positive pour <i>efo</i>
APECO2					
BEN 0073	B2		BEN 0149	D	
BEN 0074	A		BEN 0151	D	
BEN 0082	B2		BEN 0152	D	
BEN 0087	B2		BEN 0157	D	
BEN 0097	B2		BEN 0159	D	
BEN 0098	B2		BEN 0160	D	
BEN 0099	B2		BEN 1656	B2	
BEN 0100	B2		BEN 1715	D	
BEN 0102	B2		BEN 1716	D	
BEN 0104	B2		BEN 1717	D	
BEN 0106	B2		BEN 1730	B2	◇
BEN 0109	B2		BEN 1756	D	
BEN 0114	B2		BEN 1957	B2	
BEN 0117	A		BEN 1967	D	
BEN 0119	B2		BEN 2269	A	
BEN 0122	B2		BEN 2329	B2	
BEN 0126	B2		BEN 2332	B2	
BEN 0132	B2		BEN 2419	B2	
BEN 0133	D		BEN 2431	B2	◇
BEN 0137	B2		BEN 2432	B2	
BEN 0138	D		BEN 2905	D	
BEN 0139	D		BEN 2979	D	◇
BEN 0142	D				
APECO78					
BEN 0001	A	+	BEN 1980	A	+
BEN 0002	A	◇	BEN 1986	A	+
BEN 0003	A	◇	BEN 2031	A	+
BEN 0004	A	+	BEN 2032	A	+
BEN 0005	D		BEN 2374	A	◇
BEN 0006	A	+	BEN 2382	A	+
BEN 0007	A	◇	BEN 2383	A	+
BEN 0008	A	◇	BEN 2384	D	
BEN 0009	A	◇	BEN 2387	A	◇
BEN 0010	A	◇	BEN 2392	A	+
BEN 0011	A	◇	BEN 2393	D	
BEN 0012	A	◇	BEN 2400	A	◇
BEN 0014	D		BEN 2401	A	+
BEN 0016	A	+	BEN 2411	B2	◇
BEN 1683	A		BEN 2428	D	
BEN 1970	A	◇	BEN 2430	D	
BEN 1978	A	◇	BEN 2433	A	+
BEN 1979	A	+	BEN 2434	A	◇

Souches	Groupe phylogénétique	Positive pour <i>efo</i>	Souches	Groupe phylogénétique	Positive pour <i>efo</i>
BEN 2435	A	+	BEN 2956	A	◇
BEN 2436	A	+	BEN 2957	A	+
BEN 2437	A	◇	BEN 2958	A	+
BEN 2438	A	+	BEN 2959	A	◇
BEN 2439	B2	◇	BEN 2960	A	◇
BEN 2440	A	◇	BEN 2961	A	◇
BEN 2441	A	◇	BEN 2962	A	+
BEN 2443	A	◇	BEN 2963	A	◇
BEN 2447	A	◇	BEN 2964	A	◇
BEN 2448	B2	◇	BEN 2965	A	◇
BEN 2452	A	+	BEN 2966	A	◇
BEN 2455	A	◇	BEN 2967	A	◇
BEN 2470	A	+	BEN 2968	A	◇
BEN 2940	A	◇	BEN 2969	A	◇
BEN 2943	A	◇	BEN 2970	D	
BEN 2944	A	◇	BEN 2971	A	◇
BEN 2945	A	◇	BEN 2972	D	
BEN 2946	A	+	BEN 2973	A	◇
BEN 2947	A	+	BEN 2974	A	◇
BEN 2948	A	◇	BEN 2975	A	+
BEN 2949	B2		BEN 2978	A	+
BEN 2950	B2		BEN 2980	A	◇
BEN 2954	A	◇	BEN 2981	A	+
BEN 2955	A	◇	BEN 2982	A	+

ANNEXE IV

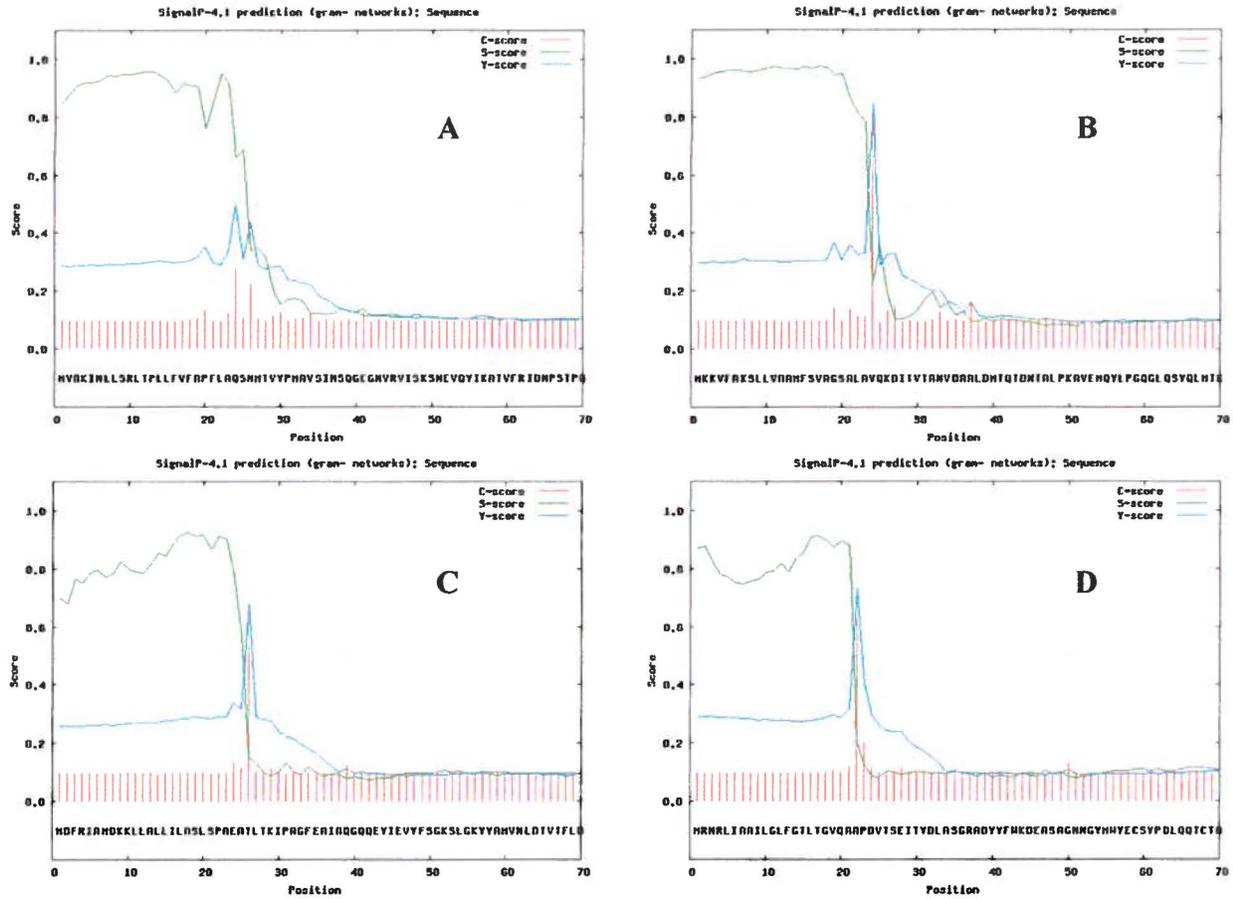


Figure 35 : Prédiction des peptides signaux pour les protéines EfoABCD

Utilisation du logiciel en ligne SignalP 4.1, à l'adresse <http://www.cbs.dtu.dk/services/SignalP/> (376). **A** : Prédiction du peptide signal pour la protéine EfoA, entre les positions 23 et 24 (entre les aa A et Q). **B** : Prédiction du peptide signal pour la protéine EfoB, entre les positions 23 et 24 (entre les aa A et V). **C** : Prédiction du peptide signal pour la protéine EfoC, entre les positions 25 et 26 (entre les aa A et T). **D** : Prédiction du peptide signal pour la protéine EfoD, entre les positions 21 et 22 (entre les aa A et A).

ANNEXE V

>ECOR72

Length of sequence- 200
Threshold for promoters - 0.20
Number of predicted promoters - 1
Promoter Pos: 166 LDF- 6.23
-10 box at pos. 151 ATCTCTAAT Score 38
-35 box at pos. 131 TTGACT Score 61

Oligonucleotides from known TF binding sites:

For promoter at 166:

arcA:	TTAACTAA	at position	108	Score -	12
argR:	TTTTTTAT	at position	122	Score -	13
rpoD16:	TTTTTATA	at position	123	Score -	13

>IAI39

Length of sequence- 200
Threshold for promoters - 0.20
Number of predicted promoters - 1
Promoter Pos: 166 LDF- 7.99
-10 box at pos. 151 ACATCTAAT Score 25
-35 box at pos. 131 TTGACT Score 61

Oligonucleotides from known TF binding sites:

For promoter at 166:

arcA:	TTAACTAA	at position	108	Score -	12
rpoD17:	AATAAATA	at position	140	Score -	11
fnr:	ATAAATAT	at position	141	Score -	9
dnaA:	TATACACA	at position	146	Score -	14
metJ:	ACATCTAA	at position	151	Score -	12
ntrC:	TGCACTAA	at position	159	Score -	8
fnr:	TTTTTTGA	at position	168	Score -	9

>536

Length of sequence- 200

Threshold for promoters - 0.20
 Number of predicted promoters - 1
 Promoter Pos: 102 LDF- 9.28
 -10 box at pos. 87 TTGTATTTT Score 53
 -35 box at pos. 62 GTGAAA Score 27

Oligonucleotides from known TF binding sites:

For promoter at 102:

argR2:	CATATTTT	at position	71	Score -	8
lrp:	TATTTTTT	at position	90	Score -	11
lrp:	ATTTTTTT	at position	91	Score -	11
fnr:	TTTTTTGA	at position	93	Score -	9
rpoD17:	AATAAATA	at position	102	Score -	11
lexA:	ATAAATAA	at position	103	Score -	14
rpoD16:	AAATAATT	at position	105	Score -	15
deoR:	AATTTTAT	at position	109	Score -	8
fis:	TTATCTAA	at position	113	Score -	11

Résultat de l'analyse du promoteur du cluster *efo*, chez les souches ECOR72, IA139 et 536, avec le logiciel en ligne BPRM de SoftBerry (<http://linux1.softberry.com/berry.phtml>). Plusieurs séquences de régulation par la protéine ArgR ont été révélées (notées argR et argR2).

ANNEXE VI

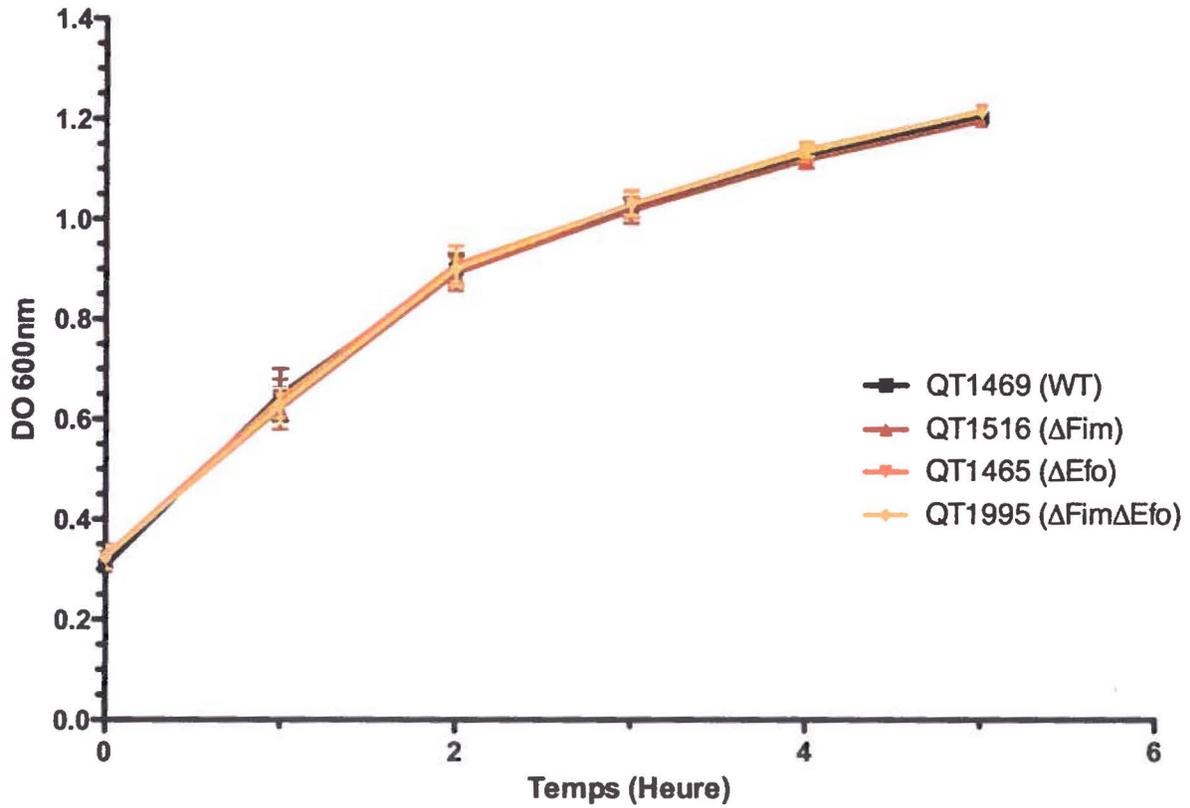


Figure 36 : Courbe de croissance de ECOR72 et ses mutants pour les fimbriae Type 1 et Efo
Aucune différence n'est observée entre les quatre souches pour leur capacité de croissance en milieu riche.