



HAL
open science

Cystites récidivantes à *Escherichia coli*: facteurs microbiens de la persistance

Nicolas Vautrin

► **To cite this version:**

Nicolas Vautrin. Cystites récidivantes à *Escherichia coli*: facteurs microbiens de la persistance. Médecine humaine et pathologie. Normandie Université, 2022. Français. NNT : 2022NORMR087 . tel-04033771

HAL Id: tel-04033771

<https://theses.hal.science/tel-04033771>

Submitted on 17 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité SCIENCES DE LA VIE ET DE LA SANTE

Préparée au sein de l'Université de Rouen Normandie

Cystites récidivantes à Escherichia coli : facteurs microbiens de la persistance et nouvelles cibles thérapeutiques.

Présentée et soutenue par
NICOLAS VAUTRIN

**Thèse soutenue le 09/12/2022
devant le jury composé de**

MME LAURENCE ARMAND-LEFEVRE	PROFESSEUR DES UNIV - PRATICIEN HOSP., UNIVERSITE PARIS 7 PARIS DIDEROT	Rapporteur du jury
M. NICOLAS BARNICH	PROFESSEUR DES UNIVERSITES, UNIVERSITE CLERMONT AUVERGNE CLERMONT AUVERGNE	Rapporteur du jury
M. THIERRY BERTHE	MAITRE DE CONFERENCES, Université de Rouen Normandie	Membre du jury
MME NATHALIE CAROFF	PROFESSEUR DES UNIVERSITES, UNIVERSITE NANTES	Membre du jury
MME MARTINE PESTEL-CARON	PUPH DES DISCIPLINES PHARMACEUTIQUES, Université de Rouen Normandie	Directeur de thèse
M. FRANÇOIS CARON	PROFESSEUR DES UNIV - PRATICIEN HOSP., Université de Rouen Normandie	Co-directeur de thèse

Thèse dirigée par MARTINE PESTEL-CARON (Dynamique Microbienne associée aux Infections Urinaires et Respiratoires) et FRANÇOIS CARON (Dynamique Microbienne associée aux Infections Urinaires et Respiratoires)

Remerciements

À *Madame le professeur Laurence Armand-Lefevre*, À *Monsieur le professeur Nicolas Barnich*. Je vous remercie de l'honneur que vous me faites en ayant accepté d'être rapporteur de mon travail de doctorat. Veuillez trouver ici le témoignage de mes remerciements les plus sincères.

À *Madame le professeur Nathalie Caroff*. Je vous remercie doublement d'avoir accepté de faire partie de mon jury de thèse et de mon comité de suivi individuel. Nos échanges étaient toujours très intéressants et m'ont permis de prendre du recul vis-à-vis de ce travail.

À *Monsieur le docteur Thierry Berthe*. Vous me faites l'honneur d'examiner aujourd'hui ce travail de thèse après m'avoir enseigné la microbiologie pendant trois années à l'université. Je vous remercie pour votre gentillesse et pour avoir accepté d'apporter votre expertise scientifique à ce travail.

À *Madame le professeur Martine Pestel-Caron*. Je vous suis extrêmement redevable pour tout ce que vous avez fait pour moi, non seulement ces dernières semaines, mais tout au long de mon cursus universitaire. Ces quelques lignes ne suffiront pas à vous témoigner toute la reconnaissance que j'ai à votre égard. Vous m'avez fait confiance, confié un projet de recherche et ce depuis mon Master 1, et vous n'avez jamais cessé de croire en moi. Du plus profond de mon cœur je souhaite vous dire, merci pour tout ce que vous avez fait pour moi.

À *Monsieur le professeur François Caron*. Je vous remercie d'avoir accepté de co-diriger cette thèse. Votre regard clinique et votre culture impressionnante ont été absolument indispensable à la réalisation de ces travaux. Vous avez toujours su rappeler que si nous faisons de la recherche, c'est avant tout, pour les patients ! Je souhaite également vous remercier pour votre bonne humeur et votre humour à toute épreuve, qui a été une bouffée d'oxygène pendant ces longs mois de rédaction.

À *Monsieur le professeur Jean-Christophe Plantier*. Je vous remercie pour votre accueil au sein du laboratoire Dynamicure où vous avez toujours eu à cœur de valoriser les étudiants qui passent par votre laboratoire.

À *mes collègues et amis du laboratoire Dynamicure (Ex : GRAM 2.0. d'où votre surnom de gramy's !)*

À *Sandrine*, je pense comme pour Martine ne pas être capable de t'exprimer toute ma gratitude. Tu es non seulement une personne impressionnante dans le domaine professionnel mais également quelqu'un d'incroyable sur le plan personnel. Ta gentillesse, ton sourire et ta bienveillance ont toujours su me redonner foi en moi dans les moments de doute. De notre passion commune pour les lectures d'antibiogrammes est né une amitié que j'espère voir durer longtemps.

À *Agnès*, je pense n'avoir jamais rencontré quelqu'un d'une douceur et d'une gentillesse pareille ! Toi et moi avons partagé les mêmes ennuis de « patientes », « nombre de souches », « milieu urinaire » et « statistiques sous R ». Rien ne peut avoir raison de ton sourire et de ta bonne humeur qui sont autant de choses qui savaient me remotiver.

À *Marie*, ma super colloc' de bureau. Ne nous voilons pas la face, sans tes compétences en bioinformatique et ta bonne humeur, j'aurais déjà renoncé depuis longtemps face à la complexité du génome de *Escherichia coli*.

À *Kévin*, ta capacité à avoir toujours un coup d'avance sur le projet en cours et à toujours proposer des projets innovants ne cessera jamais de m'étonner. Sans toi, ce projet de recherche n'existerait pas.

À *Stéphanie*, il n'y a aucun doute sur le fait que sans toi, tout le monde serait perdu ! Merci pour ton soutien, ta rigueur, ta gentillesse et toute l'aide que tu as su m'apporter tout au long de ce projet.

À *Alice et Élodie*, merci pour ces pauses « café » dans votre bureau à discuter de tout et de rien. Merci pour votre gentillesse et votre bienveillance.

À mes collègues de laboratoire, *Laurie, Camille, Vincent, Chervin* merci pour votre bonne humeur contagieuse et votre bienveillance.

Merci également à *François et Simon Le Hello* pour votre aide sur la partie génomique de ce travail de doctorat.

Je glisse ici un petit mot d'encouragement pour la relève, *Alice et Caroline*, soyez les bienvenus dans l'équipe et merci pour vos encouragements de ces dernières semaines !

À tout le personnel de l'UMR Inserm 1073 ADEN, merci pour votre accueil et votre gentillesse.

À *David*, merci beaucoup pour ton aide dans ce travail de doctorat. Ton expertise en culture cellulaire a été indispensable pour mener à bien ce projet. Merci pour ta disponibilité, ta bienveillance et ta gentillesse (même quand je cassais le matériel !!!).

À mes amis.

À *Maxime*, nous avons commencé en tant que collègues et en l'espace de quelques mois, nous avons noué une relation amicale très solide. Merci pour toute l'aide que tu m'as apporté dans ce travail (notamment pour les analyses avec R !). Merci pour tous les bons moments que nous avons passés ensemble et pour tous ceux à venir.

À *Solveig*, merci pour ta gentillesse, ta bienveillance et tous les bons moments, passés et à venir.

À *Julien*, tu m'as connu enfant, et d'aussi loin que je me souviens tu as toujours été là pour moi. Je te remercie du fond du cœur. Merci à *Mélanie et Pauline (dit Paupau)* pour tous ces bons moments.

À *Delphine, Clément et Thibaut* merci pour ces bons moments passés ensemble.

À *ma famille*.

À *mes parents*, merci pour tout. Vous avez toujours cru en moi et vous m'avez toujours fait confiance. Vous avez su m'aiguiller sans pour autant interférer avec mes choix. Je n'aurai pu rêver meilleurs modèles. Je vous aime infiniment.

À *mon frère*, le premier docteur de la famille, merci pour tout. Tu n'es pas seulement mon frère, tu es mon confident, mon meilleur ami, mon modèle. Ta gentillesse, ton humour et ta persévérance à tout épreuve sont autant de sources d'inspiration pour moi. À *Marion*, ma belle-sœur merci pour ta gentillesse et ta bienveillance.

À *Kiwi*, mon petit chat, qui a su, à sa façon, me remonter le moral et me sentir moins seul pendant cette période de rédaction.

À ma belle-famille, *Isabelle, Christophe, Loïc, Marie-Hélène*. Merci pour votre gentillesse et votre bienveillance à toute épreuve.

À *tous les membres de ma famille*, je ne vous dirai jamais assez combien je vous aime.

À *ma femme, Manon* merci pour tout. Sans toi je n'aurais jamais eu la force d'aller aussi loin. Merci pour ton soutien, ta patience, pour avoir relu ce manuscrit. Ta bonne humeur et ta gentillesse ont su éclairer ma route quand j'étais perdu.

Table des matières

Remerciements	2
Table des matières	5
Liste des figures	8
Liste des tableaux	10
Liste des abréviations	11
Contexte de la thèse.....	12
Introduction.....	13
I. <i>Escherichia coli</i>	13
I.1. Historique	13
I.2. Principaux caractères microbiologiques.....	13
I.3. Données génomiques	14
I.4. <i>E. coli</i> : une espèce versatile.....	15
I. 4. 1. Une espèce commensale ou mutualiste ?	15
I. 4. 2. Une espèce pathogène	16
I.5. Typage moléculaire	18
I. 5. 1. Ère pré-moléculaire	18
I. 5. 2. Phylogroupage par réaction de polymérisation en chaîne (PCR)	20
I. 5. 3. <i>Multi-Locus Sequence Typing</i> (MLST).....	23
I. 5. 4. <i>CH typing</i>	24
I. 5. 5. Méthodes basées sur le séquençage à haut débit.....	25
I.6. Structure de l'espèce.....	27
II. Infections urinaires.....	30
II.1. Appareil urinaire.....	30
II. 1. 1. Généralités.....	30
II. 1. 2. Histologie de la vessie.....	31
II. 1. 3. Composition de l'urine.....	33
II.2. Microbiote urinaire	34

II.3.	Généralités sur les infections urinaires	36
II. 3. 1.	Définitions et diagnostic	36
II. 3. 2.	Epidémiologie et étiologie des infections urinaires.....	37
II. 3. 3.	Physiopathologie des IU.....	38
II.4.	<i>E. coli</i> uropathogène (UPEC).....	39
II. 4. 1.	Définition du pathovar.....	39
II. 4. 2.	Épidémiologie de la résistance aux antibiotiques des UPEC	40
II. 4. 3.	Facteurs de virulence.....	42
II. 4. 4.	Notion de souches de <i>E. coli</i> pathogènes hybrides	49
III.	Cystites récurrentes à <i>E. coli</i>	51
III.1.	Fardeau des cystites récurrentes.....	51
III.2.	Rechutes ou réinfections ?	51
III.3.	Facteurs de risque de récurrence.....	52
III.4.	Physiopathologie des cystites récurrentes : hypothèses actuelles	52
III. 4. 1.	Persistance vésicale	53
III. 4. 2.	Réservoir digestif des UPEC	58
III. 4. 3.	Réservoir vaginal d'UPEC	60
III. 4. 4.	Influence du microbiote urinaire sur les IU récurrentes.....	61
	Travaux personnels	65
I.	Première étude : intérêt de l'antibiogramme et du <i>CH typing</i> pour l'identification des rechutes de cystite à UPEC.....	65
I.1.	Contextualisation du travail	65
I.2.	Publication n°1	67
I.3.	Principaux résultats	68
I. 3. 1.	Description de la cohorte VITALE.....	68
I. 3. 2.	Diversité génétique des isolats d'UPEC	68
I. 3. 3.	Prévalence des rechutes – apport du <i>CH typing</i> et de l'antibiotype dans l'identification des rechutes.....	68

II.	Second travail personnel	70
II.1.	Contextualisation de l'étude	70
II.2.	Publication n°2	72
II.3.	Principaux résultats	73
II. 3. 1.	Caractéristiques génomiques des souches d'UPEC à l'origine de rechutes de cystites	73
II. 3. 1.	Caractéristiques phénotypiques des souches d'UPEC à l'origine de rechutes de cystites	73
III.	Discussion générale	75
III.1.	Constitution de la cohorte.....	75
III.2.	Evaluation de la diversité génotypique des souches d'UPEC.....	76
III.3.	Identification des rechutes et évaluation de leur place dans le phénomène de récurrence	76
III.4.	Comparaison génomique et phénotypique de souches récidivantes et sporadiques	77
III. 4. 1.	Recherche de déterminants génomiques associés à la récurrence	77
III. 4. 2.	Recherche de déterminants phénotypiques associés à la récurrence	78
III. 4. 3.	Suivi longitudinal des souches responsables de rechute.....	80
IV.	Conclusion générale	82
	Bibliographie.....	83

Liste des figures

Figure 1 : Structure de la paroi d'une bactérie à Gram négatif.....	14
Figure 2 : Structure et plasticité du génome de <i>E. coli</i>	15
Figure 3 : Phylogénie de 72 souches de <i>E. coli</i> construite à partir des données de <i>Multi-locus Enzyme Electrophoresis</i> (MLEE).....	20
Figure 4 : Profils obtenus à l'aide de la PCR triplex décrite par Clermont <i>et al.</i> (A) et démarche dichotomique permettant d'assigner une souche à l'un des quatre principaux phylogroupes de <i>E. coli</i> à partir du profil obtenu (B)	21
Figure 5 : Profils obtenus avec la PCR quadruple permettant d'assigner une souche à l'un des huit phylogroupes de <i>E. coli</i> à l'aide d'une PCR quadruple.....	22
Figure 6 : Localisation génomique des sept gènes de ménage utilisés pour le schéma MLST de <i>E. coli</i>	23
Figure 7 : Déroulement du séquençage d'un génome bactérien par NGS.....	25
Figure 8 : Phylogénie du genre <i>Escherichia</i> basée sur les SNPs identifiés entre 72 souches appartenant à ce genre	28
Figure 9 : Contribution des transferts horizontaux dans l'émergence des différents pathovars de <i>E. coli</i>	29
Figure 10 : Arbre urinaire féminin (A) et masculin (B)	30
Figure 11 : Organisation histologique de la paroi de la vessie.....	31
Figure 12 : Structure détaillée de l'urothélium vésical	32
Figure 13 : (A) Observation des structures cristallines d'uroplakine (UP) à la surface de cellules urothéliales murine en microscopie électronique à balayage. (B) Schéma représentant l'hétérotétramère d'uroplakine assemblé à la membrane plasmique des cellules urothéliales.	32
Figure 14 : Composés majoritaires (> 1 mM) présents dans l'urine	33
Figure 15 : Diagramme de Venn comparant les espèces composant le microbiote urinaire, intestinal et vaginal	35
Figure 16 : Épidémiologie des infections urinaires communautaires	37
Figure 17 : Épidémiologie des IU associées aux soins en Amérique du Nord et en Europe	38
Figure 18 : Physiopathologie des infections urinaires causées par un UPEC	39
Figure 19 : Principaux facteurs de virulence des UPEC classés selon leur fonction.....	42
Figure 20 : Structure du <i>pilus</i> de type I	44
Figure 21 : Représentation d'une bactérie remontant contre un flux antagoniste	45
Figure 22 : Fonctionnement des deux types de systèmes de captation du fer chez les UPEC	47

Figure 23 : Invasion d'une cellule urothéliale par des UPEC observée en microscopie électronique à transmission dans un modèle d'IU murin	53
Figure 24 : Filament bactérien interagissant avec deux cellules adjacentes observé en microscopie électronique à balayage	54
Figure 25 : Bactérie en cours d'internalisation par une cellule urothéliale observée en microscopie électronique à balayage	55
Figure 26 : Conséquences de l'interaction FimH – Uroplakine	56
Figure 27 : Mécanisme d'internalisation d'un UPEC	57
Figure 28 : Formation des communautés bactériennes intracellulaires et des réservoirs quiescents intracellulaires	58
Figure 29 : Groupes de bactéries dont l'abondance relative varie significativement dans le microbiote des patientes atteintes de cystites récidivantes par rapport à un groupe contrôle	62
Figure 30 : Implication des trois microbiotes dans la pathogénèse des infections urinaires.	63

Liste des tableaux

Tableau 1 : Liste des pathovars de <i>E. coli</i> pathogènes pour l'Homme.....	17
Tableau 2 : Liste des milieux urinaires artificiels publiés	34
Tableau 3 : Seuils de significativité de bactériurie selon l'espèce bactérienne et le sexe du patient ..	36
Tableau 4 : Gènes de virulence ciblés par des PCRs permettant d'identifier les différents pathovars d'InPEC.....	40
Tableau 5 : Données de résistance des UPEC aux principaux antibiotiques constituant l'arsenal thérapeutique contre les IU	41
Tableau 6 : Liste des gènes ou opérons codant des adhésines identifiés chez les UPEC.....	43
Tableau 7 : Liste des toxines pouvant être produites par les UPEC.....	46
Tableau 8 : Liste des pathovars hétéro-pathogènes et hybrides ainsi que leurs principales caractéristiques	50

Liste des abréviations

CC : Complexe clonal

CR : Cystite récidivante

cgMLST : *Core genome multi-locus sequence typing*

EHEC : *E. coli* entérohémorragiques

EIEC : *E. coli* entéro-invasif

ExPEC : *Extra-intestinal pathogenic E. coli* – *E. coli* pathogènes extra-intestinaux

Gb : Gigabase (10^9 bases)

IBC : Communauté bactérienne intracellulaire

InPEC : *Intestinal pathogenic E. coli* – *E. coli* pathogènes intestinaux

IU : Infection urinaire

Kb : Kilobase (10^3 bases)

LPS : Lipopolysaccharide

Mb : Mégabase (10^6 bases)

MEB : Microscopie électronique à balayage

MET : Microscopie électronique à transmission

MLEE : *Multi-Locus Enzyme Electrophoresis*

MLST : *Multi-Locus Sequence Typing*

mM : Milimolaire (10^{-3} mol.L⁻¹)

MMU : milieu mimant l'urine

NCBI : *National Center for Biotechnology Information*

NGS : *Next generation sequencing*

ONT : *Oxford Nanopore Technologies*

Pb : Paires de bases

PCR : Réaction de polymérisation en chaîne

PGFE : Électrophorèse en champ pulsé

PI3K : *Phosphoinositide-3-kinase*

QIR : Réservoir intracellulaire quiescent

SI : Système immunitaire

SNP : *Single nucleotide polymorphism*

ST : *Sequence type*

UFC : Unité formant des colonies

Contexte de la thèse

Ce travail de thèse s'inscrit dans l'axe de recherche « infections urinaires » du laboratoire DYNAMICURE (Dynamique microbienne associée aux infections urinaires et respiratoires), Inserm UMR 1311. Un des objectifs de cet axe est d'explorer la physiopathologie des cystites récidivantes, d'étudier les conséquences des traitements antibiotiques itératifs sur la santé des patients, et de développer des alternatives à ces derniers.

Dans ce contexte, l'UMR 1311 étudie l'agent étiologique principal de ces cystites récidivantes : *Escherichia coli*, et explore les déterminants microbiens associés à la récurrence dans le but de mieux la comprendre puis la prévenir.

La première partie de ce manuscrit consiste en un état des lieux des connaissances concernant *E. coli*, en se focalisant sur les souches dites « uropathogènes », sur les infections urinaires et la physiopathologie des cystites récidivantes.

La seconde partie est une présentation des résultats de mes travaux personnels dont les objectifs étaient : (i) l'étude de la dynamique des cystites récidivantes à *E. coli* et (ii) la recherche de marqueurs génomiques ou phénotypiques associés à la récurrence.

Introduction

I. *Escherichia coli*

I.1. Historique

Escherichia coli a été isolé pour la première fois à partir d'un prélèvement de selles de nouveau-né par le microbiologiste et pédiatre Theodor Escherich en 1884 (Escherich, 1988). Initialement nommé *Bacterium coli commune* par ce dernier, ce micro-organisme fut renommé *Escherichia coli* en 1958 (Henry, 2015). Étant facile à isoler, cultiver et manipuler, il s'est rapidement imposé en tant qu'organisme modèle pour l'étude des procaryotes. *E. coli* est aujourd'hui incontournable en biologie et les travaux le concernant ont été à l'origine d'avancées majeures non seulement en physiologie bactérienne, mais également en biologie moléculaire, médecine et biotechnologies (Blount, 2015).

I.2. Principaux caractères microbiologiques

L'espèce *E. coli* appartient au phylum des *Proteobacteria* (renommé *Pseudomonadota* depuis 2021) et à la famille des *Enterobacteriaceae* (Oren and Garrity, 2021). À ce titre, *E. coli* se présente sous la forme de bacille à Gram négatif et ne possède pas de cytochrome C oxydase. Cette espèce est mobile grâce à une ciliature péritriche. Elle est aérobie-anaérobie facultative sans exigence culturale et sa température optimale de croissance se situe aux alentours de 37°C (bactérie mésophile). La capacité de *E. coli* à fermenter le lactose, à produire de l'indole via une activité tryptophanase ainsi que son activité β -D-glucuronidase le distinguent des autres *Enterobacteriaceae*. Cette dernière caractéristique est exploitée par certains milieux chromogènes pour l'identification rapide de cette espèce.

La paroi des bactéries à Gram négatif est composée de deux membranes lipidiques : la membrane interne délimitant la cellule, et la membrane externe délimitant le périplasme. Au sein du périplasme se trouve une fine couche de peptidoglycane rigidifiant la paroi. Le lipopolysaccharide (LPS), structure caractéristique des bactéries à Gram négatif, est ancré dans la membrane externe. Certaines souches de *E. coli* sont capsulées ; elles présentent alors, en plus des antigènes flagellaire (H) et somatique (O), un 3^{ème} antigène de surface, le polysaccharide constitutif de la capsule (K) (Sande and Whitfield, 2021) (Figure 1).

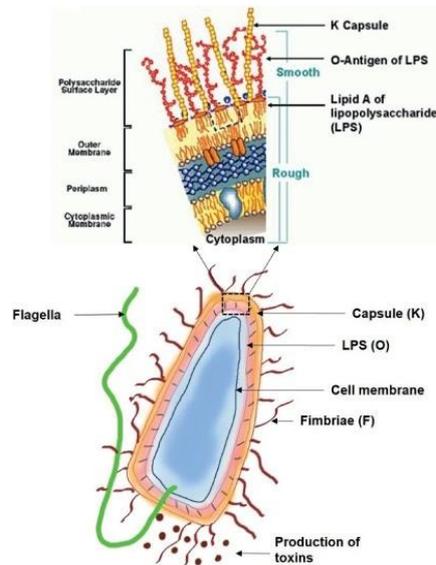


Figure 1 : Structure de la paroi d'une bactérie à Gram négatif

À la surface d'une bactérie à Gram négatif, trois structures antigéniques peuvent être retrouvées : l'antigène flagellaire (H), l'antigène somatique (O) et l'antigène capsulaire (K). Le détail de la paroi bactérienne est représenté dans la partie haute de la figure (Reis et al., 2021).

1.3. Données génomiques

En août 2022, 31 186 assemblages de génomes de *E. coli* étaient disponibles sur la banque de données *Genome* du *National Center for Biotechnology Information* (NCBI). Cependant, l'ensemble des réplicons n'a été séquencé entièrement que pour 7% d'entre eux. D'après ces données, la taille du génome de *E. coli* est comprise entre 4,17 et 6,16 mégabases (Mb), avec une taille moyenne estimée à 5,11Mb et un pourcentage de guanine et cytosine moyen (GC%) de 50,6%. Ce génome est constitué d'un chromosome circulaire accompagné ou non de plasmides. En moyenne, un génome de *E. coli* contiendrait 4 730 gènes codant des protéines (<https://www.ncbi.nlm.nih.gov/genome/167>).

Avec l'accumulation de génomes séquencés, il est devenu possible d'établir le *pan genome* de l'espèce (ensemble des gènes identifiés au sein de l'espèce). Parmi ces gènes, certains peuvent être retrouvés chez toutes les souches ; ils constituent le *core genome* (Figure 2). Les autres appartiennent au génome dit « accessoire ». Parmi ces derniers, certains sont spécifiques d'un environnement donné (urine vs intestin), d'un type d'interaction avec l'hôte (commensalisme versus pathogénie) etc. En 2009, une étude portant sur un corpus de 14 souches de *E. coli* a établi un *pan genome* de 18 000 gènes, pour un *core genome* d'environ 2 000 gènes (Touchon et al., 2009) (Figure 2). Cependant, la taille de ces deux ensembles est fortement influencée par le nombre de souches étudiées et les méthodes mises en œuvre. En effet, une étude datant de 2019 portant sur 491 génomes a mis en évidence un *pan genome* de 43 415 gènes pour un *core genome* de seulement 867 gènes (soit 2%) (Yang et al., 2019).

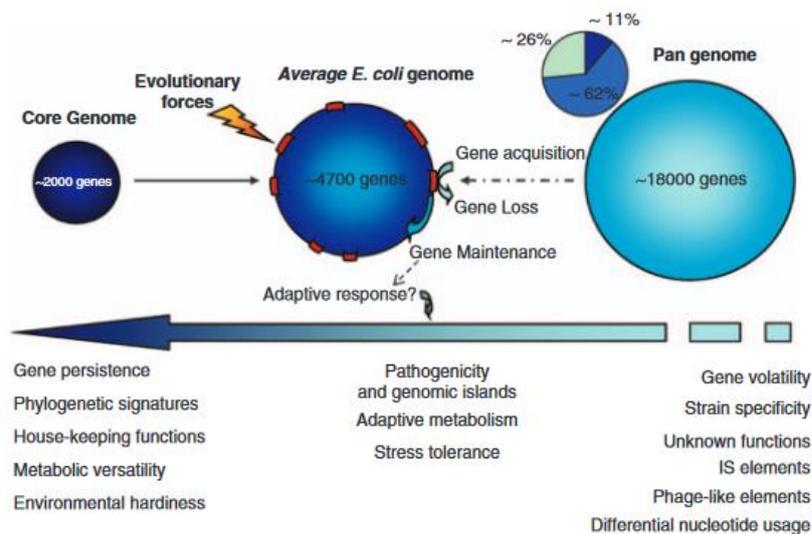


Figure 2 : Structure et plasticité du génome de *E. coli*

Le génome d'une souche de *E. coli* comporte en moyenne 4 700 gènes. Touchon et al. ont identifié 18 000 gènes dans l'ensemble des génomes qu'ils ont étudié (pan genome). Parmi eux, 2 000 gènes (soit 11% du pan genome) ont été identifiés chez toutes les souches (core genome). En moyenne, 2 000 gènes sont partagés entre toutes les souches de *E. coli*. Ces gènes constituent le core genome. Au sein du pan genome, 62% des gènes ont été identifiés dans au moins 90% des souches constituant le génome dit persistant et 26% des gènes ont été identifiés chez moins de 20% des souches constituant le génome dit volatile (Touchon et al., 2009). La flèche représente un gradient de caractéristiques associées aux différents ensembles de gènes (van Elsas et al., 2011).

I.4. *E. coli* : une espèce versatile

I.4.1. Une espèce commensale ou mutualiste ?

E. coli est la principale espèce aérobie identifiée dans le microbiote intestinal humain et est l'une des premières à coloniser l'intestin du nouveau-né (Mueller et al., 2015). Au sein du tractus gastro-intestinal, *E. coli* représente entre 0,1 et 5% de la population bactérienne totale, avec une concentration variant de 10^7 à 10^9 unités formant des colonies (UFC) par gramme de matière fécale (Foster-Nyarko and Pallen, 2022 ; Tenailon et al., 2010).

Le terme commensalisme désigne une interaction durable entre des individus d'espèces différentes, où l'un des partenaires retire un bénéfice de l'association, tandis que l'autre n'y trouve ni avantage ni véritable inconvénient. Le mutualisme quant à lui, se définit comme une association dans laquelle les deux partenaires tirent un bénéfice réciproque (Casadevall and Pirofski, 2000).

E. coli est souvent qualifié de bactérie commensale (Lescat et al., 2013 ; Nielsen et al., 2018 ; Rasko et al., 2008 ; Tenailon et al., 2010). Or, s'il est vrai qu'il trouve chez son hôte un environnement stable et des nutriments, la relation qu'il entretient avec ce dernier se rapproche davantage du mutualisme. En effet, *E. coli* jouerait un rôle dit de « flore de barrière » en empêchant les organismes pathogènes de coloniser le tube digestif. De plus, il permettrait de favoriser l'implantation de la flore anaérobie chez le nouveau-né, en consommant l' O_2 et pourrait ainsi contribuer à la mise en place d'une

anaérobiose. Enfin, *E. coli* fournirait à son hôte certaines molécules essentielles dont la vitamine K (Foster-Nyarko and Pallen, 2022).

I. 4. 2. Une espèce pathogène

Bien que *E. coli* soit une espèce mutualiste du tube digestif, certaines souches peuvent être responsables d'infections. Ces infections sont de localisations diverses, intestinales ou extra-intestinales (**Tableau 1**). La présence de facteurs de virulence spécifiques (toxines, facteurs de colonisation, etc.), la symptomatologie associée à l'infection, ainsi que l'organe cible, sont à la base d'une classification des souches en pathovars. Il existe à ce jour 10 pathovars « humains » de *E. coli* classés en 2 groupes : les *intestinal pathogenic E. coli* (InPEC) et les *Extra-intestinal pathogenic E. coli* (ExPEC) (Denamur et al., 2021).

À ce jour, sept pathovars d'InPEC ont été décrits (**Tableau 1**). Ils sont définis par leur mécanisme de pathogénicité et par leur arsenal de facteurs de virulence. Certains de ces pathovars, comme celui des *E. coli* entérohémorragiques (EHEC), sont étroitement surveillés du fait de la gravité des infections qu'ils causent, bien que celles-ci soient rares. De plus, les EHEC sont fréquemment associés à des toxi-infections alimentaires collectives. Ces dernières faisant l'objet d'une déclaration obligatoire, des chiffres relativement précis sont disponibles (Santé publique France, 2021).

Les souches d'ExPEC sont classées en trois pathovars (**Tableau 1**) : les *E. coli* uropathogènes (UPEC), responsables d'infections urinaires et potentiellement de sepsis, les *E. coli* responsables de méningites chez le nouveau-né (NMEC) et, les souches de *E. coli* responsables de pneumonies (Denamur et al., 2021). Les souches d'ExPEC sont responsables de la majorité des infections à *E. coli*. En effet, cette bactérie est responsable de 80% des infections urinaires communautaires, et de respectivement 50% et 23,6% des infections associées aux soins, aux États-Unis et en France (Desvaux et al., 2020 ; Russo, 2003 ; Santé publique France, 2017).

Tableau 1 : Liste des pathovars de *E. coli* pathogènes pour l'Homme

(Croxen and Finlay, 2010 ; Denamur et al., 2021).

Pathovar	Acronyme	Principaux tableaux cliniques	Principaux organes cibles	Principaux facteurs de virulence
<i>Extra-Intestinal Pathogenic Escherichia coli (ExPEC)</i>				
<i>Uropathogenic E. coli</i>	UPEC	Cystite, pyélonéphrite, urosepsis	Vessie, rein	Adhésines (<i>fim</i> , <i>pap</i>) toxines (<i>hly</i> , <i>cnf1</i>), sidérophores (<i>iro</i> , <i>fyu</i>), capsule
<i>Newborn meningitis E. coli</i>	NMEC	Méningite, sepsis	Méninges, LCR, sang	Capsule
<i>Pneumonia-associated E. coli</i>		Pneumonie	Poumon	Hémolysine (<i>hly</i>), adhésine (<i>sfa</i>)
<i>Intestinal Pathogenic Escherichia coli (InPEC)</i>				
<i>Enterohemorrhagic E. coli</i> ou <i>Shiga toxin-producing E. coli</i>	EHEC ou STEC	Diarrhée hémorragique, syndrome hémolytique et urémique	Côlon	Shiga-toxines (<i>stx1</i> , <i>2</i>), Intimine (<i>eae</i>) Tir (<i>espE</i>), Entérohémolysine (<i>ehxA</i>)
<i>Enteropathogenic E. coli</i>	EPEC	Diarrhées infantiles	Intestin grêle	Intimine (<i>eae</i>), Tir (<i>espE</i>), Type IV bundle forming pilus (<i>bfp</i>)
<i>Enterotoxigenic E. coli</i>	ETEC	Diarrhée du voyageur, syndrome cholériforme	Intestin grêle	Entérotoxines thermostable (<i>estA</i>) et thermolabile (<i>eltB</i>)
<i>Enteroinvasive E. coli</i>	EIEC	Syndrome dysentérique, diarrhées hémorragiques	Côlon	<i>Invasion plasmid antigens (lpa)</i> , <i>Iron-sulfur cluster (isc)</i> , régulateurs transcriptionnels (<i>vir</i>)
<i>Enteraggregative E. coli</i>	EAEC	Diarrhée du voyageur (syndrome cholériforme)	Côlon	Pili d'adhésion et d'agrégation (<i>aaf</i>), régulateur transcriptionnel AggR
<i>Diffusely adherent E. coli</i>	DAEC	Diarrhées infantiles	Intestin grêle	Adhésines (<i>afa/dra</i>)
<i>Enteraggregative and hemorrhagic E. coli</i>	EAHEC	Diarrhée profuse et hémorragique, syndrome hémolytique et urémique	Côlon	Pili d'adhésion et d'agrégation(<i>aaf</i>), Shiga-toxines (<i>stx 1, 2</i>)

E. coli est au premier rang des bactéries responsables d'infections humaines, les souches d'UPEC étant responsables d'une large majorité d'entre elles. Ainsi, dans les relevés d'activité du CHU de Rouen, *E. coli* est de loin l'espèce bactérienne considérée comme cliniquement significative, avec en moyenne plus de 6 000 souches dédoublonnées (*i.e.* un seul prélèvement par patient pour le même épisode infectieux) relevant de la réalisation d'un antibiogramme dont 85% d'isolats urinaires, 5% d'isolats d'hémocultures et 10% de situations très diverses (péritonites, méningites...) (Caron, 2014). Aux États-Unis, les infections urinaires sont de même considérées comme les plus fréquentes, estimées entre six et huit millions de cas par an (Russo, 2003).

I.5. Typage moléculaire

Comme vu précédemment, *E. coli* est responsable d'infections diverses. Pour suivre l'émergence et la diffusion de clones de virulence et de résistance particulières, il a été nécessaire de développer des méthodes de typage permettant de distinguer les souches de *E. coli* entre elles. Les applications en sont multiples. À un niveau local, le typage ouvre la possibilité d'études de micro-épidémiologie permettant par exemple, la comparaison de souches lors d'investigations d'épidémies hospitalières pour attester de la clonalité. À plus large échelle et à plus long terme, une surveillance macro-épidémiologique permet d'étudier la distribution des clones au niveau mondial, de déterminer la structure d'une espèce bactérienne et son mode d'évolution génomique, et ainsi permet de différencier les espèces clonales (évoluant principalement par mutations) et panmictiques (évoluant principalement par recombinaisons).

Au fil du temps, et avec le développement considérable de la biologie moléculaire et du séquençage d'ADN, les méthodes de typage se sont perfectionnées, permettant de suivre avec plus de précision l'évolution des souches pathogènes de *E. coli* et d'affiner les données relatives à la structure de cette espèce.

I. 5. 1. Ère pré-moléculaire

Sérotypage

Historiquement, le sérotypage par agglutination est l'une des premières méthodes qui a été utilisée pour typer les souches de *E. coli* (Kauffmann, 1947). Le sérotypage est fondé sur l'étude du polymorphisme des 2 antigènes pariétaux H et O et de l'antigène capsulaire K. Cette méthode, basée sur la formation de complexes anticorps-antigènes, présente l'avantage d'être techniquement facile à mettre en œuvre. Elle est cependant laborieuse car manuelle (nécessité de disposer de nombreux anti-*sera*) et source de nombreuses erreurs. De plus, seuls quelques laboratoires possédaient à l'époque, la capacité de typer l'antigène capsulaire K. Pour ces raisons, et grâce à l'avènement de la biologie moléculaire, cette méthode a évolué et aujourd'hui le sérotypage est réalisé par séquençage des gènes

impliqués dans la synthèse des différents antigènes (*Fratamico et al., 2016*). Le sérotypage reste utilisé en France par le centre national de référence des *E. coli* pour l'investigation rapide des cas groupés à EHEC ou *E. coli* entéro-invasif (EIEC).

Multi-Locus Enzyme Electrophoresis (MLEE)

Cette technique contemporaine du sérotypage est basée sur un profil de migration électrophorétique de 20 enzymes codées par des gènes de ménage. En théorie, toute modification dans l'un des gènes devrait modifier le profil de migration, et permettre ainsi de distinguer deux souches. Cette technique a contribué à améliorer la compréhension de la structure de l'espèce *E. coli* en mettant en évidence l'existence de 4 groupes phylogénétiques, ou phylogroupes (A, B1, B2 et D). Un cinquième groupe hétérogène de souches (groupe C) a été identifié par cette méthode mais ne sera réellement défini que plus tard par des méthodes plus résolutive (**Figure 3**) (*Escobar-P et al., 2004 ; Herzer et al., 1990*). Il a par la suite été démontré qu'il existait une association de ces phylogroupes avec la virulence des souches, les UPEC appartenant majoritairement aux phylogroupes B2 et D, alors que la plupart des souches mutualistes appartiennent au groupe A (*Boyd and Hartl, 1998*).

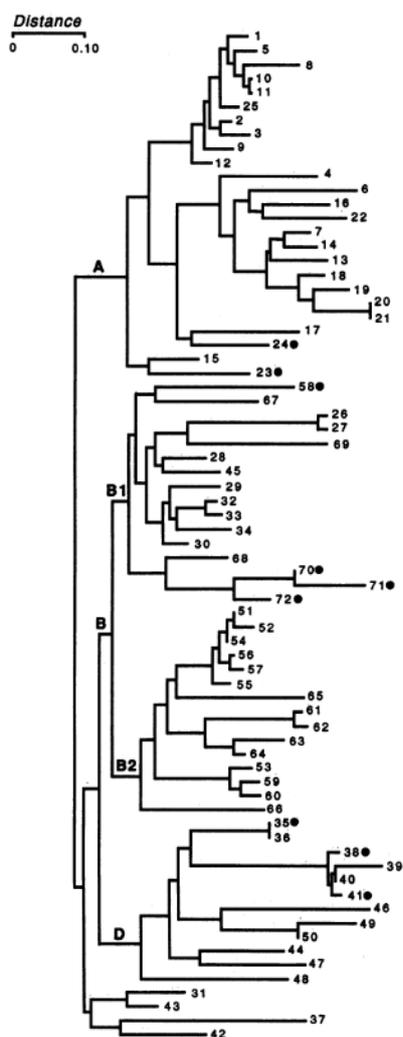


Figure 3 : Phylogénie de 72 souches de *E. coli* construite à partir des données de *Multi-locus Enzyme Electrophoresis* (MLEE)

Les phylogroupes A, B1, B2 et D sont désignés sur l'arbre par leur lettre. Les souches appartenant au groupe C ont été réassignées aux autres groupes (Herzer et al., 1990).

La MLEE présente cependant de nombreux inconvénients. Le plus évident d'entre eux est qu'une mutation pouvant être silencieuse, un changement d'allèle n'engendre pas obligatoirement une modification du profil en MLEE. De plus, cette méthode est chronophage et particulièrement complexe à mettre en œuvre, avec de nombreux facteurs pouvant influencer le résultat final. Il est ainsi difficile de comparer des résultats issus d'expériences indépendantes (Stanley and Wilson, 2003).

1. 5. 2. [Phylogroupage par réaction de polymérisation en chaîne \(PCR\)](#)

PCR triplexe

L'assignation d'une souche à son phylogroupe par MLEE présentant de nombreux inconvénients, une PCR triplexe a été développée en 2000 par Clermont et al., permettant d'attribuer rapidement un isolat de *E. coli* à l'un des quatre phylogroupes A, B1, B2 ou D (Clermont et al., 2000). Cette méthode

se base sur la présence ou l'absence de trois marqueurs génétiques : *chuA*, *yjaA* et TspE4.C2 (**Figure 4**).

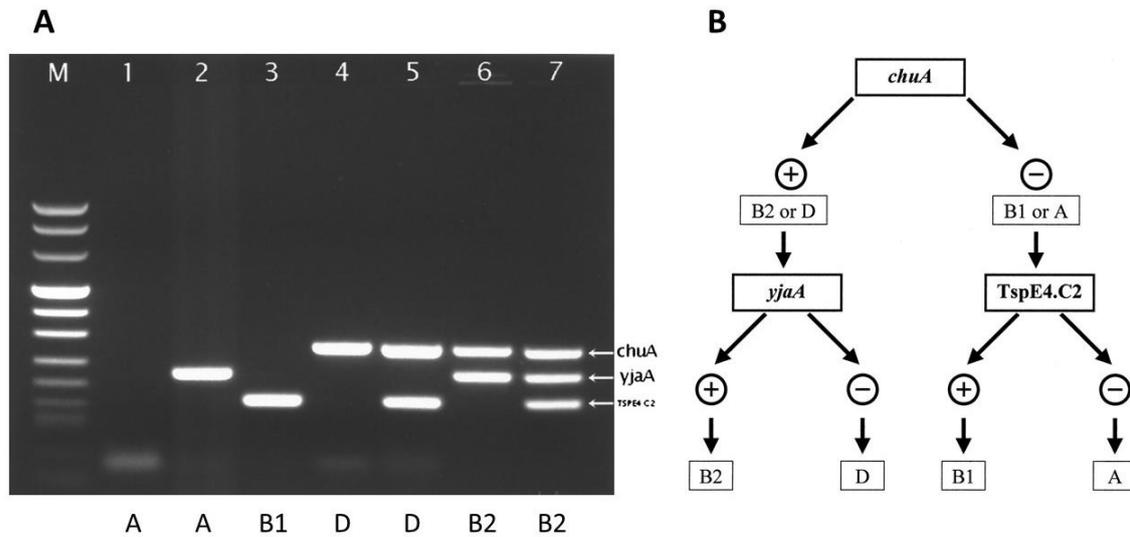


Figure 4 : Profils obtenus à l'aide de la PCR triplexe décrite par Clermont *et al.* (A) et démarche dichotomique permettant d'assigner une souche à l'un des quatre principaux phylogroupes de *E. coli* à partir du profil obtenu (B)

chuA : gène codant le récepteur d'hème *ChuA*, *yjaA* : gène codant la protéine *YjaA* induite par le stress, *TspE4.C2* : Fragment d'ADN anonyme (Clermont *et al.*, 2000).

Le premier d'entre eux, le gène *chuA*, identifié chez la souche entérohémorragique O157:H7 (groupe E), est un gène codant une protéine impliquée dans le transport des hèmes (Torres and Payne, 1997). Ce gène aurait été acquis par les groupes B2 et D très tôt après leur émergence, ce qui expliquerait son absence chez les souches des groupes B1 et A (Lecointre *et al.*, 1998). Le second, *yjaA*, est un gène initialement identifié dans le génome de *E. coli str.* K-12 (Blattner *et al.*, 1997) et dont la fonction n'était pas connue à cette époque. En 2010, une étude a mis en évidence expérimentalement l'implication de la protéine *YjaA*, dans la réponse au stress acide, à l'exposition au peroxyde d'hydrogène, et dans la formation de biofilm (J. Lee *et al.*, 2010). Enfin, *TspE4.C2* correspondait initialement à un fragment d'ADN « anonyme », mis en évidence par la digestion du génome de *E. coli str.* C5 (pathovar des NMEC) par une enzyme de restriction : *Tsp509I* (Bonacorsi *et al.*, 2000). Plus tard, ce fragment sera identifié comme codant une lipase estérase putative (Gordon *et al.*, 2008).

Dans la publication initiale décrivant la PCR triplexe, 230 souches de *E. coli* ont été typées (Clermont *et al.*, 2000). La PCR a démontré une excellente corrélation (228 souches sur 230, 99%) avec les méthodes de référence (MLEE et ribotypage). Plus tard, une étude montrera que 85 à 90% des souches de *E. coli* peuvent être attribuées à un phylogroupe par cette méthode, et qu'elle permet d'assigner correctement une souche à son phylogroupe dans seulement 80 à 85% des cas (comparaison avec les résultats de *Multi-Locus Sequence Typing*) (Gordon *et al.*, 2008).

PCR quadruple

Par la suite, le nombre croissant de données de séquençage multi-*locus* et de données de génomes complets ont permis d'affiner la compréhension de la structure des phylogroupes de *E. coli*, et de définir huit groupes : sept (A, B1, B2, C, D, E, F) appartenant à *E. coli sensu stricto*, tandis que le huitième est le clade cryptique I d'*Escherichia*. Parmi les nouveaux phylogroupes définis, le groupe E correspond à un petit ensemble de souches (dont les souches *E. coli* O157:H7), les souches du groupe F dérivent du groupe B2 et étaient initialement classées dans le groupe D (Jaureguy et al., 2008), et les souches appartenant au groupe C, sont très proches mais distinctes de celles du groupe B1 (Moissenet et al., 2010).

La découverte de ces nouveaux phylogroupes a conduit à l'amélioration de la PCR triple en PCR quadruple par la même équipe en 2013 (Clermont et al., 2013). La principale modification réside dans l'ajout d'un quatrième marqueur : le gène *arpA*, de fonction inconnue. Ce gène permet d'une part de différencier les souches du groupe F de celles du groupe D, et d'autre part d'évaluer la qualité de l'ADN étudié, puisque toutes les souches de *E. coli* et de clade I possèdent au moins l'un des quatre marqueurs (Figure 5). De plus, en s'appuyant sur les données de séquençage disponibles, la séquence des amorces des trois autres marqueurs a été modifiée pour limiter l'impact d'éventuelles mutations sur l'hybridation des amorces.

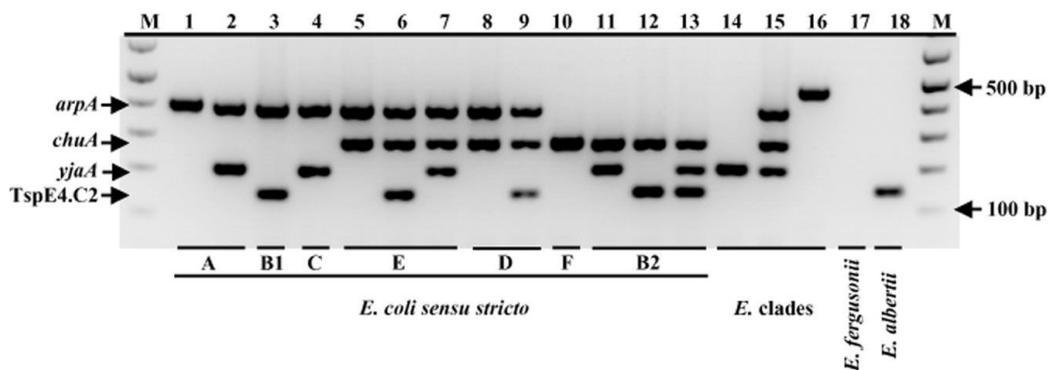


Figure 5 : Profils obtenus avec la PCR quadruple permettant d'assigner une souche à l'un des huit phylogroupes de *E. coli* à l'aide d'une PCR quadruple

arpA : gène codant une protéine non annotée ; *chuA* : gène codant le récepteur d'hème *ChuA*, *yjaA* : gène codant la protéine *YjaA* induite par le stress, *TspE4.C2* : Fragment d'ADN anonyme (Clermont et al., 2013).

Cependant, certains profils obtenus avec cette PCR quadruple peuvent correspondre à plusieurs phylogroupes. Ce problème a été résolu par l'ajout de PCRs complémentaires spécifiques de groupes (Lescat et al., 2013). Ces dernières sont basées sur la détection de variants spécifiques de certains marqueurs : *arpA* pour distinguer les souches du groupe E de celles du groupe D, et *trpA* pour distinguer les souches du groupe C de celles du groupe A.

Cette technique est toujours utilisée dans des études récentes, car elle permet d'investiguer rapidement l'appartenance aux différents groupes phylogénétiques, et donc la virulence potentielle des souches de *E. coli*. Par exemple, une étude de 2022 portant sur des souches de *E. coli* responsables de cystites chez des patientes en pédiatrie, a mis en évidence que la prévalence des souches appartenant aux groupes B2 et D est de 82% chez les patientes sans anomalie de l'arbre urinaire, et de seulement 23% chez celles présentant des anomalies de l'arbre urinaire (Domingos *et al.*, 2022).

1. 5. 3. Multi-Locus Sequence Typing (MLST)

La généralisation du séquençage d'ADN par la méthode Sanger a imposé le *Multi-Locus Sequence Typing* (MLST) comme la technique de référence pour le typage, et l'étude de la diversité des isolats de *E. coli*. Cette approche consiste à séquencer des fragments internes (500 paires de bases [pb] environ) de sept gènes domestiques. L'analyse des polymorphismes de chacun de ces *loci* génère un profil allélique à sept chiffres correspondant à un *Sequence type* (ST). Les STs proches peuvent être ensuite regroupés en complexes clonaux (CCs). Pour *E. coli*, trois schémas ont été décrits entre 2000 et 2006 (Escobar-P *et al.*, 2004 ; Reid *et al.*, 2000 ; Wirth *et al.*, 2006). Aujourd'hui, seuls deux schémas sont toujours utilisés, dont celui décrit par Wirth *et al.* (Figure 6) qui semble être le plus congruent avec les données phylogénétiques produites par le séquençage de génomes complets (Chaudhuri and Henderson, 2012).

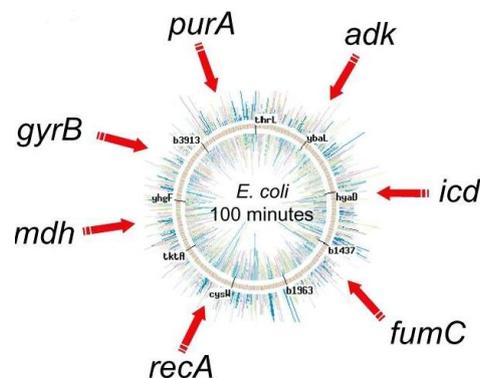


Figure 6 : Localisation génomique des sept gènes de ménage utilisés pour le schéma MLST de *E. coli*. Dans ce schéma sont inclus les gènes suivants : *adk* : gène codant l'adénylate kinase, *icd* : gène codant l'isocitrate déshydrogénase, *fumC* : gène codant la fumarate hydratase, *recA* : gène codant la recombinaise A, *mdh* : gène codant la malate déshydrogénase, *gyrB* : gène codant la sous-unité B de l'ADN gyrase, *purA* : gène codant l'adénylosuccinate synthetase (Wirth *et al.*, 2006).

Le typage par MLST présente l'avantage de fournir des résultats non ambigus et comparables d'un laboratoire à un autre, grâce à l'indexation de chaque allèle dans des banques de données internationales en ligne, comme PubMLST (<https://pubmlst.org/organisms/escherichia-spp>). Cependant, cette méthode reste chronophage, coûteuse, et nécessite des compétences en séquençage, ce qui limite son utilisation en routine.

Cette méthode a contribué à affiner la compréhension de la structure de l'espèce *E. coli*. En effet, elle a permis de démontrer l'importance des phénomènes de recombinaisons génétiques dans l'évolution du génome de *E. coli*, malgré son caractère clonal (Chaudhuri and Henderson, 2012), et a affiné la définition des phylogroupes (Tenailon et al., 2010).

Le typage par MLST s'est également avéré être un outil particulièrement utile pour suivre l'émergence et la diffusion de clones d'intérêt. Le ST131 est par exemple la lignée d'UPEC la plus étudiée, en raison de son émergence rapide et de sa propagation dans le monde au cours des 20 dernières années. Ce clone est également caractérisé par son urovirulence et sa résistance aux antibiotiques, tels que les fluoroquinolones et les β -lactamines (Nicolas-Chanoine et al., 2014). Une méta-analyse datant de 2019 portant sur 215 études utilisant la MLST, a mis en évidence la grande diversité génétique des isolats causant des infections extra-intestinales. Parmi eux, 20 STs majeurs sont le plus fréquemment identifiés à travers le monde, dont le ST131 (identifié dans 90% des études), suivi par les ST69 et ST10 (identifiés dans 50% des études) (Manges et al., 2019). D'autres génotypes, tels que les ST12, ST117, ST127, ST405, ST648 et ST1193 émergent dans plusieurs régions du monde (Riley, 2020).

1. 5. 4. CH typing

Une alternative plus rapide et moins onéreuse au typage par MLST a vu le jour en 2012, reposant sur le séquençage de seulement deux gènes : le *CH typing*. Cette méthode s'appuie sur un travail de 2010, qui démontrait l'intérêt pour le typage d'isolats cliniques, de séquencer un gène hautement polymorphe et très répandu chez les ExPEC, le gène *fimH*. Celui-ci code l'adhésine terminale du *pilus* de type I et présente un niveau de polymorphisme très élevé (Dias et al., 2010). En 2012, Weissman et al. associent au séquençage du gène *fimH* le séquençage du gène *fumC*, gène appartenant au schéma de MLST de Wirth et al. et présentant la plus grande variabilité allélique, pour développer le schéma à deux loci nommé *CH (fumC/fimH) typing*. Le typage de 1 044 souches de *E. coli* a démontré le pouvoir discriminant égal voire supérieur du *CH typing*, comparativement au typage par MLST, tout en permettant d'en prédire les résultats (prédiction des STs ou des CCs dans 95% des cas) (Weissman et al., 2012).

Cette méthode a été utilisée récemment dans une étude en 2022 portant sur 100 isolats d'UPEC. Dans cette étude, le *CH typing* a servi à typer les UPEC, en combinaison avec le phylogroupage, pour identifier rapidement 11 lignées de clones d'intérêt, dont la lignée CH 14-64 appartenant au ST1193, et présentant une résistance particulière vis-à-vis des fluoroquinolones. Cette lignée a pu par la suite être étudiée plus en détail par des méthodes plus résolutive (séquençage haut débit notamment)

difficiles à mettre en œuvre pour l'étude de larges populations. Cette étude souligne l'intérêt du *CH typing* pour l'investigation de larges populations (García-Meniño et al., 2022).

I. 5. 5. Méthodes basées sur le séquençage à haut débit

Principe

L'émergence des techniques de séquençage haut débit, ou *next generation sequencing* (NGS), a rendu possible l'accès à la séquence complète d'un génome bactérien. Les coûts décroissants des technologies, et l'amélioration des outils bioinformatiques, aident à une expansion rapide de cette méthode pour le typage d'isolats bactériens (Deurenberg et al., 2017).

Le séquençage d'un génome bactérien par NGS comprend 3 grandes étapes : (i) l'extraction de l'ADN génomique, (ii) le séquençage, dont le protocole varie selon la technologie utilisée, et (iii) l'exploitation des données (Figure 7).

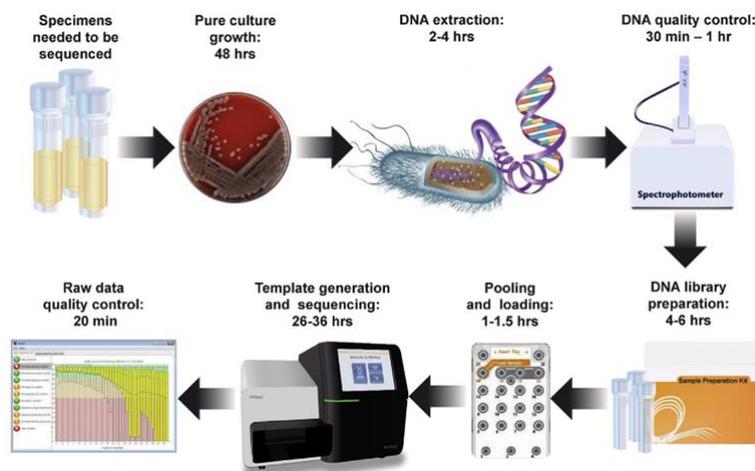


Figure 7 : Déroulement du séquençage d'un génome bactérien par NGS

La démarche comprend trois étapes : (i) l'extraction d'ADN à partir de la culture et le contrôle de sa qualité, (ii) le séquençage *sensu stricto* qui va de la préparation de la librairie jusqu'à l'obtention des séquences et (iii) l'analyse des données qui commence par le contrôle qualité des séquences obtenues. (Besser et al., 2018).

Actuellement, deux types de technologies de séquençage existent dont la distinction réside dans la taille des *reads* générés (fragments de séquence d'ADN générés par le séquenceur). Les technologies dites *short-reads* génèrent des *reads* courts (< 500pb), et les technologies dites *long-reads* génèrent des *reads* longs (> 1 kilobase [kb]). Les représentants actuels de ces deux types de séquençage sont respectivement Illumina pour les *short-reads* et PacBio ainsi que Oxford Nanopore Technologies (ONT) pour les *long-reads* (Besser et al., 2018).

Le séquençage par Illumina peut générer jusqu'à 2,5 milliards de *reads* par expérience, dont la taille est comprise entre 150 et 300 pb. Le coût actuel estimé est entre 20 et 200 € par Gigabase (Gb) séquencée. Cette méthode présente l'avantage de produire des séquences fiables (taux d'erreur

d'environ 0,1%) (Goodwin, 2016), mais les *short-reads* compliquent l'assemblage *de novo* notamment, au niveau des régions répétées. Ainsi, cette méthode ne génère pas réellement la séquence complète du génome, mais une séquence consensus dans laquelle il reste des *gaps* (régions non couvertes par le séquençage) (Besser et al., 2018 ; Maiden et al., 2013).

Les approches *short* et *long-reads* sont complémentaires. Les combiner permet en théorie de conserver la fiabilité de l'information générée par Illumina, tout en profitant des *long-reads* pour faciliter l'assemblage, et obtenir ainsi le génome le plus contigu possible. En 2017, un *pipeline* d'assemblage dit « hybride » nommé Unicycler a ainsi été développé (Wick et al., 2017), permettant de générer des assemblages à partir de *reads* provenant d'Illumina et de technologies *long-reads* (PacBio ou ONT).

Analyse des données

Les données de génome complet peuvent être utilisées de diverses façons selon le type d'analyse souhaitée et l'objectif visé. Elles permettent d'abord d'accéder à de nombreuses informations collectées via différentes méthodes de typage (sérotypage, pathovar, ST, phylogroupe, *CH type*...) en une seule expérience de séquençage, via l'utilisation d'outils adaptés (Robins-Browne et al., 2016). Ce faisant, de nouvelles méthodes de typage exploitant la richesse des données génomiques ont vu le jour. Deux approches majeures se distinguent : (i) les approches basées sur la phylogénie reposant sur l'analyse des variations ponctuelles de nucléotides ou *single nucleotide polymorphisms* (SNPs) et (ii) les approches dites « gène par gène » ou « *MLST-like* » (Maiden et al., 2013). Ces différentes méthodes n'ont pas le même degré d'exigence en matière de débit de séquençage, de longueur des *reads*, et de précision de la séquence établie. Dès lors, le choix de la technologie de séquençage ainsi que de divers paramètres de séquençage sont cruciaux et doivent être adaptés à l'application désirée (Vincent et al., 2017).

Typage basé sur les SNPs

Cette méthode est sans conteste la plus résolutive et permet de distinguer sans ambiguïté deux isolats. Elle consiste à aligner le génome de l'isolat d'intérêt sur un génome de référence, et à compter le nombre de positions variantes (Besser et al., 2018). Cette méthode est parfaitement adaptée au suivi longitudinal de souches, et a été appliquée dans le cadre du suivi de patientes atteintes d'infection urinaire récidivante à *E. coli* (Nielsen et al., 2021), ou encore pour déterminer *in vitro* la vitesse de divergence de souches de *E. coli* après 100 jours de *subculture* (Petronella et al., 2019). Cette méthode a également fait ses preuves lors d'une épidémie à *E. coli* O104:H4 ayant eu lieu en Allemagne et en France en 2011, en permettant de différencier des souches qui ne pouvaient être distinguées par les méthodes de typage conventionnelles (Grad et al., 2012). Elle reste cependant assez complexe à

utiliser en routine, essentiellement pour deux raisons. Tout d'abord, l'utilisation de cette méthode nécessite d'avoir une référence cohérente avec les isolats d'intérêt, de façon à identifier des SNPs réellement informatifs et à éliminer ceux dus au polymorphisme naturel de certains gènes. Deuxièmement, il n'existe actuellement pas de protocole standardisé pour la détermination des SNPs et le choix de la référence, ce qui rend les résultats difficilement comparables d'un laboratoire à l'autre (Besser et al., 2018).

Typage par approche dite « gène par gène »

Ce type d'approche reprend le principe de la MLST et l'étend à un plus grand nombre de gènes, permettant ainsi d'exploiter le potentiel des données génomiques (Maiden et al., 2013). Il existe plusieurs types d'approches « gène par gène » différant par le nombre de gènes étudiés. Parmi elles, les plus connues sont le *core genome* MLST (cgMLST) étudiant le polymorphisme des gènes présents dans la majorité des isolats d'une espèce donnée (> 97%), et le *whole genome* MLST qui s'intéresse cette fois à tous les gènes de la population étudiée (*core genome* et génome accessoire) (Besser et al., 2018). Le principal avantage de ce type d'approche, par rapport au typage par SNPs, est l'existence de schémas publiés et de banques de données en ligne qui permettent de standardiser les résultats et de les comparer d'un laboratoire à un autre (Besser et al., 2018). De plus, les outils développés pour le typage MLST sont compatibles avec ces nouvelles approches (Deurenberg et al., 2017). Il existe à ce jour un seul schéma de cgMLST pour *E. coli*, ciblant 2 513 gènes (Zhou et al., 2020). Cette approche a par exemple été utilisée dans une étude de 2021, afin d'identifier les sources de contamination à *E. coli* dans une usine de fabrication de produits alimentaires (Nakamura et al., 2021).

1.6. Structure de l'espèce

L'accumulation des données produites par les méthodes de typage décrites précédemment, ont considérablement amélioré la compréhension de la structure particulière de l'espèce *E. coli*. Cette espèce peut en effet être divisée en deux sous-espèces. La première sous-espèce contient les *E. coli sensu stricto* qui sont des souches appartenant à l'un des 9 phylogroupes décrits à ce jour (A, B1, B2, C, D, E, F, G, H) (Figure 8). Ces groupes s'organisent en deux clusters : l'un responsable de pathologies extra-intestinales (B2, D, F et son dérivé H, et G), et l'autre comportant plutôt des souches intestinales commensales et pathogènes (A, B1, C & E) (Denamur et al., 2021). La deuxième sous-espèce est composée des souches du clade I d'*Escherichia*. Ce sont des souches qui présentent la particularité d'échanger très fréquemment du matériel génétique avec les souches de *E. coli sensu stricto* (Luo et al., 2011).

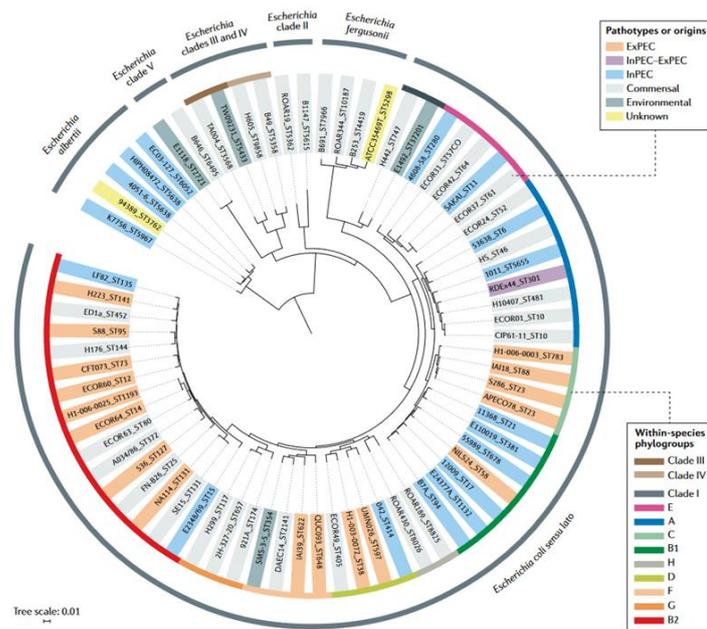


Figure 8 : Phylogénie du genre *Escherichia* basée sur les SNPs identifiés entre 72 souches appartenant à ce genre

Les isolats sont colorés en fonction de leur pathovar et/ou leur origine. L'anneau extérieur désigne les différentes espèces d'*Escherichia* sur la base de l'identité nucléotidique moyenne. *E. coli sensu stricto* désigne les souches classées dans l'un des phylogroupes connus de *E. coli* tandis que le terme de *E. coli sensu lato* désigne les *E. coli sensu stricto* et les souches appartenant au clade I du genre *Escherichia* (Denamur et al., 2021).

Malgré son caractère clonal, le taux de recombinaison dans le génome de *E. coli* peut être jusqu'à deux fois supérieur au taux de mutation. Cependant, ces recombinaisons ne concernent majoritairement que de très courts fragments du chromosome (environ 50 pb). Le taux de recombinaison est équivalent au taux de mutation pour des fragments dont la taille est comprise entre 2 et 4 kb, et est beaucoup plus faible pour des fragments de taille importante (autour de 100 kb) (Chaudhuri and Henderson, 2012 ; Denamur et al., 2021 ; Tenailon et al., 2010).

Les transferts horizontaux sont le moteur de l'émergence de lignées pathogènes chez *E. coli*. En effet, de nombreux facteurs de virulence décrits chez les InPEC sont portés par des éléments génétiques mobiles, comme des plasmides (entérotoxine thermolabile des ETEC), des transposons (entérotoxine thermostable des ETEC), des prophages (shiga-toxine des EHEC), et des îlots de pathogénicité (PAI 1 et 2 des UPEC ou le locus d'effacement des entérocytes des EPEC et EHEC) (Figure 9).

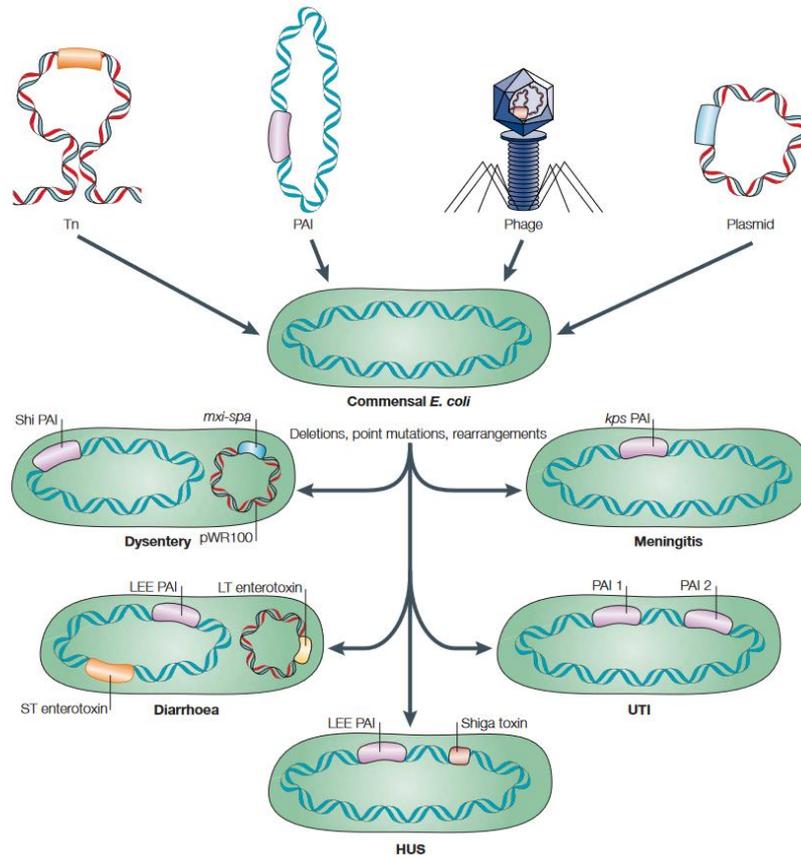


Figure 9 : Contribution des transferts horizontaux dans l'émergence des différents pathovars de *E. coli*

L'émergence des différents pathovars de *E. coli* est due à trois mécanismes : (i) l'acquisition de facteurs de virulence via des éléments génétiques mobiles (transposons, îlots de pathogénicité, phages, plasmides), (ii) l'inactivation de certains facteurs bridant la virulence et (iii) divers remaniements génétiques. Tn : transposon, PAI : îlot de pathogénicité, LEE : locus d'effacement des entérocytes, HUS : syndrome hémolytique et urémique, UTI : infection urinaire, ST : entérotoxine thermostable, LT : entérotoxine thermolabile (Kaper et al., 2004)

II. Infections urinaires

II.1. Appareil urinaire

II. 1. 1. Généralités

L'arbre urinaire est une suite d'organes creux dont la fonction est de collecter, transporter, stocker et évacuer l'urine. Ce système joue un rôle central dans le maintien de l'homéostasie hydrique et électrolytique du corps, ainsi que dans l'élimination des sous-produits du métabolisme de l'azote (*Hickling et al., 2015 ; Pellatt, 2007*).

L'arbre urinaire débute par les reins, comportant chacun environ un million de néphrons. Ces derniers sont le siège de la production de l'urine qui est le résultat de deux phénomènes. Le premier est la filtration du sang au niveau des glomérules de Malpighi, et le second est la réabsorption tubulaire se produisant au niveau des tubules rénaux. Les reins filtrent environ 150L de sang par jour et produisent en moyenne 1,5L d'urine définitive par jour (*Pellatt, 2007*). Cette dernière est acheminée via les uretères dans la vessie (**Figure 10**), qui peut contenir en moyenne au maximum 500mL d'urine. Grâce aux muscles lisses qui composent en partie sa paroi, la vessie peut se contracter pour provoquer l'évacuation de l'urine via l'urètre. Ce phénomène est appelé miction (*Hickling et al., 2015*).

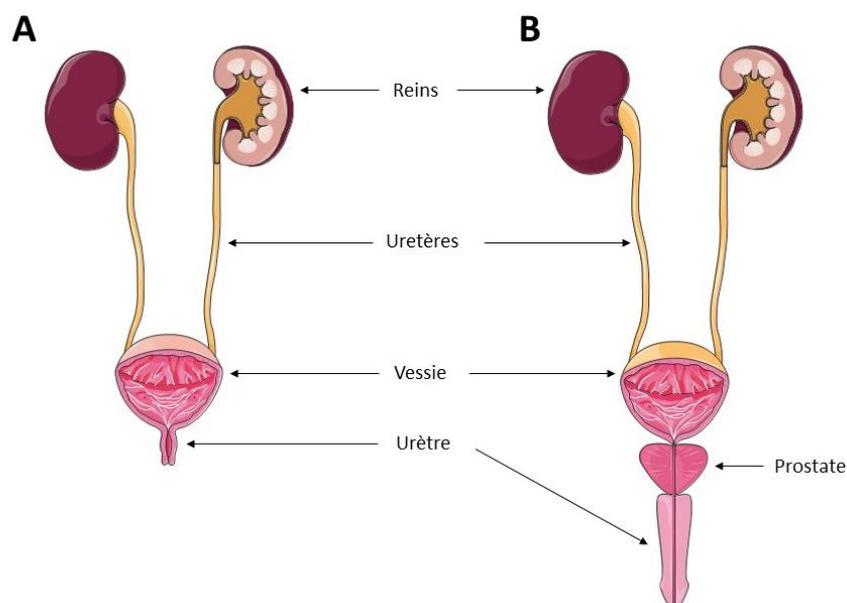


Figure 10 : Arbre urinaire féminin (A) et masculin (B)

L'arbre urinaire masculin (B) se distingue du féminin (A) par la longueur de l'urètre (18 à 20cm vs 3 à 4cm) et par la présence de la prostate qui engaine une partie de l'urètre. (www.smart.servier.com)

L'arbre urinaire masculin se distingue du féminin principalement par la longueur de l'urètre et l'existence de la prostate (**Figure 10**). En effet, l'urètre d'une femme adulte mesure en moyenne 3 à 4 cm de long, alors que l'urètre chez l'homme adulte mesure entre 18 et 20cm. De plus, la segmentation

de ce dernier est différente chez l'homme, en partie du fait de la présence de la prostate qui engaine une partie de l'urètre (Abelson et al., 2018).

II. 1. 2. Histologie de la vessie

Organisation générale

Sur le plan histologique, la vessie est composée de quatre tuniques de l'intérieur vers l'extérieur de la vessie : un épithélium, une *lamina propria* qui contient entre autres des fibres élastiques, une musculuse pour permettre la contraction de la vessie lors de la miction, et une adventice (**Figure 11**) (Bolla et al., 2022). Le tissu le plus remarquable de la paroi vésicale est son épithélium, nommé urothélium.

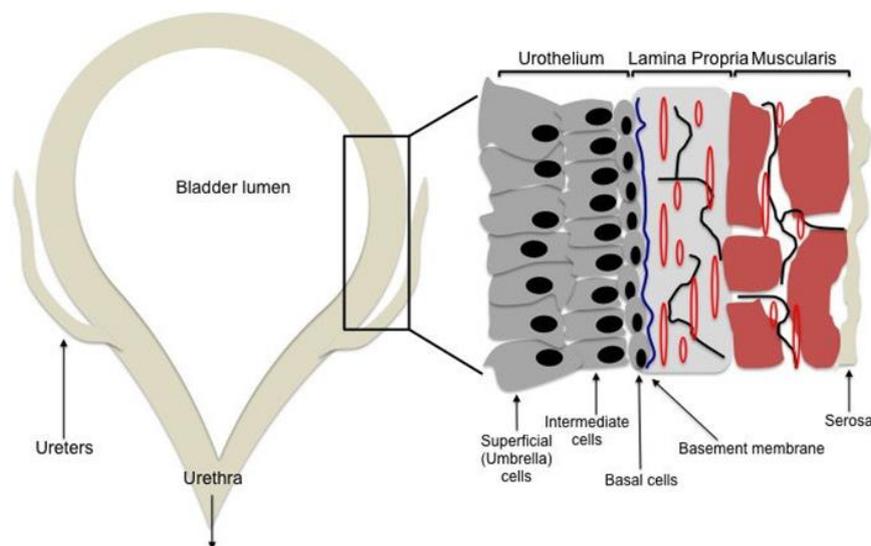


Figure 11 : Organisation histologique de la paroi de la vessie

La paroi de la vessie est composée de quatre tuniques : un urothélium constituant une barrière étanche, une lamina propria contenant de nombreuses fibres élastiques permettant à la paroi de la vessie de se déformer, une musculuse permettant la contraction de la vessie lors de la miction et une séreuse ou adventice. (Chan et al., 2017)

Structure de l'urothélium

Le terme d'urothélium peut désigner en réalité l'ensemble des épithéliums retrouvés le long de l'arbre urinaire, à l'exception de l'urètre antérieur (du fait de sa structure éloignée des autres épithéliums de l'arbre urinaire). Comme conseillé par certains auteurs (Hickling et al., 2015), le terme d'urothélium sera employé pour désigner uniquement l'épithélium vésical, lequel est de surcroit le mieux caractérisé.

L'urothélium est un épithélium pluristratifié à renouvellement très lent (200 jours en moyenne vs moins de 30 jours pour l'épiderme cutané). Il est composé d'une couche de cellules basales dites « cuboïdes », de quelques couches de cellules intermédiaires polygonales, et d'une unique couche de

cellules hautement différenciées souvent binucléées appelées « cellules en parapluie » (Bolla et al., 2022) (Figure 12).

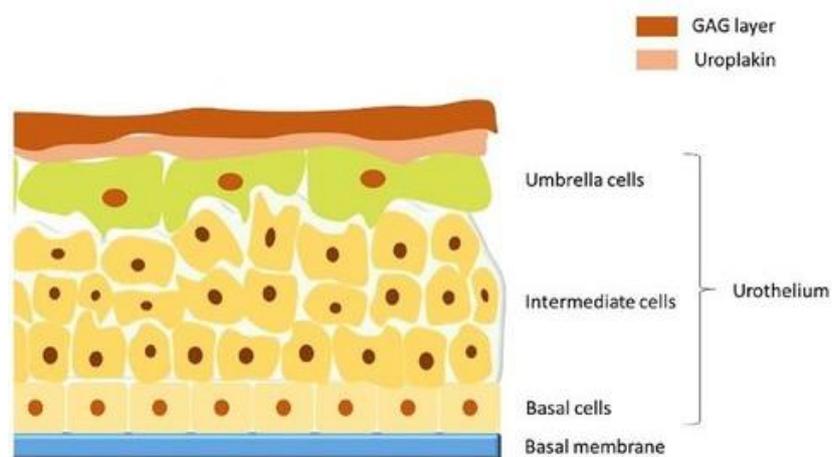


Figure 12 : Structure détaillée de l'urothélium vésical

L'urothélium est composé d'une couche de cellules basales assurant le renouvellement de ce tissu, de plusieurs couches de cellules intermédiaires et d'une couche de cellules différenciées : les cellules parapluies. GAG : Glycosaminoglycane (Lopez and Mangir, 2021).

Les cellules en parapluie sont liées entre elles par des jonctions serrées, afin de diminuer au maximum la perméabilité inter-cellulaire, et sont recouvertes de plaques cristallines hexagonales d'une glycoprotéine appelée uroplakine (UP). Il existe quatre UP majeures (UP Ia, Ib, II et IIIa) et une mineure (UP IIIb). Ces protéines s'assemblent en hétérotétramères à la membrane plasmique, pour former une couche cristalline d'uroplakine (Figure 13). Les UPs jouent par ailleurs un rôle important de récepteur cellulaire lors des infections urinaires (IU) à *E. coli* (cf. paragraphe III. 4. 1) (Hickling et al., 2015).

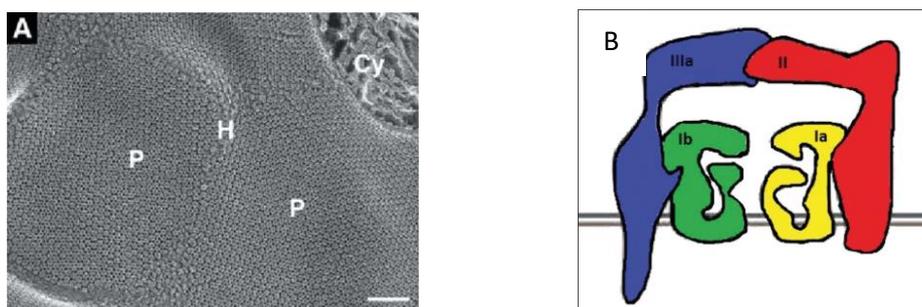


Figure 13 : (A) Observation des structures cristallines d'uroplakine (UP) à la surface de cellules urothéliales murine en microscopie électronique à balayage. (B) Schéma représentant l'hétérotétramère d'uroplakine assemblé à la membrane plasmique des cellules urothéliales.

(A) P : Plaque d'uroplakine, H : point de suture entre deux plaques adjacentes, Cy : cytosquelette apparent suite à la fracture de la membrane plasmique causée par la préparation de l'échantillon (Stoffler et al., 2005). (B) L'unité de base des plaques observées est l'hétérotétramère d'uroplakine. Ce dernier est composé de deux hétérodimères : UP IIIa & UP Ib associés à UP II & UP Ia (Matuszewski et al., 2016).

II. 1. 3. Composition de l'urine

La fonction première de l'urine est l'évacuation des déchets solubles comme l'urée (dégradation des acides aminés), la créatinine (dégradation de la créatine musculaire), l'acide urique (dégradation des acides nucléiques) et l'urobiline (dégradation de l'hémoglobine), cette dernière donnant sa couleur à l'urine. L'urine contient également des ions, des sucres et beaucoup d'autres composés, ce qui en fait un fluide biologique extrêmement complexe (Bouatra et al., 2013).

Une étude de 2013 a étudié la composition de l'urine du premier jet d'un corpus de 22 adultes volontaires (14 hommes et 8 femmes), âgés de 19 à 67 ans, et a identifié de façon certaine 2 651 métabolites, répertoriés dans une banque de données en ligne : *The Urine Metabolome Database* (Bouatra et al., 2013). Cette étude a également mis en évidence une variabilité extrêmement importante de la composition de l'urine d'un individu à l'autre, et au sein d'un même individu au cours du temps. Selon les composés, les variations observées vont de $\pm 50\%$ à $\pm 350\%$. Quatorze composés dits majeurs (concentration > 1 mM) ont été identifiés (Figure 14). Parmi eux, l'urée et les ions sodium, chlorure et potassium représentent à eux seuls environ 85% des solutés totaux (Reitzer and Zimmern, 2019).

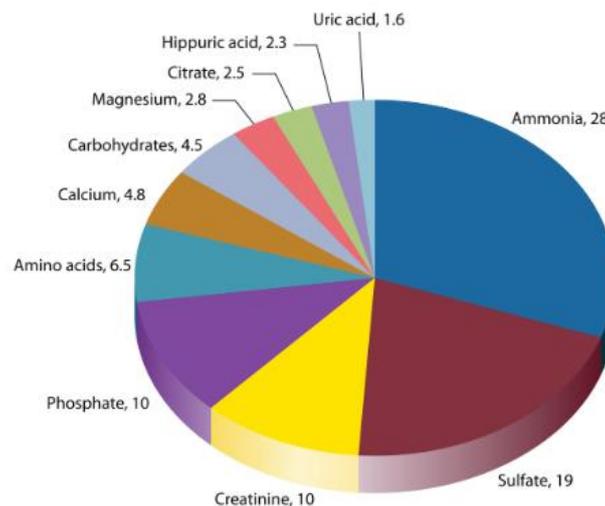


Figure 14 : Composés majoritaires (> 1 mM) présents dans l'urine

Le nombre à côté de chaque composé correspond à sa concentration en millimolaire (mM) dans l'urine. Pour des raisons de lisibilité, l'urée (290 mM), le sodium (123 mM), le chlorure (115 mM) et le potassium (46 mM) ne sont pas représentés sur le graphique (Reitzer and Zimmern, 2019).

Le pH de l'urine varie entre 5,0 et 8,0 et est en moyenne de 6,0. Du fait de sa faible teneur en espèces tampon, le pH de l'urine peut être impacté par divers facteurs, dont le métabolisme bactérien.

L'urine, du fait de sa composition, présente un certain nombre de caractéristiques (milieu oligotrophique, pauvre en certains oligoéléments) qui rendent la croissance bactérienne difficile.

Pour étudier le comportement des bactéries dans l'urine, deux approches peuvent être mises en place. La première consiste à organiser un recueil d'urine auprès de personnes saines afin d'utiliser cette dernière comme milieu. Cependant du fait de la grande variabilité inter-individuelle existant en termes de composition, utiliser de l'urine implique nécessairement de récolter de grandes quantités d'urine provenant de nombreuses personnes, afin de constituer un lot avec lequel toutes les expérimentations pourront être menées. En plus de la variabilité dans la composition se pose le problème de la conservation de cette urine dans le temps. Ainsi, des milieux tentant de reproduire *in vitro* les conditions physico-chimiques de l'urine ont été mis au point (**Tableau 2**). Parmi eux, certains ont été développés afin d'étudier la croissance des bactéries uropathogènes, comme celui publié par Brooks et Keevil en 1997 (*Brooks and Keevil, 1997*). Ces milieux présentent l'avantage non négligeable d'être de composition connue et rendent ainsi la comparaison des résultats entre laboratoires possible (*Hogins et al. 2022*).

Tableau 2 : Liste des milieux urinaires artificiels publiés

SHU = Synthetic Human Urine. Adapté de (*Ipe et al., 2016*)

A. Original study ¹	Focus of the research application	Microbes or cell types	Relevance
1. Physiology (Aurora et al., 1980)	Understanding the development of calcium oxalate monohydrate precipitation	N/A	Urolithiasis
2. Infection (Brooks and Keevil, 1997)	Development of SHU to investigate the growth of urinary pathogens	<i>E. coli</i> , <i>P. aeruginosa</i> , <i>P. mirabilis</i> , <i>S. epidermidis</i> , <i>C. albicans</i>	UTI
3. Physiology (Burns and Finlayson, 1980)	Description of standard SHU for <i>in vitro</i> urolithiasis assays	N/A	Urolithiasis
4. Cell Biology (Chutipongtante and Thongboonkerd, 2010)	Comparison of multiple SHU media, and study in epithelial cell culture assays	Kidney epithelial cells	Cell Biology
5. Infection and Physiology (Griffith et al., 1976)	Investigation of infection-induced urinary stones	<i>E. coli</i> , <i>P. aeruginosa</i> , <i>P. mirabilis</i> , <i>Citrobacter koseri</i> , <i>Proteus rettgeri</i> , <i>Providencia stuartii</i> , <i>Morganella morganii</i> , <i>Klebsiella oxytoca</i> ,	UTI, and Urolithiasis
6. Infection (Minuth et al., 1976)	Measurement of the antimicrobial efficacy of gentamicin in SHU	<i>E. coli</i> , <i>P. aeruginosa</i>	UTI
7. Physiology (Putnam et al., 1971)	Characterization of urinary constituents	N/A	General
8. Physiology (Gardner and Doremus, 1978)	Study of calcium oxalate crystallization	N/A	Urolithiasis

II.2. Microbiote urinaire

L'urine a longtemps été considérée comme un fluide biologique stérile (*Bouatra et al., 2013; Hickling et al., 2015*). Cela est en partie dû à la faible abondance des bactéries dans l'urine comparativement aux autres microbiotes (10^4 à 10^5 UFC/mL d'urine) (*Pearce et al., 2014*), et au fait que les méthodes classiques de culture de l'urine favorisent principalement la croissance des micro-organismes aérobies et non exigeants comme les entérobactéries (*Whiteside et al., 2015*). Cependant, l'avènement du séquençage haut-débit (*Wolfe et al., 2012*) ainsi que la mise au point de méthodes de culture étendue quantitative d'urine (*Hilt et al., 2014*), ont permis de mettre en évidence l'existence d'un microbiote urinaire.

Le microbiote urinaire est très variable d'un individu à un autre, et varie selon le sexe et l'âge. Le microbiote urinaire « sain » chez l'homme et la femme, présente de nombreux genres bactériens communs, comme *Prevotella*, *Escherichia*, *Enterococcus*, *Streptococcus* ou encore *Citrobacter*. La principale différence de composition du microbiote entre les sexes réside dans l'abondance de certains genres : les genres *Streptococcus* et *Corynebacterium* sont plus abondants chez l'homme, alors que le genre *Lactobacillus* domine chez la femme (Perez-Carrasco et al., 2021).

L'origine de ce microbiote n'est pas encore clairement élucidée. Chez la femme, le microbiote vaginal pourrait en être la principale source de par sa proximité anatomique. Deux études récentes ont en effet émis l'hypothèse d'une interconnexion entre les microbiotes urinaire et vaginal chez la femme (Komesu et al., 2020 ; Thomas-White et al., 2018). L'analyse phylogénétique de génomes entiers d'isolats identifiés chez des mêmes femmes dans les prélèvements vaginaux et urinaires, a en effet mis en évidence des souches de *Lactobacillus iners*, *L. crispatus*, *Streptococcus anginosus* et *E. coli* très similaires (Thomas-White et al., 2018). Cependant, une étude de 2020 a également montré chez la femme, l'existence d'une forte ressemblance entre le microbiote intestinal et le microbiote urinaire, avec 62,5 % d'espèces communes entre ces deux microbiotes, contre 32 % en commun avec le microbiote vaginal (Perez-Carrasco et al., 2021) (Figure 15).

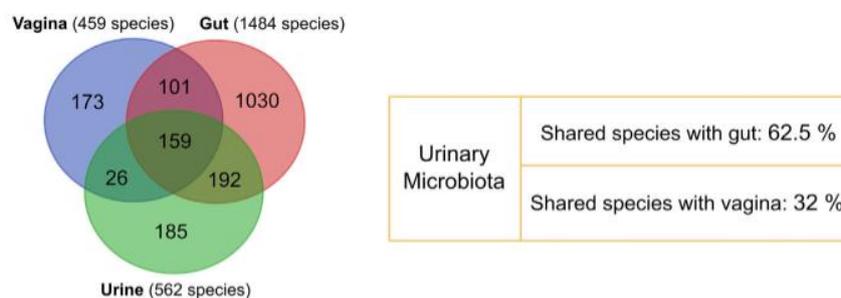


Figure 15 : Diagramme de Venn comparant les espèces composant le microbiote urinaire, intestinal et vaginal

Le microbiote urinaire (en vert) partage 351 espèces (62,5%) avec le microbiote intestinal (en rouge) vs 185 espèces (32%) avec le microbiote vaginal (en bleu) (Perez-Carrasco et al., 2021).

La place du microbiote urinaire dans la physiopathologie des IU est encore assez mal connue. Certaines espèces de *Lactobacillus*, comme *L. crispatus*, joueraient un rôle protecteur vis-à-vis des IU grâce à leur capacité à inhiber la croissance de *E. coli* (Hudson et al., 2020). Ainsi, une altération du microbiote urinaire réduisant l'abondance des *Lactobacillus* protecteurs, pourrait rendre le tractus urinaire vulnérable à la colonisation par des uropathogènes (Garretto et al., 2020).

II.3. Généralités sur les infections urinaires

II. 3. 1. Définitions et diagnostic

Une IU est définie par la présence de micro-organismes (le plus souvent des bactéries) dans l'urine, associée à des signes cliniques et biologiques (Caron et al., 2018). Ainsi, le diagnostic d'une IU repose sur la présence de signes cliniques très variables dans leur gravité, selon la localisation anatomique de l'infection associée à la mise en évidence d'une bactériurie significative, accompagnée ou non d'une leucocyturie ($\geq 10^4$ leucocytes/mL d'urine) (Caron et al., 2018 ; Geerlings, 2016). Des seuils de significativité de bactériurie ont été définis afin d'aider au diagnostic biologique d'une IU, en fonction de l'espèce bactérienne (caractère uro-pathogène plus ou moins marqué), ainsi que du sexe du patient (Tableau 3) (Caron et al., 2018).

Tableau 3 : Seuils de significativité de bactériurie selon l'espèce bactérienne et le sexe du patient (Caron et al., 2018)

Espèces bactériennes	Seuil de significativité (UFC/mL)	
	Homme	Femme
<i>E. coli</i> , <i>S. saprophyticus</i>	$\geq 10^3$	$\geq 10^3$
Entérobactéries autres que <i>E. coli</i> , entérocoque, <i>C. urealyticum</i> , <i>P. aeruginosa</i> , <i>S. aureus</i>	$\geq 10^3$	$\geq 10^4$

Le terme d'IU regroupe les cystites (infections de la vessie) et les pyélonéphrites (infections du rein). Les IU masculines représentent une entité à part du fait d'une hétérogénéité clinique. Certaines patientes présentent des facteurs de risques pouvant rendre les IU plus sévères ou plus difficiles à traiter. Ces facteurs de risques incluent les anomalies de l'arbre urinaire, la grossesse, l'âge (> 65 ans) ou encore l'immunodépression grave. Dans le cas où une patiente présentant au moins l'un de ces facteurs de risque de complication contracte une IU, cette dernière sera alors qualifiée d'IU à risque de complication (ou compliquée dans les pays anglosaxons) (Caron et al., 2018 ; Geerlings, 2016).

Pour la stratégie diagnostique et thérapeutique, les IU sont qualifiées de récidivantes en France par la survenue d'au moins quatre épisodes sur une période de douze mois consécutifs (Caron et al., 2018). Cependant, il n'existe pas de consensus international autour de cette définition. Dans la littérature scientifique, notamment portant sur la physiopathologie, les IU récidivantes sont souvent définies par la survenue d'au moins trois épisodes sur une période de douze mois, ou au moins deux épisodes sur une période de six mois (Glover et al., 2014).

II. 3. 2. Epidémiologie et étiologie des infections urinaires

Les IU font partie des infections bactériennes les plus fréquentes chez l'Homme (Flores-Mireles et al., 2015). Les IU sont plus fréquentes chez la femme (12,6% d'incidence annuelle vs 3% chez les hommes) du fait de facteurs anatomiques : un urètre plus court, la proximité entre le méat urinaire et l'anus, favorisant la contamination de l'arbre urinaire par des bactéries intestinales, et la contiguïté avec la filière génitale, ces facteurs étant résumés par certains auteurs en « réservoir recto-vaginal des cystites » (Foxman and Brown, 2003). Le nombre d'IU communautaires par an est estimé à 150 millions à travers le monde. Les IU associées aux soins sont la première cause d'infections nosocomiales en France (Santé Publique France, 2017 ; 25% des cas) et la 3^{ème} cause des infections associées aux soins en Europe (19.6% des cas) (European center for diseases control, 2013).

L'étiologie des IU varie en fonction du type d'infection (communautaire ou associée aux soins). Dans le cas des IU communautaires, une écrasante majorité (entre 74 et 83% selon la population étudiée) sont causées par des entérobactéries, et plus spécifiquement par *E. coli* (65 à 75%) (Figure 16) (Flores-Mireles et al., 2015).

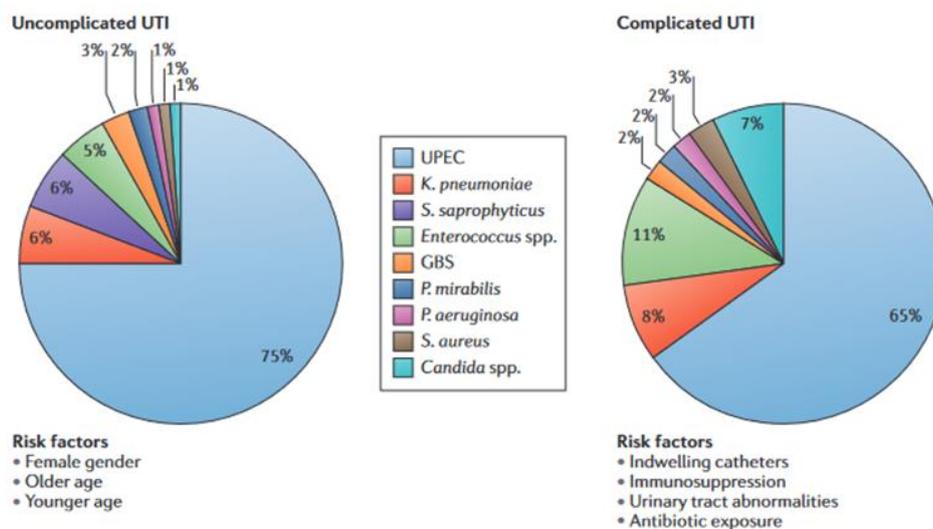


Figure 16 : Épidémiologie des infections urinaires communautaires

Diagramme de gauche : IU simples, diagramme de droite : IU à risque de complication. UTI : Urinary tract infection, GBS : Group B Streptococcus, UPEC : Uropathogenic Escherichia coli (Flores-Mireles et al., 2015).

Dans les IU associées aux soins, si *E. coli* reste l'agent étiologique principal, sa proportion se retrouve considérablement réduite au profit d'autres espèces (Figure 17) (Tandogdu and Wagenlehner, 2016).

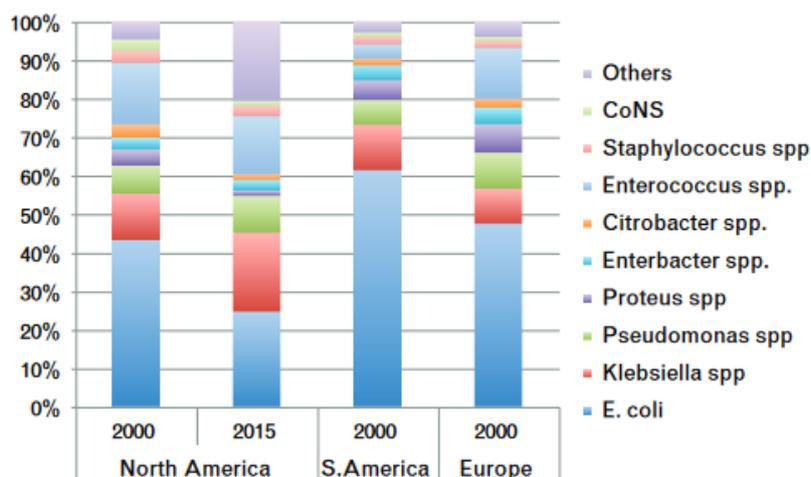


Figure 17 : Épidémiologie des IU associées aux soins en Amérique du Nord et en Europe
CoNS : Coagulase Negative Staphylococcus (Tandogdu and Wagenlehner, 2016).

II. 3. 3. Physiopathologie des IU

Les IU se développent par voie ascendante et sont principalement causées par des souches UPEC. Ces dernières proviennent généralement de l'intestin et ont commencé par coloniser les zones vaginale et péri-urétrale. Les bactéries vont ensuite remonter le long de l'urètre pour atteindre la vessie qu'elles vont coloniser. Dans cette dernière, *E. coli* va causer des dommages tissulaires et engendrer une réponse inflammatoire, causant donc une cystite. Par la suite, la souche d'UPEC peut atteindre les reins via les uretères et causer une pyélonéphrite. Enfin, les bactéries peuvent rejoindre la circulation sanguine et causer une bactériémie en traversant la paroi de l'un des nombreux vaisseaux sanguins irriguant le rein (**Figure 18**).

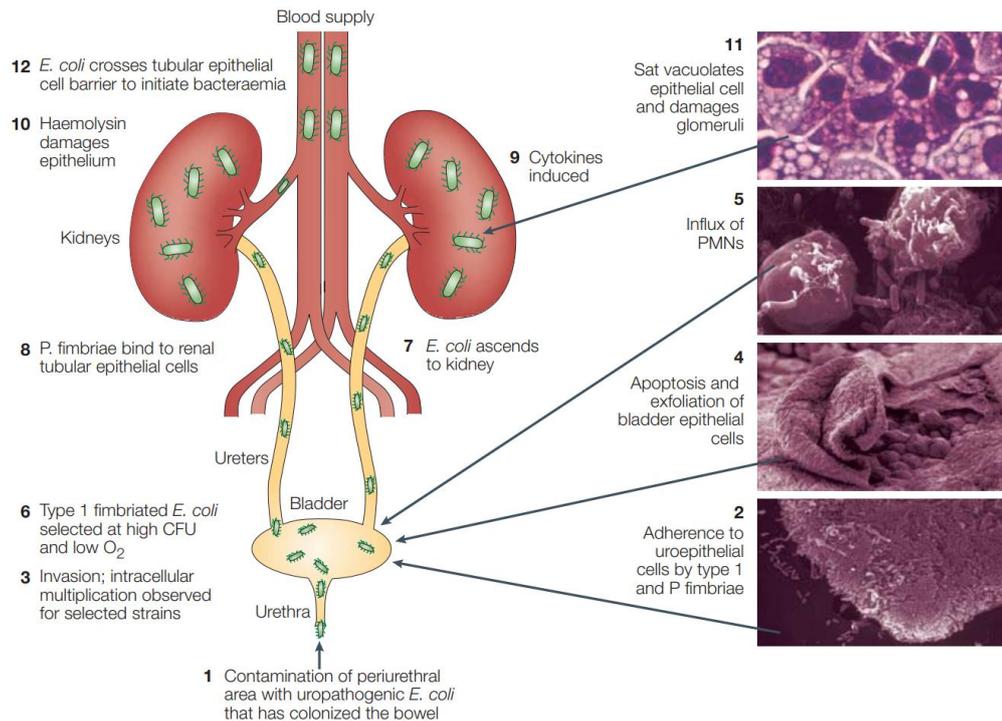


Figure 18 : Physiopathologie des infections urinaires causées par un UPEC

Le point de départ d'une IU est la colonisation de la région péri-urétrale par une souche d'UPEC qui va ensuite remonter le long de l'urètre, atteindre la vessie et coloniser cette dernière. Dans la vessie, les UPEC vont causer une cystite en provoquant des dommages tissulaires (4) ainsi qu'en déclenchant la mise en place d'une réponse inflammatoire (5). Par la suite, les bactéries peuvent atteindre les via les uretères et engendrer une pyélonéphrite, puis traverser la paroi de l'un des vaisseaux sanguins irriguant les reins pour rejoindre la circulation sanguine et causer une bactériémie. PMN : Polynucléaire neutrophile, CFU : Unité formant colonie (Kaper et al., 2004).

II.4. *E. coli* uropathogène (UPEC)

II. 4. 1. Définition du pathovar

S'il est possible d'identifier les pathovars appartenant aux InPEC via des PCRs ciblant les gènes codant leurs principaux facteurs de virulence (**Tableau 4**) (Riley, 2020), il n'existe pas de telles méthodes pour identifier les pathovars d'ExPEC. Ces derniers se définissent donc en fonction du site d'isolement de la souche (Denamur et al., 2021 ; Foxman, 2010). Ainsi, une souche d'UPEC se définit comme étant une souche de *E. coli* isolée des urines d'un patient souffrant d'une IU (Denamur et al., 2021).

Tableau 4 : Gènes de virulence ciblés par des PCRs permettant d'identifier les différents pathovars d'InPEC

(Riley, 2020)

<i>Escherichia coli</i> group	Gene target
Enteropathogenic <i>E. coli</i> (EPEC)	<i>bfpA, eae</i>
atypical EPEC (aEPEC)	<i>eae</i>
Enterotoxigenic <i>E. coli</i> (ETEC), heat stable (ST) producer	ST gene
ETEC, heat labile (LT) producer	LT gene
Enteroinvasive <i>E. coli</i> (EIEC)	<i>ipaC, ipaH</i>
Shigatoxin-producing <i>E. coli</i> (STEC)	<i>stx1, stx2</i>
Enterohemorrhagic <i>E. coli</i> (EHEC)	<i>stx1, stx2, eae,</i>
Enteraggregative <i>E. coli</i> (EAEC)	<i>aatA</i>
Enteraggregative-Enterohemorrhagic <i>E. coli</i> (EAEC-EHEC)	<i>aatA, eae, stx2</i>

II. 4. 2. Épidémiologie de la résistance aux antibiotiques des UPEC

E. coli est une espèce naturellement sensible à de nombreux antibiotiques utilisés en clinique comme la plupart des β -lactamines ou les fluoroquinolones. Cependant, la plasticité de son génome lui permet d'acquérir de nombreux mécanismes de résistance, médiés par l'apparition de mutations de gènes chromosomiques ou par des transferts horizontaux (Poirel et al., 2018). C'est pourquoi *E. coli* figure, au même titre que d'autres entérobactéries, sur la liste des agents pathogènes prioritaires établie par l'organisation mondiale de la santé, c'est-à-dire des bactéries pour lesquelles il est urgent de développer de nouveaux antibiotiques (OMS, communiqué de presse du 27/02/2017).

Les IU font partie des principales causes de prescription d'antibiotiques en France et dans le monde. Elles représentaient 15% des prescriptions d'antibiotiques en 2015 en ville en France selon le rapport de l'Agence Nationale de Sécurité du Médicament et des produits de santé (ANSM, Rapport 2016). Aux États-Unis, les IU constituaient le 2nd pôle de prescription d'antibiotique en 2010 (Foxman, 2010).

Les antibiotiques recommandés pour le traitement des cystites varient selon les pays en fonction des molécules disponibles, de l'analyse de leur balance bénéfice-risque, et l'épidémiologie locale. Quasiment partout le triméthoprime et l'association triméthoprime-sulfaméthoxazole (abrégé en cotrimoxazole dans la suite du manuscrit) ont été abandonné en première intention du fait d'une prévalence de la résistance supérieure à 20%, d'où des taux d'éradication clinique insuffisants (Caron et al. 2018) En France, la fosfomycine est positionnée en première intention et le pivmécillinam en second dans la cystite simple. La toxicité rarissime mais gravissime de la nitrofurantoïne (fibrose pulmonaire laissant des insuffisances respiratoires à vie, ou hépatotoxicité pouvant conduire à la greffe hépatique) a condamné la molécule malgré son excellente efficacité *in vitro* comme *in vivo* (Caron et

al., 2018). Bien d'autres pays dont les États-Unis d'Amérique (*Gupta et al., 2011*) ont conservé la nitrofurantoïne en première intention dans le traitement de toutes les formes de cystites (*Gupta et al., 2011*).

Les taux de résistance des souches de *E. coli* responsables d'IU sont donc faibles (<10%) pour les molécules de premières intention réservées aux cystites que sont la fosfomycine, le pivmécillinam et la nitrofurantoïne, contrairement à d'autres molécules comme le cotrimoxazole et les fluoroquinolones qui connaissent des indications thérapeutiques nombreuses au-delà des IU. La résistance aux fluoroquinolones, principalement médiée par des mutations dans les gènes *gyrA* ou *parC*, est devenue préoccupante en France, avec un surrisque documenté pour les patients ayant été exposés à cette classe dans les six derniers mois (*Caron et al., 2018*). D'autre part, la résistance aux céphalosporines de 3^{ème} génération (prescrites en probabiliste dans le traitement des pyélonéphrites) a largement progressé, concernant environ 5% des souches (*Onerba, 2018*). Cette résistance est principalement liée à la production de bêta-lactamases à spectre élargi (BLSE), enzymes codées par des gènes plasmidiques. Les données d'antibiorésistance des UPEC aux principales molécules antibiotiques, en France et aux États-Unis sont présentées dans le **Tableau 5**.

Tableau 5 : Données de résistance des UPEC aux principaux antibiotiques constituant l'arsenal thérapeutique contre les IU

ND : Non disponible. Les chiffres pour la France proviennent de l'Observatoire National de l'Épidémiologie de la Résistance Bactérienne aux Antibiotiques (*ONERBA, 2018*) et ceux pour les États-Unis proviennent de l'étude de Sanchez et al. (*Sanchez et al., 2016*)

Antibiotique	% souches résistantes	
	France	États-Unis
Amoxicilline	41,8%	41,3%
Amoxicilline – Acide Clavulanique	16,9%	3,9%
Céphalosporines de 3 ^e génération	3,2%	4,5%
Cotrimoxazole	19,8%	22,3%
Nitrofurantoïne	0,6%	0,9%
Fluoroquinolones	11,0%	11,8%
Fosfomycine	0,9%	4,2%
Pivmécillinam	6,7%	ND

Parmi les UPEC, les souches du ST131 sont considérées comme étant en grande partie responsables de la dissémination de l'antibiorésistance chez *E. coli*. Ce clone constitue en effet la principale lignée d'UPEC productrice de BLSE et résistante aux fluoroquinolones dans le monde (*Lindblom et al., 2022 ; Nicolas-Chanoine et al., 2014*), suivie par le ST1193 qui commence à prendre le pas sur le ST131 dans certaines régions du monde (*Pitout et al., 2022*). Le ST69 a également été identifié comme étant, avec

le ST131, fortement pourvoyeur de souches résistantes au cotrimoxazole (*Tchesnokova et al., 2015 ; Yamaji et al., 2018*).

Le caractère récidivant des IU a été identifié comme facteur de risque de l'antibiorésistance chez les UPEC du fait des expositions thérapeutiques individuelles itératives (*Ikram et al., 2015 ; G. Lee et al., 2010*). Par exemple, dans une étude de 2015 menée au Brésil, l'antibiorésistance de souches de *E. coli* responsables d'IU sporadiques ($n = 233$) a été comparée à celle de souches responsables d'IU récidivantes ($n = 187$). Les isolats responsables d'IU récidivantes étaient significativement plus résistants à la quasi-totalité des molécules testées (*Hisano et al., 2015*).

La prescription antibiotique a également un effet collectif sur l'antibiorésistance des UPEC comme cela est démontré par Kahlmeter à l'échelon européen avec une parfaite corrélation entre les consommations nationales et le risque local de multirésistance (*Kahlmeter et al. 2003*).

II. 4. 3. Facteurs de virulence

Un grand nombre de facteurs de virulence ont été identifiés chez les UPEC. Ils peuvent être classés selon leur fonction (*Johnson and Russo, 2018*). Quatre grands groupes ont ainsi été définis : les adhésines, les toxines, les systèmes de captation du fer, et les facteurs permettant d'échapper au système immunitaire (SI) (**Figure 19**) (*Lüthje and Brauner, 2014*).

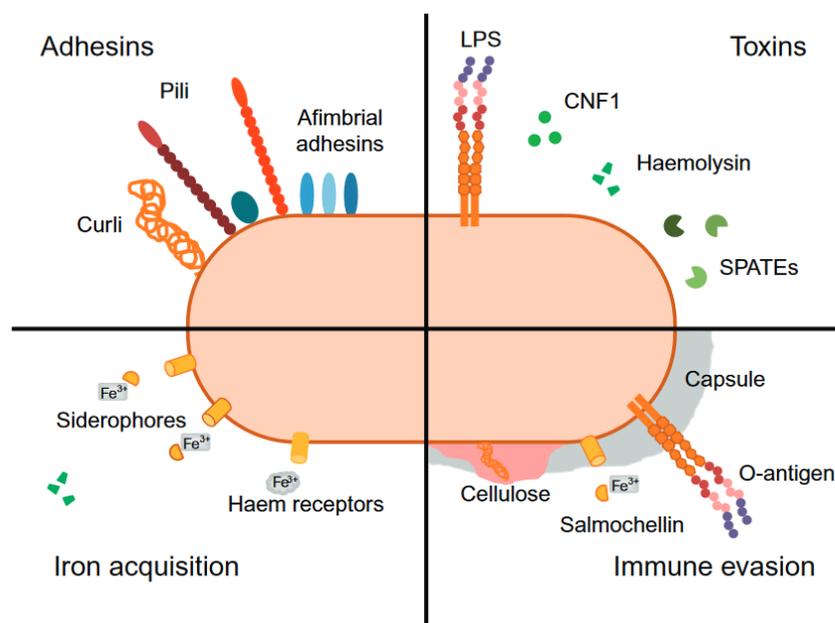


Figure 19 : Principaux facteurs de virulence des UPEC classés selon leur fonction (*Lüthje and Brauner, 2014*)

Adhésines

Les adhésines sont des protéines membranaires permettant aux bactéries d'adhérer aux cellules hôtes ou à diverses surfaces. Il en existe deux sortes : les adhésines fimbriales, composées de plusieurs protéines qui s'organisent sous forme d'appendices filamenteux, et les adhésines afimbriales, qui sont directement ancrées à la paroi bactérienne (**Figure 19**).

À ce jour, 19 adhésines ont été identifiées chez les UPEC (**Tableau 6**), leur permettant d'adhérer à la plupart des tissus constituant l'arbre urinaire. Si le rôle de chacune de ces adhésines n'est pas encore élucidé, leur multiplicité chez les UPEC semble être un avantage sélectif (*Subashchandrabose and Mobley, 2015*). Il a été montré en 2011 qu'en moyenne les souches d'UPEC possèdent plus d'opérons codant des adhésines que les souches fécales de *E. coli*. C'est par exemple le cas de la souche d'UPEC *E. coli str. CFT073* (isolat de pyélonéphrite) qui possède 12 opérons codant des *fimbriae* (*Spurbeck et al., 2011*).

Tableau 6 : Liste des gènes ou opérons codant des adhésines identifiés chez les UPEC
(*Johnson and Russo, 2018*)

Category	Gene(s) or operon	Comment	Evidence for trait as a virulence factor (reference[s])	
			Epidemiological ^b	Experimental (<i>in vivo</i>) ^c
Adhesins ^d	<i>afa/dra</i>	Dr antigen-binding adhesins (AFA I-III, Dr, F1845)	Yes (176 , 266 , 267)	Yes (268)
	<i>afaE-8</i>	Afimbrial adhesin VIII	Yes (175)	No
	<i>auf</i>	Surface structure of unclear function	Yes (269)	No (269)
	<i>bmaE</i>	Blood group M-specific adhesin	Yes (270)	No
	<i>clpG</i>	CS31A adhesin (K88-related)	Yes (271)	No
	<i>csgA</i>	Curli	Yes (272)	Yes (272)
	<i>ecp</i>	Pilus	Yes (273)	No
	<i>fim</i>	D-mannose-specific adhesin, type-1 fimbriae	Yes (267 , 274 , 275)	Yes (89 , 276)
	<i>foc</i>	F1C fimbriae	Yes (277)	No
	<i>gafD</i>	GlcNac-specific (G, F17c) fimbriae adhesin	Yes (270 , 271)	No
	<i>iha</i>	Iron-regulated-gene-homologue adhesin	Yes (21 , 111 , 267)	Yes (96)
	<i>nfa</i>	NFA-1, -2, -3, -4 (nonfimbrial adhesins)	No (278-280)	No
	<i>pap</i>	Pilus associated with pyelonephritis (P fimbriae)	Yes (9 , 151 , 267)	Yes (281); no (282)
	<i>pil</i>	Type IV pilus	Yes (283)	Yes (284)
	<i>sfa</i>	S fimbriae (sialic acid-specific)	Yes (111)	Yes (285)
	<i>sfa/foc</i>	S and F1C fimbriae	Yes (286)	Yes (285)
	<i>yad</i>	Fimbria	Yes (283)	Yes (283)
	<i>ygi</i>	Fimbria	Yes (283)	Yes (283)
	<i>yvc</i>	Fimbria	Yes (283)	No

Les adhésines les plus étudiées chez les UPEC sont les *pili* de type I codés par l'opéron *fim* et les *pyelonephritis associated pili* codés par l'opéron *pap* (*Subashchandrabose and Mobley, 2015*). Ces deux *pili* s'assemblent via un mécanisme dit *chaperone-usher* (*Geibel and Waksman, 2014*).

Pilus de type I

Le pilus de type I est assemblé à la membrane externe par une protéine transmembranaire qualifiée de placier ou *usher* (FimD). Les sous-unités de piline (FimA) sont repliées dans le périplasme et sont

acheminées vers FimD par la protéine chaperonne FimC. À l'extrémité du pilus, FimH joue le rôle d'adhésine terminale. Cette protéine est constituée de deux domaines : le domaine FimH_L (lectine) pouvant interagir de façon spécifique avec les résidus D-mannose présents de façon abondante à la surface de l'urothélium (Geibel and Waksman, 2014), et le domaine FimH_P permettant d'ancrer l'adhésine terminale au reste du pilus via les protéines FimG et FimF (Figure 20). Il est codé par l'opéron *fim* appartenant au *core genome* de *E. coli* (Oelschlaeger et al., 2002 ; Subashchandrabose and Mobley, 2015).

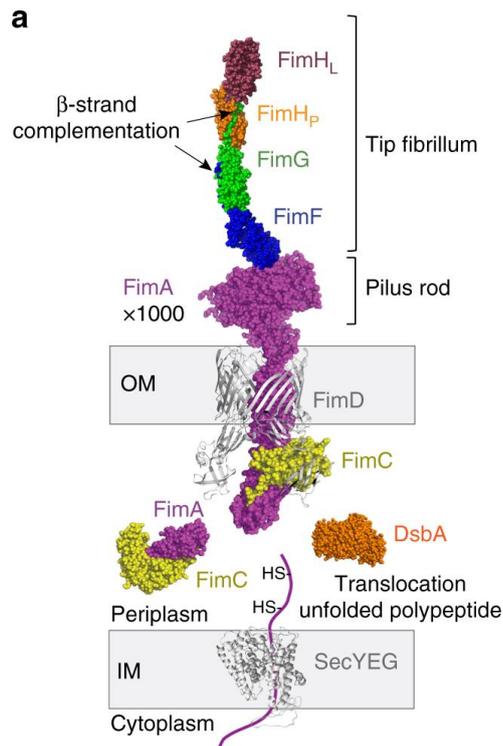


Figure 20 : Structure du pilus de type I

Le système SecYEG (*Sec*) permet l'export des sous-unités de piline (*FimA*) vers le périplasme où elles seront repliées et prises en charge par la protéine chaperonne *FimC*. Cette dernière va ensuite acheminer les sous-unités *FimA* vers la protéine placier (*FimD*) permettant d'assembler le pilus à la membrane externe. À l'extrémité du pilus, la protéine *FimH* joue le rôle d'adhésine terminale permettant aux UPEC de se lier aux résidus mannose à la surface de l'urothélium. *FimH* est reliée au reste du pilus via les protéines *FimG* et *FimF*. (Alonso-Caballero et al., 2018)

Le pilus de type I joue un rôle important dans la virulence des UPEC. Il permet la colonisation de la vessie en conférant à *E. coli* la capacité d'adhérer aux cellules urothéliales de la paroi de l'urètre et de la vessie. Une étude a également démontré que cette adhésine permet aux UPEC de résister efficacement au flux urinaire et de remonter contre ce dernier. En effet, le pilus de type I, sous l'effet d'un flux antagoniste, est capable, à la façon d'un ressort, d'une part de s'étirer pour réduire les contraintes auxquelles la bactérie est soumise, et d'autre part de se contracter pour permettre à la bactérie de remonter contre le flux (Figure 21) (Rangel et al., 2013).

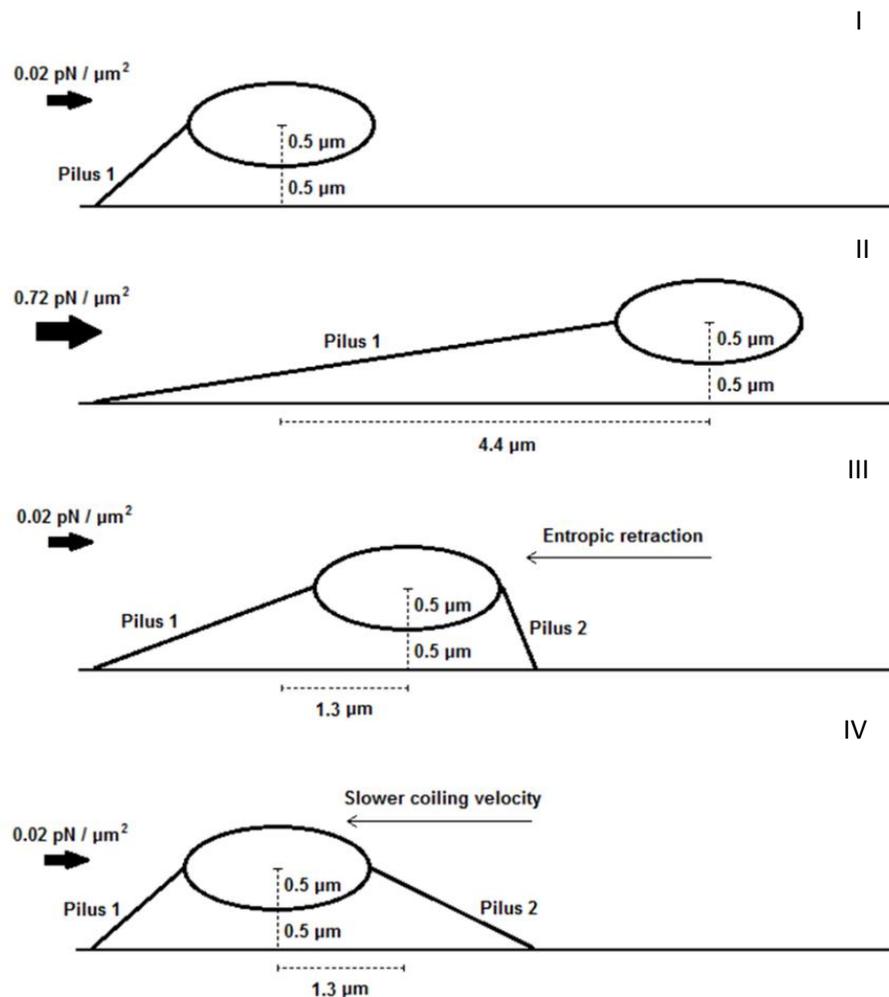


Figure 21 : Représentation d'une bactérie remontant contre un flux antagoniste

Sous l'effet d'un stress lié aux forces de cisaillement causé par un flux liquide, le pilus 1 passe de l'état normal (I) à un état étiré (II) pour réduire les contraintes exercées sur la bactérie. (III) Lorsque le flux diminue, le pilus 1 se rétracte passivement permettant à la bactérie de remonter contre ce dernier. (IV) Enfin, la fixation d'un second pilus provoque l'enroulement du premier permettant à la bactérie de revenir à sa position initiale (Rangel et al., 2013)

Le pilus de type I est également associée à un autre mécanisme de virulence extrêmement important des UPEC : l'internalisation dans les cellules urothéliales. En effet, l'interaction de FimH avec des protéines présentant un domaine cytoplasmique comme l'UP IIIa ou les intégrines $\alpha 3 \beta 1$ semble déclencher une cascade de transduction de signal, conduisant à l'internalisation d'*E coli* dans les cellules superficielles de l'urothélium vésical (Eto et al., 2007 ; Martinez, 2000 ; Martinez and Hultgren, 2002). La capacité du pilus à se rétracter favoriserait cette internalisation en permettant un contact rapproché entre les cellules urothéliales et les bactéries. En effet, dès 1998, Mulvey et al. ont constaté que lorsqu'une souche d'UPEC adhère à l'urothélium, les pili produits sont plus courts (Mulvey et al., 1998). Par la suite, une étude parue en 2018 a montré que certaines mutations de FimA peuvent altérer la capacité du pilus à s'étirer et se rétracter, altérant ainsi la capacité de la bactérie d'une part à coloniser l'arbre urinaire, et d'autre part à infecter les cellules urothéliales (Spaulding et al., 2018).

Le *pyelonephritis associated pilus*, ou *pilus P*, dont l'opéron correspondant (*pap*) est localisé dans un îlot de pathogénicité, est, contrairement au *pilus* de type I spécifique des UPEC. Il est particulièrement exprimé par les souches d'UPEC responsables de pyélonéphrites. Sa structure est similaire à celle du *pilus* de type I. La lectine terminale, PapG, se fixe spécifiquement à un motif α -D-galactopyranosyl-(1-4)- β -D-galactopyranose présent sur certains glycolipides à la surface de l'urothélium rénal (Lund et al., 1987 ; Subashchandrabose and Mobley, 2015). Il existe quatre variantes de l'adhésine PapG (classes I à IV), chacune présentant une spécificité de liaison différente et conférant donc un tropisme légèrement différent. PapG III est plutôt associé avec l'adhésion à la paroi vésicale tandis que PapG II est associé aux souches responsables de pyélonéphrites (Kudinha and Kong, 2022). Il a également été montré que l'expression de ce *pilus* est inversement corrélée à celle du *pilus* de type I, permettant d'aider à la colonisation séquentielle de la vessie puis des reins (Snyder et al., 2005).

Toxines

À ce jour, neuf toxines ont été identifiées dans l'arsenal de virulence des UPEC (Tableau 7) (Johnson and Russo, 2018).

Tableau 7 : Liste des toxines pouvant être produites par les UPEC
(Johnson and Russo, 2018)

Toxins ^d	<i>astA</i>	EAST1, heat-stable enteroaggregative <i>E. coli</i> cytotoxin	Yes (79, 287)	No (189, 288)
	<i>cdtB</i>	Cytolethal distending toxin, CDT	Yes (111)	No
	<i>cnf1</i>	Cytotoxic necrotizing factor 1, CNF-1	Yes (289)	Yes (98, 290, 291)
	<i>hly</i>	α -Hemolysin (Hly)	Yes (292)	Yes (100, 293)
	<i>pic</i>	Protein associated with intestinal colonization, PIC	Yes (294)	Yes (18, 284, 294, 295)
	<i>sat</i>	Secreted autotransporter toxin (serine protease)	Yes (20, 267)	Yes (296)
	<i>tsh</i>	Temperature-sensitive hemagglutinin, TSH	Yes (68, 294)	Yes (18, 294)
	<i>upxA (tosA)</i>	RTX toxin	Yes (297)	Yes (298)
	<i>vat</i>	Vacuolating autotransporter toxin	Yes (299, 300)	Yes (284)

Parmi ces toxines, certaines sont classiquement retrouvées chez les *Enterobacteriaceae* comme Pic, Sat et Vat qui appartiennent à la famille des *Serine Protease Autotransporter Toxine of Enterobacteriaceae* (SPATE). Le rôle de ces trois toxines dans la physiopathologie des IU n'est pas clairement établi, mais elles sont toutes trois préférentiellement présentes chez les souches d'UPEC et sont moins représentées chez les souches commensales (Lüthje and Brauner, 2014).

La toxine HlyA est probablement la plus étudiée. HlyA est une toxine dite « *pore-forming* » qui cible les hématies, les leucocytes et les cellules urothéliales. L'opéron *hly* code l'ensemble des protéines nécessaires à la production, l'activation et l'export de la toxine (codée par *hlyA*). Cette toxine joue un rôle pivot dans la virulence des UPEC au sein de l'arbre urinaire. En effet, il a été montré que HlyA permet aux UPEC d'une part d'interférer avec la réponse inflammatoire en provoquant la dégradation de protéines impliquées dans la voie NF κ B, et d'autre part de provoquer l'exfoliation des cellules

urothéliales exposant ainsi les cellules sous-jacentes (Dhakal and Mulvey, 2012). Il a également été montré que cette toxine est surexprimée par les bactéries intracellulaires (Subashchandrabose and Mobley, 2015).

CNF-1 est une toxine produite par certaines UPEC qui est impliquée dans leur fitness au cours d'une IU. Si elle ne joue pas véritablement de rôle dans la colonisation, elle améliore la résistance des UPEC face au système immunitaire en inactivant des effecteurs clés de la phagocytose et du *burst* oxydatif chez les polynucléaires neutrophiles, comme les GTPases RhoA et Rac1. Elle est retrouvée préférentiellement associée aux souches de récurrence (Ejrnæs et al., 2011 ; Subashchandrabose and Mobley, 2015).

Captation du fer

Comme vu précédemment, l'urine est un environnement carencé en certains microéléments, dont le fer. En effet, chez l'Homme, le fer est très peu biodisponible car il est souvent associé à des protéines (lactoferrine, transferrine) ou séquestré dans les hèmes de certaines protéines comme l'hémoglobine ou la myoglobine (Subashchandrabose and Mobley, 2015). En réponse à cela, les UPEC possèdent un certain nombre de systèmes leur permettant de mobiliser le fer de leur environnement. Deux types de systèmes existent, les sidérophores et les hémophores (Figure 22).

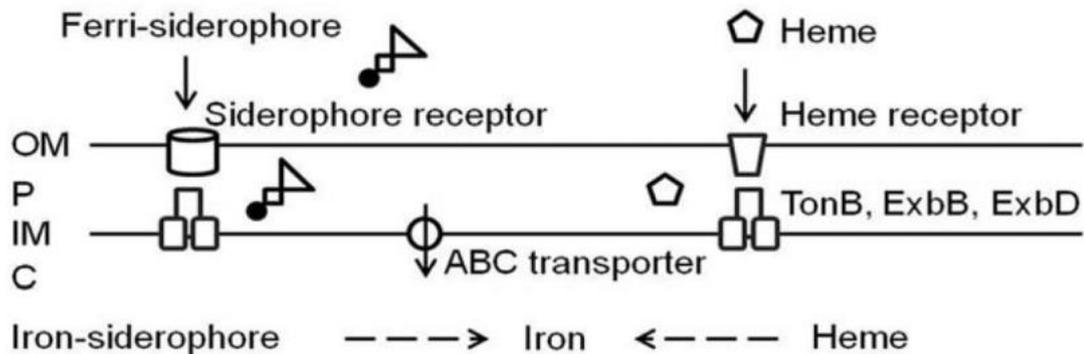


Figure 22 : Fonctionnement des deux types de systèmes de captation du fer chez les UPEC

Les UPEC ont deux moyens de mobiliser le fer présent dans leur environnement. Ils peuvent sécréter une molécule appelée sidérophore qui va chélater le fer présent dans l'environnement puis recapturer cette molécule via un récepteur spécifique. Ils peuvent également exprimer à leur membrane externe un hémophore qui va fixer spécifiquement les hèmes issus de protéines de l'hôte et permettre de faire rentrer ces derniers dans la cellule bactérienne où ils seront dégradés pour libérer le fer qu'ils contiennent. TonB, ExbB et ExbD = système exploitant la force proton-motrice pour l'import des hèmes/ferri-sidérophores dans le périplasme. ABC transporter = ATP Binding Cassette Transporter. (Subashchandrabose and Mobley, 2015)

Sidérophores

Les sidérophores sont des petites molécules sécrétées par une bactérie ayant la capacité de séquestrer le fer. Une fois le fer capturé, le complexe fer-sidérophore se lie à la cellule bactérienne via un récepteur spécifique, puis est importé dans le périplasme. Les UPEC possèdent les systèmes

permettant la synthèse, la sécrétion et la recapture de 4 sidérophores : l'aérobactine (opérons *iuc* & *iutA*), l'entérobactine (opéron *fep*), la salmochéline (opéron *iro*) et la yersiniabactine (opéron *fyu*) (*Subashchandrabose and Mobley, 2015*).

Le fer étant essentiel pour la croissance des UPEC, les sidérophores sont souvent la cible du système immunitaire. Les polynucléaires neutrophiles produisent par exemple la lipocaline II qui séquestre l'entérobactine et empêche sa recapture par la cellule bactérienne (*Subashchandrabose and Mobley, 2015*). Les UPEC contournent cette réponse en produisant la salmochéline, une version glycosylée de l'entérobactine, qui n'est donc pas reconnue par la lipocaline II.

Une étude métabolomique datant de 2009 a mis en évidence que les souches responsables d'IU produisent la salmochéline et la yersiniabactine de façon synchronisée, et à des niveaux significativement plus élevés que les souches intestinales (*Henderson et al., 2009*). De plus, les opérons codant ces deux systèmes sont significativement plus fréquemment identifiés chez les souches responsables d'IU récidivantes (*Ejrnæs et al., 2011*).

Hémophores

Chez les UPEC, deux hémophores ont été décrits : ChuA et Hma. S'il existe des homologues de ChuA chez les EHEC et *Shigella dysenteriae*, le système Hma semble être spécifique des UPEC. Cependant, *in vivo*, des mutants pour le gène *chuA* présentent un fitness réduit par rapport aux mutants du gène *hma*, indiquant que le système ChuA est très important pour la croissance des UPEC dans l'urine. De plus, *chuA* est surexprimé lors de la formation de communautés bactériennes intracellulaires dans les cellules urothéliales (*Subashchandrabose and Mobley, 2015*) ; ceci pourrait expliquer que ce gène soit préférentiellement identifié chez les souches d'UPEC causant des IU récidivantes (*Ejrnæs et al., 2011*).

Biofilms

Les biofilms sont des communautés microbiennes complexes adhérees à une surface et incluses dans une matrice d'exopolymères. Ce mode de vie se distingue du mode de vie planctonique, entre autres par une résistance accrue des cellules bactériennes aux antimicrobiens et au système immunitaire de l'hôte. La présence d'un biofilm peut être à l'origine d'infections chroniques chez l'Homme (*Donlan, 2001 ; Lebeaux et al., 2014*).

Chez les UPEC, la formation de biofilm serait exacerbée par les conditions rencontrées dans l'arbre urinaire. En effet, une étude datant de 2017, a montré que *E. coli* UTI89 produisant plus de biofilm dans un modèle *in vitro*, dans une atmosphère contenant 4% d'O₂, concentration correspondant à celle observée dans la vessie (*Eberly et al., 2017*). Une autre équipe a montré en 2018 que la combinaison

de certains stress mécaniques comme le flux urinaire pourraient exacerber, la production de biofilm par les UPEC (Chu et al., 2018).

Échappement à l'immunité

Les UPEC possèdent deux stratégies pour échapper au système immunitaire. La première consiste à modifier ou masquer leurs structures antigéniques de surface pour ne pas être reconnus par l'immunité de l'hôte, et l'autre consiste à être intrinsèquement moins immunogènes.

Pour masquer leurs antigènes de surface, la plupart des souches d'UPEC sont capsulées. Des études ont montré que la capsule joue un rôle dans la protection des UPEC vis-à-vis du complément (Buckles et al., 2009) ou encore vis-à-vis des polynucléaires neutrophiles (Burns and Hull, 1999). La formation de biofilm et notamment la production des exopolysaccharides de la matrice comme la cellulose, permet de masquer les antigènes de surface des UPEC (Kai-Larsen et al., 2010).

Pour ne pas être reconnu par le système immunitaire, les UPEC ont également acquis des versions modifiées de certains facteurs de virulence. C'est l'exemple de la salmocheline qui est une version modifiée de l'entérobactine (Cf. paragraphe captation du fer).

En étudiant *in vitro* la réponse de l'urothélium à différentes souches de *E. coli*, il est apparu que les souches d'UPEC sont moins efficacement reconnues par les cellules urothéliales que des souches non pathogènes comme *E. coli* K12 (Hilbert et al., 2008). Cette particularité est liée au LPS et au peptidoglycane des UPEC qui semblent être moins immunogènes (Billips et al., 2008). Par la suite, il a également été montré que certains sérotypes associés aux UPEC et notamment les sérotypes O6 (*E. coli* CFT073) et O18 (*E. coli* UTI89 et NU14) résistaient mieux à la phagocytose (Lüthje and Brauner, 2014).

II. 4. 4. Notion de souches de *E. coli* pathogènes hybrides

L'identification en 2011 de la souche de *E. coli* O104:H4 présentant à la fois les caractéristiques d'un EHEC et d'un EAEC (Grad et al., 2012) a fait émerger les notions de souche hétéro-pathogène et de souche pathogène hybride, pour désigner les souches de *E. coli* présentant des caractéristiques de virulence provenant de deux pathovars distincts (Santos et al., 2020). Le terme de souche hétéro-pathogène est réservé aux souches d'InPEC possédant des facteurs de virulence caractéristiques d'au moins deux pathovars intestinaux tandis que le terme de souche pathogène hybride désigne les souches de *E. coli* possédant des facteurs de virulence d'InPEC et d'ExPEC, ou bien sont des InPEC ayant causé une infection extra-intestinale (**Tableau 8**) (Santos et al., 2020).

Tableau 8 : Liste des pathovars hétéro-pathogènes et hybrides ainsi que leurs principales caractéristiques
(Santos et al., 2020)

	Isolation source	Virulence traits		Expression of the virulence traits	Disease	Outbreaks	Serogroup or serotype	MLST-phylogroup ^a	References
EAEC/STEC	Intestinal infections	Stx and aggregative fimbriae with <i>aggR</i> regulator		Yes	Diarrhea, bloody diarrhea, and HUS	Yes	O104:H4 O69:H- Orough:H- O111:H2 O89:H- O23:H28	ST678-B1 ST26-B1 ST1136-B1	Morabito et al., 1998; Iyoda et al., 2000; Frank et al., 2011; Prager et al., 2014; Lang et al., 2018
EPEC/STEC	Intestinal infections in humans, animals' gut, environment, and food	LEE PAI and Stx	Intimin (various subtypes) and Stx	Yes	Diarrhea, bloody diarrhea, and HUS	Yes	O157:H7 O145 O103 O26 O111	ST11-E ST32cplx-D ST20cplx-B1 ST29cplx-B1	Eichhorn et al., 2015
	Normal fecal sample	LEE PAI, Stx-2f, and BFP	Identified only in birds	Yes		No	O137:H6	ST2678-B2	Gioia-Di Chiacchio et al., 2018
EPEC/ETEC	Intestinal infections	LEE PAI and LT or ST	Regardless of the presence of BFP	Yes	Watery diarrhea	No	Unknown	ST278cplx-B1 ST1788-A	Dutta et al., 2015; Askari Badouei et al., 2016; Hazen et al., 2017
ExPEC/STEC and ExPEC/EHEC ^b	Diarrhea and extraintestinal infections simultaneously or extraintestinal infections only	Stx and ExPEC intrinsic virulence factors		Yes—both characteristics are expressed In general, harbor Stx2 variant	UTI, hemorrhagic cystitis, HUS, bacteremia	Yes ^c	O2:H6 O76:H19 Ont:H- O80:H2 O145:H-	ST141-B2 ST675-B1 ST10cplx-A ST165cplx-A ST32cplx-D	Mariani-Kurkdjian et al., 2014; Tovall et al., 2014b; Gati et al., 2019
ExPEC/EPEC	Extraintestinal infections only or diarrhea followed by extraintestinal infection	LEE PAI only or LEE PAI and ExPEC intrinsic virulence factors	Regardless of the presence of BFP	LA or LAL pattern was observed in some studies	Cystitis, pyelonephritis, UTI-related bacteremia, and diarrhea with multiple organ dysfunction	No	O71 O78:H- O4:H1 O12:K1:H-	Unknown ST2018-B2 ST12-B2 Unknown	Voira et al., 2001; Abe et al., 2008; Tovall et al., 2014a; Kessler et al., 2015; Rıveros et al., 2017; Lindstedt et al., 2018; Valiatti et al., 2019
STEC/ETEC	Intestinal infections in human and animals	Stx and ST toxin	Various Stx2 variants; Stx1 is frequently low	Yes, but some strains express just one toxin	Diarrhea, bloody diarrhea, and HUS	No	Various serogroups (>40 serotypes)	ST10cplx-A ST40cplx-B1 ST325-A ST329-A	Nyholm et al., 2015b; García et al., 2018; Bai et al., 2019
ExPEC/EAEC	Extraintestinal infections	Aggregative fimbriae with <i>aggR</i> regulator only or with ExPEC intrinsic virulence factors		Yes ^d	Extraintestinal infections not related to diarrhea	Yes	Various serogroups (O5, O6, O11, O15, O18, O78, etc.)	Various ST10cplx-A is the most frequently reported ST69-D ST131-B2	Abe et al., 2008; Olesen et al., 2012; Tovall et al., 2014a; Lara et al., 2017; Rıveros et al., 2017
ExPEC/ETEC	Extraintestinal infections	ETEC virulence factors only		ST was detected in human, while LT was detected in pigs	Urinary tract infections	No	Unknown	Unknown	de Brito et al., 1999; Rıveros et al., 2017

LEE PAI, locus of enterocyte effacement pathogenicity island; HUS, hemolytic uremic syndrome; BFP, bundle-forming pilus; Stx, Shiga toxin; ST, heat-stable toxin of ETEC; LT, heat-labile toxin of ETEC; UTI, urinary tract infection; LA, localized adherence; LAL, localized adherence-like.

ExPEC intrinsic virulence factors—presence of two virulence factors or genetic markers among the following: *P* fimbriae (*fap*), *S* fimbriae (*stx/foc*), afimbrial adhesin family (*afa/dra*), aerobactin (*iut/iuc*), and capsule from capsular group II (*cpsMTII*).

^aMultilocus sequence typing (MLST) and phylogroup relationship were described as available in *Escherichia/Shigella* Enterobase of Warwick Medical School (Zhou et al., 2020); STn—sequence type number; STn^{cplx}—sequence type complex; ST complex referred to STs that are grouped into clonal complexes by their similarity to a central allelic profile.

^bEHEC is used to refer to strains that harbor LEE and Stx simultaneously.

^cOnly EHEC/ExPEC strains were related to an outbreak.

^dAn aggregative adherence pattern was classified in HeLa or HEP-2 cell lineages. The expression of extraintestinal virulence (capacity to cause UTI or sepsis) was observed for some strains in specific animal models.

III. Cystites récidivantes à *E. coli*

III.1. Fardeau des cystites récidivantes

Comme précédemment décrit, les IU récidivantes peuvent se définir par la survenue d'au moins deux épisodes d'IU sur une période de six mois, ou de trois ou quatre épisodes sur une période de 12 mois, avec entre chaque épisode une guérison documentée (*Hooton, 2001*).

L'essentiel des IU récidivantes sont à type de cystites, survenant chez la femme adulte (chez l'enfant des deux sexes les IU récidivantes sont bien plus rares et le plus souvent liées à une malformation de l'appareil urinaire, soit un cadre très différent qui ne sera pas abordé ici). En moyenne, une femme sur deux aura une IU au cours de sa vie et parmi elles, environ 30% auront un second épisode dans les six mois qui suivent (*Foxman, 2002, 2010*). Si elles n'exposent à aucun risque vital, ces cystites génèrent un inconfort et parfois un réel handicap lorsqu'elles se répètent très fréquemment (par exemple plus d'un épisode par mois, chacun symptomatique). Les IU récidivantes ont en effet un retentissement important sur la qualité de vie comme le montre une étude de 2014 portant sur 575 patients atteints de cystites récidivantes. Cette série a montré que 61,9% des patientes atteintes d'IU récidivantes présentaient des symptômes associés à la dépression (*Renard et al., 2014*). Les IU récidivantes sont également particulièrement associés à une progression alarmante de la résistance aux antibiotiques du fait des traitements itératifs, comme le montre cette étude datant de 2019 dans laquelle les isolats de cystites récidivantes étaient significativement plus résistants que les isolats de cystite simple, et ce pour la quasi-totalité des molécules testées (*Thänert et al., 2019*).

Ces épisodes d'IU récidivantes peuvent être dus à une rechute ou une réinfection, les deux entités pouvant se succéder chez une même patiente.

III.2. Rechutes ou réinfections ?

Une rechute se définit comme étant une nouvelle infection causée par la même souche que celle de l'infection initiale. Une réinfection se définit comme une nouvelle infection causée par une souche différente de celle de l'infection initiale (*Caron et al., 2018 ; Hooton, 2001; Skjøt-Rasmussen et al., 2011*).

La majorité des cystites récidivantes sont dues à des UPEC, de sorte que le distinguo entre rechute et réinfection dépend de la finesse de la méthode utilisée pour comparer les souches : simple antibiogramme en routine (comparaison des profils de sensibilité aux antibiotiques usuels des isolats successifs, ce qui comporte d'importantes limites), typage moléculaire à des fins de recherche ou dans des situations cliniques très particulières (essentiellement pour l'identification de souches épidémiques) (*Glover et al., 2014*).

III.3. Facteurs de risque de récurrence

Les facteurs de risques de cystites récurrentes ont fait l'objet de nombreux travaux du fait du « fardeau clinique » décrit ci-dessus. Plusieurs facteurs liés à l'hôte ont ainsi été décrits. Cependant, à ce jour, aucune caractéristique intrinsèque aux souches ne permet de prédire le caractère récurrent de l'infection.

Les facteurs liés à l'hôte dépendent du statut hormonal de la femme (*Glover et al., 2014*). Chez les femmes non ménopausées, les rapports sexuels fréquents, le changement de partenaire, le recours à des spermicides ou à un diaphragme (méthodes de contraception rares en France), sont des facteurs de risque comportementaux de récurrence bien établis (*Zare et al., 2022*). Il s'y ajoute un facteur anatomique : la brièveté de la distance uréthro-anales (pouvant rendre compte d'une colonisation microbienne plus forte, première étape avant l'infection). Hooton *et al.* ont démontré cet impact, en comparant 98 patientes âgées de 18 à 30 ans atteintes d'IU récurrente, à 107 patientes indemnes (*Hooton et al., 1999*). Enfin, des facteurs familiaux sont identifiés, pouvant impliquer la génétique mais aussi des déterminants comportementaux (mictions retenues par exemple). Ainsi, une étude de 2010 qui portait sur 861 femmes âgées de 18 à 49 ans, dont 431 atteintes de cystite récurrente et 430 patientes contrôle (femmes d'âge similaire sans antécédents d'IU), a montré que chez 70,9% des femmes atteintes de cystite récurrente, au moins un autre cas d'IU avait été rapporté chez une autre femme de la famille, comparativement à 42,4% chez le groupe contrôle (différence hautement significative : $p < 0,001$). De surcroît, le risque individuel de cystite récurrente augmentait avec le nombre de parentes ayant elle-même un antécédent d'IU (*Scholes et al., 2010*).

Après le premier pic de fréquence des cystites récurrentes survenant chez des femmes jeunes et très corrélé à l'activité sexuelle, survient un second pic après la ménopause (*Foxman, 2010*). Il pourrait être expliqué par l'importante diminution de la production d'œstrogènes, s'accompagnant d'une réduction de la part des *Lactobacillus* dans le microbiote vaginal. Or, ces bactéries jouent un rôle protecteur vis-à-vis des uropathogènes en diminuant le pH vaginal (*Raz, 2011*). D'autres facteurs associés à la ménopause ont également été identifiés comme des facteurs anatomiques, telle l'existence d'une cystocèle (*Raz et al., 2000*).

III.4. Physiopathologie des cystites récurrentes : hypothèses actuelles

Pour expliquer les cystites récurrentes, plusieurs hypothèses, ne s'excluant pas mutuellement, ont été formulées, essentiellement pour les UPEC qui, seuls, seront abordés ici. Au fil des décennies, différents scénarios microbiens ont été avancés, ayant pour trait commun une colonisation durable de la patiente concernée par une ou plusieurs souches d'UPEC.

Deux principaux scénarios existent :

- i. La réémergence d'une souche persistante au sein de l'arbre urinaire,
- ii. Des infections ascendantes répétées à partir d'un réservoir extérieur à l'arbre urinaire : réservoir intestinal ou vaginal

III. 4. 1. Persistance vésicale

À ce jour, la persistance vésicale d'UPEC est le mécanisme le plus largement exploré. D'abord étudié dans des modèles animaux, ce scénario a été bien plus récemment étayé par des données cliniques.

Mise en évidence

E. coli a longtemps été considéré comme un organisme pathogène extracellulaire (Silva, 2012). Cependant, dès 1979, des UPEC intracellulaires ont été identifiés dans un modèle d'IU chez le rat. Des images de microscopie électronique à transmission (MET) et à balayage (MEB) ont en effet mis en évidence la présence d'UPEC enfermées dans un endosome, ou libres au sein des cellules superficielles de l'urothélium vésical (Fukushi et al., 1979). Ce résultat a été confirmé par une autre équipe onze ans après cette première étude (Figure 23), dans un modèle d'IU chez la souris, ce qui amorcera alors un changement de paradigme (Mctaggart et al., 1990). Il était à l'époque suggéré que cette internalisation soit un mécanisme de défense de la cellule eucaryote, provoquant à terme la mort de la bactérie.



Figure 23 : Invasion d'une cellule urothéliale par des UPEC observée en microscopie électronique à transmission dans un modèle d'IU murin

La flèche noire indique une bactérie incluse dans un endosome (Mctaggart et al., 1990)

En 2001, Mulvey et al. ont pour la première fois relié l'implication de ce mode de vie intracellulaire avec la persistance des UPEC au sein de l'arbre urinaire dans un modèle murin (Mulvey et al., 2001). Les auteurs ont en effet observé que les bactéries internalisées n'étaient pas tuées par les cellules

urothéliales, et qu'elles pouvaient au contraire tirer bénéfice de cette internalisation, en étant protégée du flux urinaire, et du SI de l'hôte et ainsi persister jusqu'à 6 semaines après leur inoculation dans la muqueuse vésicale de la souris (Mulvey *et al.*, 2001). Il a été montré par la suite que les bactéries internalisées peuvent quitter la vésicule d'endocytose (Pang *et al.*, 2022) pour se multiplier dans le cytoplasme de la cellule urothéliale et former des communautés bactériennes intracellulaires dénommées IBC (*Intracellular Bacterial Communities*) (Justice *et al.*, 2004). La maturation des IBC aboutit à la perforation de la membrane plasmique et à la sortie dans le milieu extracellulaire des bactéries sous forme filamenteuse. Ces bactéries filamenteuses résistent au SI et peuvent infecter d'autres cellules hôtes, démarrant ainsi un nouveau cycle d'infection (Justice *et al.*, 2006) (Figure 24).

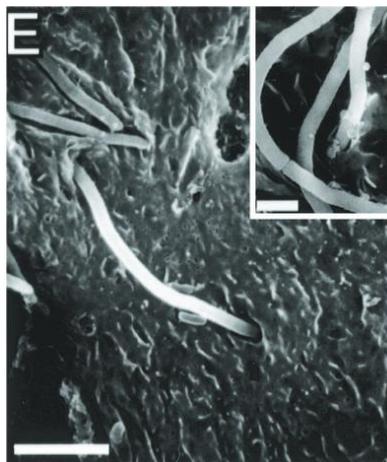


Figure 24 : Filament bactérien interagissant avec deux cellules adjacentes observé en microscopie électronique à balayage.

*L'encart est un agrandissement de l'image au niveau d'un septum partiellement formé, montrant que ces filaments bactériens sont composés de plusieurs cellules bactériennes (Mulvey *et al.*, 2001)*

Les modèles murins suggérant une implication de l'internalisation des UPEC dans la physiopathologie des IU, les études suivantes ont questionné ce phénomène chez l'Homme. En 2007, Rosen *et al.* ont étudié les urines du deuxième jet de 80 jeunes femmes (entre 18 et 41 ans) atteintes de cystite simple et de 20 femmes asymptomatiques ayant des antécédents d'IU. Ils ont mis en évidence par MET et MEB la présence de bactéries sous forme filamenteuse, et de cellules urothéliales exfoliées contenant des IBC chez 40% des femmes avec IU, alors que ces structures n'étaient identifiées chez aucune des femmes asymptomatiques (Rosen *et al.*, 2007). Cette étude a ainsi apporté la preuve de l'existence du phénomène d'internalisation des UPEC chez l'Homme au cours des IU. Puis, en 2013, Robino *et al.* ont rapporté avoir mis en évidence la présence d'UPEC intracellulaires dans les urines d'enfants atteints d'IU récidivantes (Robino *et al.*, 2013). De façon intéressante, les IBC étaient statistiquement plus visualisés dans les urines des patients souffrant d'IU récidivantes et notamment chez les enfants sans anomalies des voies urinaires.

Enfin, des preuves anatomo-pathologiques de l'internalisation d'UPEC chez des patientes souffrant de cystites récidivantes ont été apportées en 2016 par Liu *et al.* Cette équipe a pratiqué des biopsies de vessie chez 32 patientes entre deux épisodes d'IU (asymptomatiques après un traitement antibiotique adapté et présentant un ECBU négatif de moins de 14 jours). Dans 6 cas, des UPEC intracellulaires ont pu être observés (Liu *et al.*, 2016). Ce résultat atteste de la capacité des UPEC à persister au sein de l'arbre urinaire en formant des IBC dans la paroi vésicale, constituant ainsi un potentiel réservoir à l'origine d'IU récidivantes.

Mécanisme moléculaire

En 1998, Mulvey *et al.* ont étudié les mécanismes d'interaction entre UPEC et l'urothélium murin par MET. Ils ont mis en évidence que la cellule bactérienne interagit étroitement avec l'urothélium grâce à ses *pili* de type I. De plus, une souche délétée pour l'adhésine terminale du *pilus* (FimH) n'est plus capable d'adhérer aux cellules urothéliales. Au cours de cette étude, les auteurs ont également observé une internalisation des bactéries par un mécanisme de type « zipper » (interaction directe entre une protéine de la surface bactérienne et un récepteur de la cellule hôte), ressemblant fortement à celui observé pour *Listeria monocytogenes* ou les bactéries du genre *Yersinia* (Figure 25) (Mulvey *et al.*, 1998).

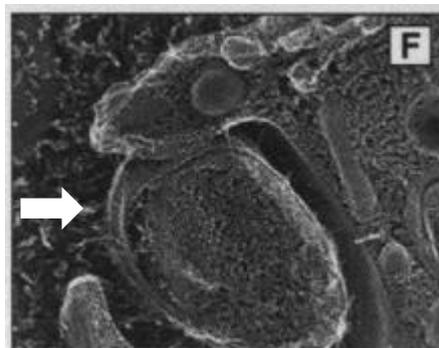


Figure 25 : Bactérie en cours d'internalisation par une cellule urothéliale observée en microscopie électronique à balayage

La bactérie est en cours d'internalisation dans cellule urothéliale dont on voit la membrane encercler l'UPEC par un mécanisme en « zipper ». (Mulvey *et al.*, 1998)

Par la suite, il a été confirmé que FimH est l'effecteur principal de cette internalisation, puisque sa présence seule sur des microbilles de latex est suffisante pour provoquer leur internalisation dans les cellules urothéliales (Martinez, 2000).

Le premier récepteur identifié comme impliqué dans le mécanisme d'internalisation des UPEC est l'hétérotétramère d'UP localisé à la surface des cellules urothéliales. La sous-unité UP la de ce tétramère présente en effet un grand nombre de résidus mannose permettant l'adhésion de FimH. L'interaction UP la - FimH provoque la phosphorylation du domaine cytoplasmique de la sous-unité

UP IIIa, point de départ d'une cascade de transduction de signal aboutissant à l'internalisation de la cellule bactérienne. Cette phosphorylation conduit également à l'apoptose de la cellule infectée, qui peut cependant être bloquée par certains signaux provenant des *Toll-Like Récepteurs* permettant à la bactérie de proliférer dans la cellule eucaryote (**Figure 26**) (Thumbikat *et al.*, 2009).

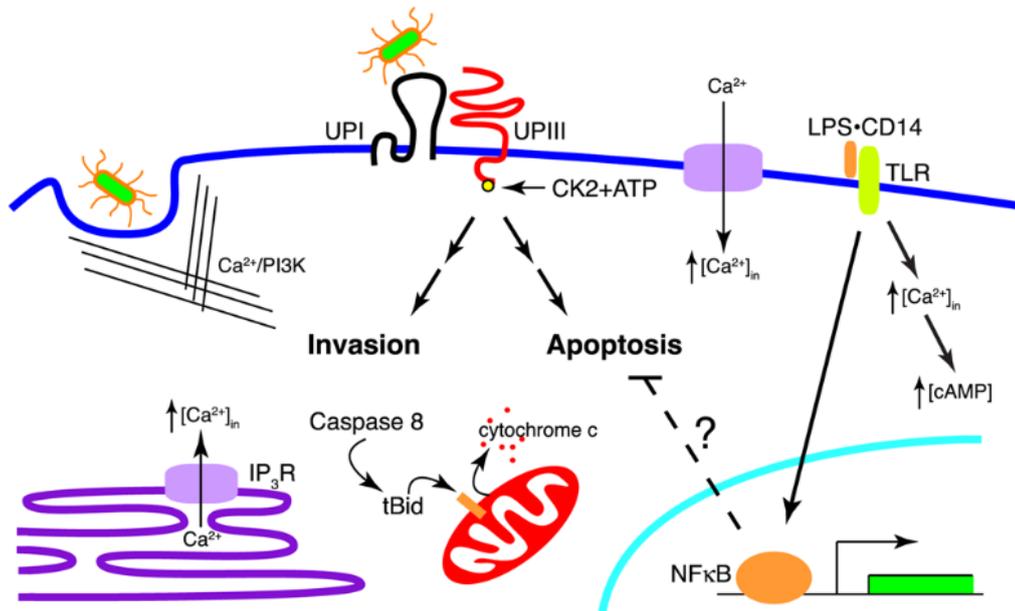


Figure 26 : Conséquences de l'interaction FimH – Uroplakine

L'interaction FimH – Uroplakine (UPI) entraîne la phosphorylation du domaine cytoplasmique de l'UP IIIa par une caséine kinase 2 (CK2). Cette phosphorylation est le point de départ de deux cascades de transduction de signaux. La première implique une élévation de la concentration intracellulaire en calcium et conduit à l'internalisation de la cellule bactérienne. La seconde est moins bien décrite et mène à l'apoptose de la cellule cible. Cette dernière peut être bloquée par des signaux provenant des Toll-like receptors (TLR) via un mécanisme inconnu (Thumbikat *et al.*, 2009).

Un second récepteur cellulaire a également été identifié, tout aussi important que l'UP IIIa. En effet, Eto *et al.* en 2007 mettent en évidence que les UPEC peuvent interagir, via FimH avec des hétérodimères d'intégrine ($\alpha3\beta1$) situés à la surface des cellules urothéliales. Cette interaction conduit à l'activation de voies intracellulaires conduisant au remaniement du cytosquelette de la cellule et à l'internalisation de la bactérie (**Figure 27**). Ce second récepteur explique la capacité des UPEC à infecter les cellules de la muqueuse vésicale qui ne présentent pas d'UP membranaire (Eto *et al.*, 2007).

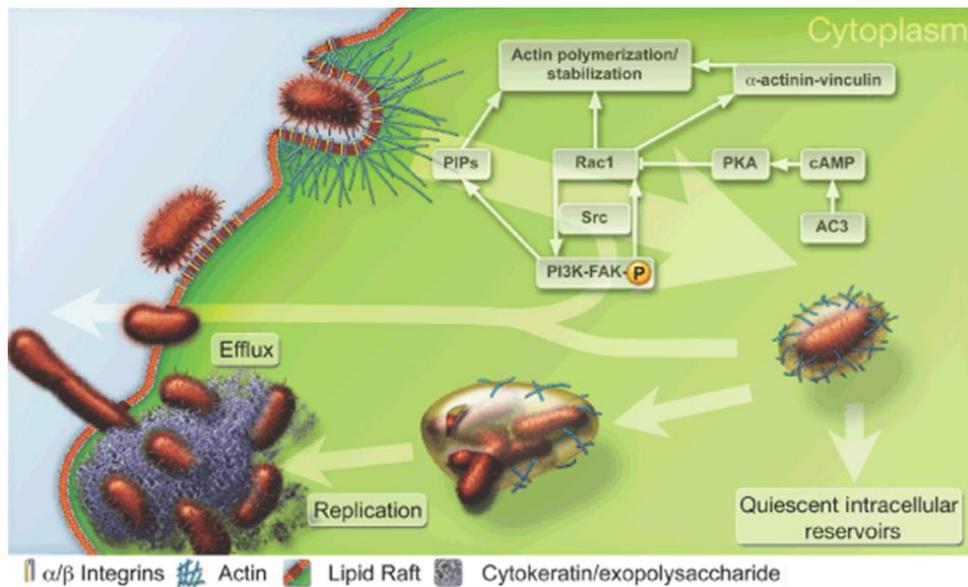


Figure 27 : Mécanisme d'internalisation d'un UPEC

La liaison d'un UPEC au dimère d'intégrines $\alpha 3 \beta 1$ conduit à l'activation de la phosphoinositide 3 kinase (PI3K) qui formera un complexe avec la Focal Adhesin Kinase (FAK). Ce complexe va produire des phosphophosphatidylinositides (PIPs) qui vont jouer le rôle de second messager conduisant au remaniement du cytosquelette d'actine de la cellule et à l'internalisation de la bactérie. En parallèle, il semble que le recrutement de certaines GTPases Rho comme Rac1 et Src contribue également à l'internalisation. (Dhakal et al., 2008)

Une fois internalisées, les bactéries sont incluses dans des vésicules portant des marqueurs caractéristiques des endosomes tardifs (Lewis et al., 2016). Les UPEC inhibent la fusion de ces vésicules d'endocytose avec les lysosomes en partie via le recrutement de la GTPase rho Rab35, empêchant ainsi leur dégradation (Miao et al., 2015).

Le mécanisme de formation des IBC a été étudié en modèle murin en utilisant la souche d'UPEC de référence UT189.

En poursuivant leurs divisions, les UPEC font éclater la vésicule d'endocytose via la production d'une phospholipase (PldA) et se retrouvent dans le cytoplasme de la cellule superficielle infectée. (Pang et al., 2022). Sous cette forme, *E. coli* peut croître de façon rapide avec un temps de génération autour de 35 minutes (versus 20 minutes en conditions nutritionnelles optimales) (Justice et al., 2004). Entre 6 et 8h post-infection, les UPEC commencent à former une communauté bactérienne dense en utilisant les filaments intermédiaires du cytosquelette de la cellule (cytokératine). Cette structuration ressemble fortement à un biofilm, au point que ces IBC sont qualifiées de « *biofilm-like* ». Douze heures après l'initiation de l'infection, les UPEC ressortent des cellules en perçant la membrane plasmique sous la forme de filaments bactériens. Ceux-ci peuvent être éliminés dans l'urine, causant une bactériurie, ou envahir les cellules proximales et recommencer un nouveau cycle d'infection (Figure 28) (Justice et al., 2004 ; Scott et al., 2015).

L'infection répétée de cellules parapluie associée à la production de toxines (Sat et HlyA) provoque une exfoliation massive des cellules urothéliales superficielles, exposant les cellules non différenciées des couches inférieures. Les UPEC peuvent ainsi infecter les cellules de transition. Cependant, leur croissance y est bloquée par le pH intra-vésiculaire acide et la présence d'actine filamenteuse associée aux vésicules d'endocytose. Dans cet état, les bactéries internalisées sont sous forme quiescente et la vésicule est qualifiée de réservoir intracellulaire quiescent (quiescent intracellular reservoir, **QIR**). Chaque QIR contient généralement quatre à dix bactéries non répliquatives. Sous cette forme, les UPEC peuvent persister plusieurs semaines au sein de la muqueuse vésicale, du fait du renouvellement très lent de l'urothélium ainsi que de sa faible perméabilité protégeant les bactéries de l'action du système immunitaire. Les QIR se réactivent ensuite à la faveur de la différenciation des cellules urothéliales en formant de nouveaux IBC et peuvent ainsi constituer des réservoirs à l'origine de récurrences (**Figure 28**) (Eto et al., 2006 ; Lewis et al., 2016 ; Mysorekar and Hultgren, 2006).

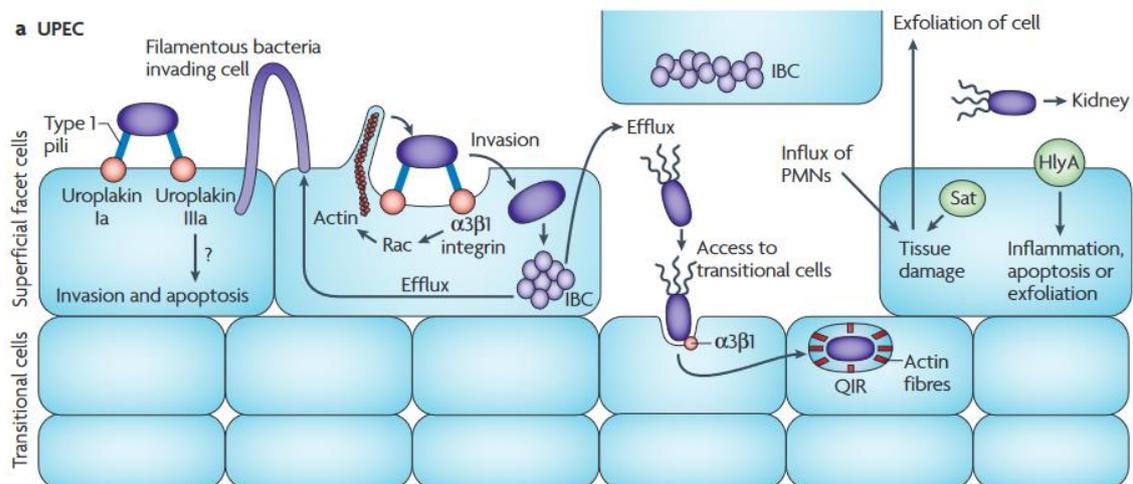


Figure 28 : Formation des communautés bactériennes intracellulaires et des réservoirs quiescents intracellulaires.

Les UPEC adhèrent à l'urothélium via leurs pili de type I. Ils peuvent adhérer à l'uroplakine, ce qui déclenche l'internalisation de la bactérie et l'apoptose de la cellule cible. Les UPEC peuvent également adhérer aux hétérodimères d'intégrine ($\alpha3\beta1$), ce qui va conduire également à leur internalisation selon un mécanisme différent. Dans tous les cas l'internalisation d'une bactérie conduit à la formation d'un « biofilm intracellulaire » appelé communauté bactérienne intracellulaire (IBC). Les UPEC peuvent aller contaminer les cellules proximales en sortant de la cellule infectée sous forme de filament bactérien. Les UPEC produisent également des toxines (Sat & HlyA) ayant pour but de provoquer l'exfoliation des cellules superficielles et rendre accessible les cellules de transition. Les UPEC peuvent ainsi pénétrer dans ces cellules non différenciées via les intégrines pour former des réservoirs intracellulaires quiescents (QIR) qui pourront se réactiver suite à la différenciation de la cellule. (Croxen and Finlay, 2010)

III. 4. 2. Réservoir digestif des UPEC

L'implication de l'intestin dans la physiopathologie des IU récurrentes a été postulée très tôt. L'une des premières hypothèses était que le tractus digestif, si riche en bactéries, joue le rôle de réservoir de bactéries uropathogènes à l'origine de réinfections successives de l'arbre urinaire (Hooton, 2001).

Par la suite, de nombreuses études concernant l'implication du microbiote intestinal dans la physiopathologie des IU ont été menées, conduisant à l'émergence de trois hypothèses (*Worby, Olson, et al., 2022*) :

- (i) Le microbiote intestinal n'influe pas sur la récurrence et n'est qu'un réservoir temporaire pour les bactéries uropathogènes (hypothèse de l'intestin « spectateur » de la récurrence)
- (ii) Une altération du microbiote intestinal facilite le développement des uropathogènes dans l'intestin (hypothèse de l'intestin « acteur » de la récurrence)
- (iii) L'interaction entre le microbiote et l'hôte conduit à une inflammation chronique de la vessie facilitant la récurrence (hypothèse de l'intestin « perturbateur »).

Intestin « spectateur »

Dans cette hypothèse, le microbiote intestinal n'a aucun lien avec la récurrence des IU et l'intestin ne serait qu'un point de passage pour les UPEC (*Worby, Olson, et al., 2022*). Cette théorie est cependant contestée du fait d'études récentes ayant mis en évidence des altérations significatives du microbiote intestinal chez des patientes atteintes d'IU récurrentes (*Worby, Schreiber, et al., 2022*). De plus, une publication de 2017 rapporte plusieurs cas de diminution significative de la fréquence des IU chez des femmes atteintes de cystites récurrentes, suite à une transplantation de flore intestinale (*Tariq et al., 2017*).

Intestin « acteur »

L'intestin contient une « flore digestive » très riche et majoritairement constituée de bactéries anaérobies. Les méthodes traditionnelles avaient quantifié les bactéries cultivables à environ 10^{10} bactéries par gramme de selles et environ 10^{13} bactéries par individu, soit le nombre de cellules composant un humain (*Sender et al., 2016*). Les Entérobactérales – dominées par *E. coli* – étaient quantifiées à environ 10 millions de bactéries par individu. Les approches de métagénomique ont révolutionné ces concepts en démontrant une bien plus grande diversité microbienne de ce qui est aujourd'hui qualifié de microbiote intestinal.

Si le microbiote sain est globalement très résistant à l'implantation des pathogènes intestinaux, il n'empêche pas la survenue d'IU puisque pouvant comporter intrinsèquement des UPEC n'exprimant de virulence que pour l'arbre urinaire. À partir de ce « réservoir digestif », les UPEC peuvent coloniser le méat puis l'arbre urinaire. Des études phylogénétiques supportent fortement ce mécanisme. Ainsi, Thänert *et al* ont mis en évidence chez des patientes atteintes de cystites récurrentes un lien très fort entre les UPEC ayant causé les épisodes et ceux isolés dans la région périnéale et dans les selles (*Thänert et al., 2019*).

Si elle n'est pas indispensable, l'altération du microbiote intestinal favorise l'IU lorsqu'elle conduit à amplifier la densité en UPEC dans l'intestin devenant ainsi acteur. Magruder *et al.* ont ainsi mis en évidence que le risque de présenter une bactériurie ou une IU à *Escherichia* augmente si ce genre est représenté dans l'intestin à plus de 1% (Magruder *et al.*, 2019).

Ainsi, les IU récidivantes peuvent survenir à la suite d'une perturbation du microbiote intestinal, favorisant la prolifération d'UPEC, leur permettant de coloniser la région périnéale puis d'infecter la vessie. Les facteurs provoquant de telles perturbations du microbiote intestinal restent toutefois encore mal connus (Worby, Olson, *et al.*, 2022).

Intestin perturbateur

Aujourd'hui, un certain nombre de pathologies peuvent être au moins en partie expliquées par des altérations du microbiote intestinal. En effet, certains métabolites produits par les bactéries ont un impact positif ou négatif avéré sur d'autres organes (Trompette *et al.*, 2014). En 2022, Worby *et al.* ont mis en évidence une diminution significative de la part des bactéries productrices de butyrate chez des femmes atteintes de cystites récidivantes (Worby, Schreiber, *et al.*, 2022). Cette diminution pourrait avoir comme conséquence une inflammation chronique de la vessie favorisant les récurrences. Le rôle perturbateur de l'intestin dans la pathogénie des IU reste pour le moment peu étudié.

III. 4. 3. Réservoir vaginal d'UPEC

De par sa proximité avec l'urètre, le rôle du réservoir vaginal est suspecté de longue date dans la physiopathologie des IU de la femme, et notamment celle des récurrences.

Le microbiote vaginal des femmes non-ménopausées est dominé par de nombreuses espèces de *Lactobacillus* (*L. crispatus*, *L. jensenii*, *L. gasseri* et *L. iners*). Un microbiote dominé par les bactéries du genre *Lactobacillus* est ainsi considéré comme « sain », alors qu'une dysbiose vaginale est associée à une réduction voire une disparition de ces lactobacilles (Meštrović *et al.*, 2020).

Le lien entre IU et dysbiose vaginale a été démontré dès 2009 (Sumati and Saritha, 2009). En effet, les *Lactobacillus* jouent un rôle protecteur vis-à-vis des uropathogènes en diminuant le pH vaginal via la production d'acide lactique, et en produisant des substances antibactériennes comme l' H_2O_2 ou des bactériocines (Borges *et al.*, 2014). Leur diminution dans le microbiote vaginal rend donc le vagin plus sensible à la colonisation par des uropathogènes comme *E. coli*, et devient dès lors un réservoir de bactéries pour de futures récurrences (Raz, 2011). Une étude a d'ailleurs mis en évidence une association entre colonisation vaginale par *E. coli*, et diminution importante de la part des *Lactobacillus* dans le microbiote vaginal de femmes souffrant d'IU récidivantes (Gupta *et al.*, 1998). À l'inverse l'administration intravaginale de probiotiques (notamment *Lactobacillus crispatus*) a conduit à une

diminution de la fréquence des récurrences chez des patientes ayant des antécédents d'IU récurrentes (*Stapleton et al., 2011*).

De plus, il a été montré récemment que les UPEC sont capables de coloniser et se maintenir efficacement dans la muqueuse vaginale par un mécanisme d'internalisation, établissant ainsi un réservoir permettant d'infecter ensuite la vessie (*Brannon et al., 2020*).

Une dysbiose vaginale est de plus associée à la prolifération de certaines espèces anaérobies dont *Gardnerella vaginalis*. Or, une étude en modèle murin a démontré qu'une exposition de la vessie à certaines souches de *G. vaginalis* provoque la réactivation de QIR formés par des UPEC en déclenchant une apoptose et une exfoliation massive des cellules urothéliales, augmentant le risque de développer une récurrence d'IU (*Gilbert et al., 2017 ; Lewis and Gilbert, 2020*).

III. 4. 4. Influence du microbiote urinaire sur les IU récurrentes

La découverte récente de l'existence d'un microbiote urinaire a engendré un certain nombre d'études mettant en évidence des associations entre la perturbation de ce dernier et diverses pathologies urinaires, comme le cancer de la vessie ou l'incontinence (*Bučević Popović et al., 2018*).

Il n'est donc pas étonnant que son rôle dans la physiopathologie des IU soit en cours d'investigation (*Meštrović et al., 2020*). Il a été observé que ce microbiote est altéré lors d'une infection et est capable de résilience, retrouvant son état originel après la résolution de l'infection (*Bossa et al., 2017*).

Peu de données existent concernant le lien entre IU récurrentes et microbiote urinaire. Une étude de 2022 a comparé les microbiotes de patientes âgées en moyenne de 50 ans atteintes d'IU récurrentes, et de patientes asymptomatiques du même âge. Les auteurs ont mis en évidence des différences significatives de la composition du microbiote urinaire entre ces deux groupes, suggérant l'importance de celui-ci dans la récurrence. L'abondance relative de 36 *taxa* était significativement supérieure dans le groupe des IU récurrentes (**Figure 29**) dont les principaux étaient : *Proteobacteria*, *Burkholderiales*, *Ralstonia*, *Prevotella*, *Dialister* et *Corynebacterium*. À l'inverse, 24 *taxa* étaient significativement moins représentés dans le groupe des IU récurrentes, principalement : *Lactobacillus*, *Gardnerella*, *Streptococci viridans* et *Ezakiella* (**Figure 29**) (*Huang et al., 2022*).

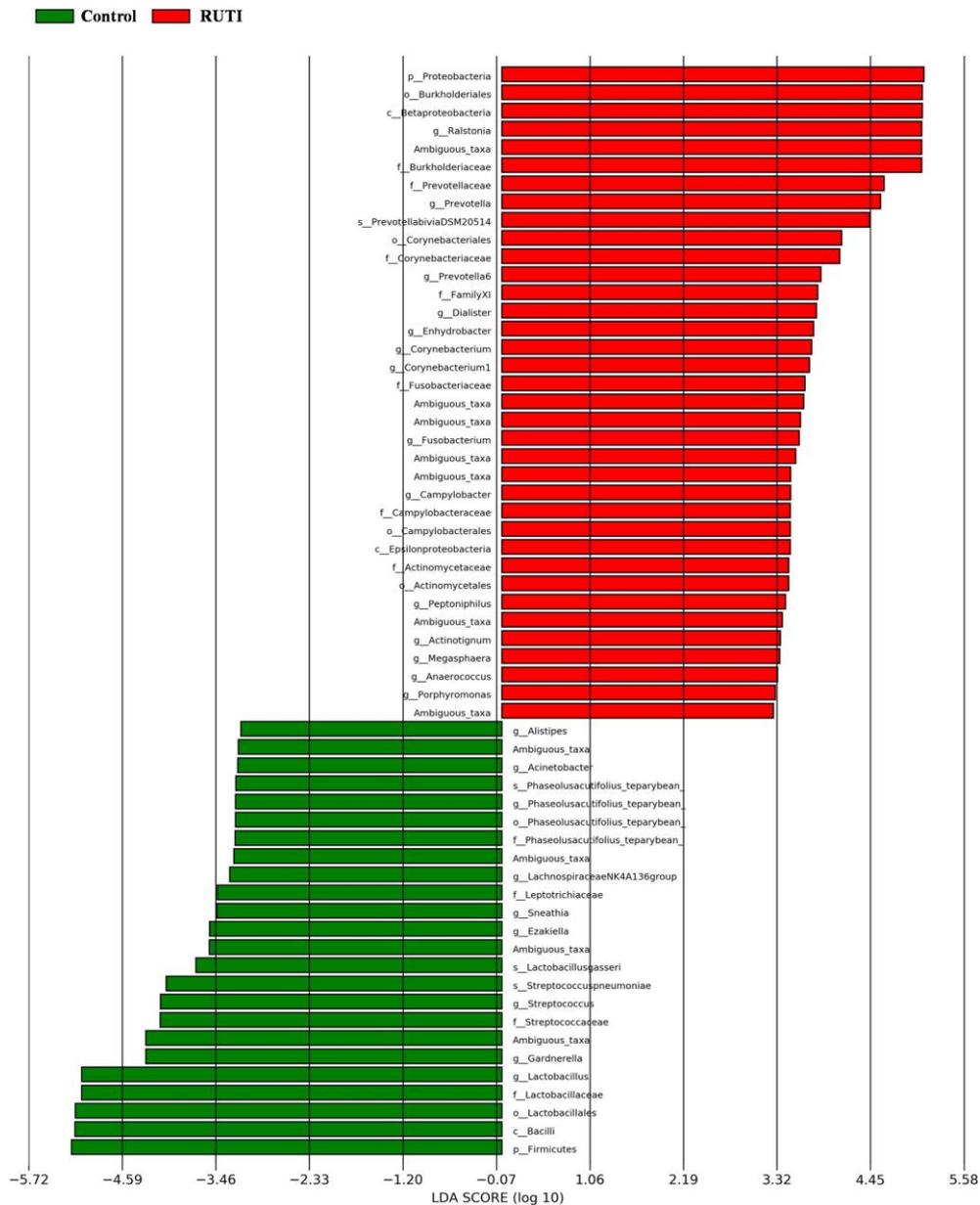


Figure 29 : Groupes de bactéries dont l'abondance relative varie significativement dans le microbiote des patientes atteintes de cystites récidivantes par rapport à un groupe contrôle

LDA score : Linear discriminant analysis score. Quand LDA score > 3 ou LDA score < -3, il y a une différence significative entre les groupes. Quand LDA score < 0, le groupe bactérien est sous-représenté dans les IU récidivantes (Vert). Quand LDA score > 0, le groupe bactérien est sur-représenté dans le groupe des IU récidivantes (Rouge). (Huang et al., 2022)

Une autre étude de 2021 a comparé le microbiote urinaire de patientes atteintes de cystite récidivante (n = 31 patientes) à celui de patientes atteintes de cystite aiguë simple (n = 11). Les auteurs ont mis en évidence une plus grande diversité microbienne chez les patientes récidivantes ainsi que des différences significatives de composition entre les deux groupes. Dans le groupe des cystites simples, *Pseudomonas*, *Acinetobacter* et les *Enterobacteriaceae* étaient plus abondants tandis que dans le groupe des cystites récidivantes, les genres *Sphingomonas*, *Staphylococcus*, *Streptococcus* et *Rothia* étaient plus abondants (Yoo et al., 2021).

Pris dans leur ensemble, ces études suggèrent l'existence d'un rôle du microbiote urinaire dans la physiopathologie des cystites récurrentes. Cependant, d'autres travaux doivent encore être entrepris à de plus larges échelles pour réellement évaluer la place de ce dernier dans la récurrence des cystites.

Il est à noter qu'il a été clairement établi que des interactions et des interconnexions existent entre le microbiote intestinal, vaginal et urinaire. L'intestin, la vessie et le vagin représentent une triade de sites anatomiques impliqués conjointement dans la pathogénèse des IU (**Figure 30**).

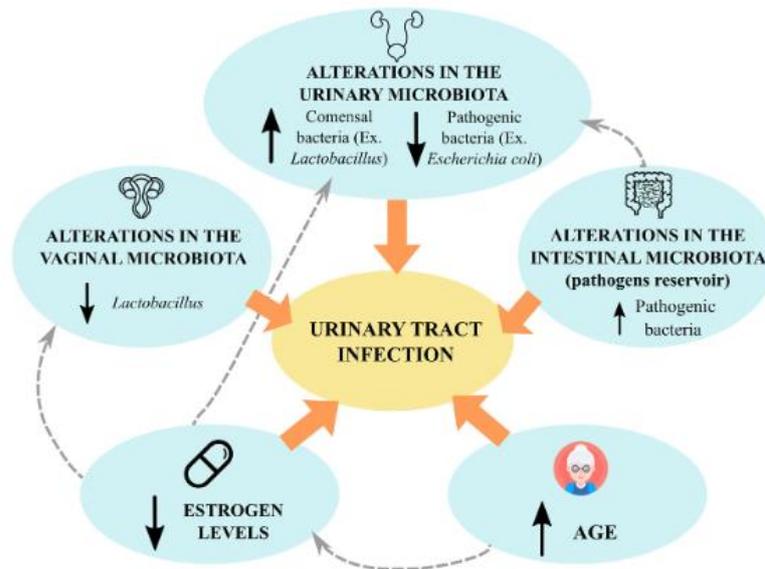


Figure 30 : Implication des trois microbiotes dans la pathogénèse des infections urinaires.

Une altération du microbiote urinaire, vaginal ou digestif peut : (i) avoir un effet délétère sur les autres microbiotes, (ii) augmenter le risque d'apparition d'une infection urinaire (Perez-Carrasco et al., 2021).

Ainsi les cystites, infections bactériennes de la vessie, sont très fréquentes chez la femme, et invalidantes lorsqu'elles récidivent à haute fréquence, pouvant générer alors un véritable handicap. Elles sont à l'origine d'expositions antibiotiques itératives, pouvant sélectionner des pathogènes résistants et transmissibles dans le microbiote digestif, ce qui est particulièrement préoccupant au regard de la pandémie d'Entérobactérales multi-résistantes.

De longue date, il est établi que leur physiopathologie mêle des facteurs favorisants liés à l'hôte et d'autres au microorganisme causal, ce dernier étant quasiment constamment une bactérie. Malgré les nombreuses études publiées, il demeure encore de nombreuses inconnues quant à la dynamique microbienne des cystites récidivantes rendant compte tantôt de rechutes (récidive au même clone) tantôt de réinfection (récidive par clone ou espèce différente).

Dans les deux cas, le principal agent étiologique est *E. coli*. Dotés d'un arsenal de facteurs de virulence particulier, ces UPEC ont en effet la capacité de coloniser puis d'infecter l'arbre urinaire de façon ascendante à partir du périnée, leur réservoir étant le microbiote digestif. De plus, les UPEC présentent la particularité de pouvoir infecter les cellules urothéliales en proliférant dans leur cytoplasme sous forme de communautés *biofilm-like* appelées IBC, impliquées dans les récives.

La découverte récente de l'existence d'un microbiote urinaire est venue complexifier le modèle, questionnant ses interconnexions avec le microbiote intestinal et aussi le microbiote vaginal dans la genèse des IU.

Si de nombreuses études ont été menées pour décrire finement les relations hôte-bactérie lors des cystites, la plupart obéissaient à un design portant sur un épisode unique soit de cystite simple, soit de cystite récidivante. D'où les travaux originaux présentés *infra* visant à analyser longitudinalement des patientes atteintes de cystites récidivantes pour lesquelles le laboratoire dispose d'une succession d'isolats chaînés à des évènements cliniques.

Travaux personnels

I. Première étude : intérêt de l'antibiogramme et du *CH typing* pour l'identification des rechutes de cystite à UPEC

I.1. Contextualisation du travail

Parmi les cystites récidivantes à UPEC qui sont épidémiologiquement très dominantes, la distinction entre rechute et réinfection est capitale pour la compréhension physiopathologique à l'échelon individuel.

Cependant, ceci pose d'emblée la question du niveau de discrimination de la méthode et aussi de sa faisabilité, si ce n'est en routine, du moins dans des délais raisonnables au regard des besoins cliniques.

Or, l'analyse de la littérature a pointé à cet égard une forme d'obsolescence.

En effet, les quelques études historiques longitudinales se sont basées sur une ancienne méthode de référence, l'électrophorèse en champ pulsé (PFGE) (*Glover et al. 2014*), mais dont la granulométrie est en fait très insuffisante, deux clones d'UPEC prouvés comme distincts par NGS pouvant avoir un profil identique en PFGE. D'où l'intérêt de chercher à identifier d'autres méthodes plus simples mais néanmoins de haute résolution, tel que le *CH typing*.

Par ailleurs, en routine la comparaison des isolats bactériens repose sur l'espèce et lorsque celle-ci est identique, sur les profils d'antibiogramme. Cela est fait souvent à l'échelon individuel pour faire la part entre rechute et réinfection ou affirmer la source d'une infection (souche identique dans les urines et le sang permettant de retenir le diagnostic de « bactériémie à porte d'entrée urinaire »), ou à l'échelon collectif pour traquer les épidémies (communautaires ou associées aux soins). Or, pour les UPEC des cystites récidivantes, les études du siècle passé avaient montré l'incapacité des antibiogrammes à distinguer rechute et réinfection, dans un contexte où la grande majorité des souches étaient de phénotype sauvage (c'est-à-dire sensibles à l'ensemble des molécules testées). Or, avec l'évolution croissante de l'antibiorésistance chez les Entérobactérales au cours de ces deux dernières décennies, le profil sauvage est devenu très minoritaire chez les UPEC. Paradoxalement, cette évolution épidémiologique préoccupante pourrait receler un avantage, en augmentant la performance de « l'antibiotypage » au sein de cette espèce, mais ceci reste à évaluer en l'absence de données bibliographiques actualisées à cet égard.

D'où, l'objectif de mon premier travail personnel visant à requestionner les performances des méthodes de typage des UPEC de cystites récidivantes pour distinguer rechute et réinfection, avec pour « gold standard » le NGS, tout en évaluant le *CH-typing* et l'antibiogramme.

A cette fin, nous avons bénéficié d'une cohorte prospective de 323 patientes atteintes de cystites, sporadiques ou récidivantes, préalablement constituée lors d'un programme de recherche sur la performance du système des soins (Protocole VITALE, NCT02292160) mené au laboratoire Dynamique UMR INSERM 1311 (Ex EA2656 GRAM 2.0) entre septembre 2015 et janvier 2017.

J'ai réalisé personnellement le typage moléculaire de l'ensemble des 347 isolats d'UPEC collectés lors de ce protocole par phylogroupage, et *CH typing*.

J'ai aussi colligé et comparé les antibiogrammes qui avaient été précédemment réalisés par automate et confirmé la production de BLSE par méthode des disques.

Réalisées en inter-groupe (cystites sporadiques versus cystites récidivantes) et en intra-patient pour les cystites récidivantes, les analyses ont apporté des résultats originaux par rapport aux études antérieures en positionnant le *CH typing* comme une méthode de haute résolution chez les UPEC, permettant de s'affranchir du NGS pour cet objectif. Mes travaux ont également souligné l'intérêt de l'antibiotypage en première approximation, dont les performances de discrimination connaissent un bond dans le contexte de l'antibiorésistance élevée des Entérobactéries

I.2. Publication n°1

Contribution of antibiotic susceptibility testing and CH typing compared to Next-generation sequencing for the diagnosis of recurrent urinary tract infections due to genetically identical *Escherichia coli*: a prospective cohort study of cystitis in women

Nicolas Vautrin, Kévin Alexandre, Martine Pestel-Caron, Estelle Bernard, Roland Fabre, Marie Leoz, Sandrine Dahyot, François Caron

Article en préparation pour une soumission dans la revue *Frontiers in Microbiology*

Ce travail a fait l'objet de deux communications orales dans deux congrès nationaux (RICAI, Paris 2017 et 2019) et d'une communication en ligne dans un congrès international (ECCMID 2020)

1 Contribution of antibiotic susceptibility testing and CH typing compared to Next-Generation
2 Sequencing for the diagnosis of recurrent urinary tract infections due to genetically identical
3 *Escherichia coli*: a prospective cohort study of cystitis in women

4 Nicolas Vautrin¹, Kévin Alexandre², Martine Pestel-Caron³, Estelle Bernard⁴, Roland Fabre⁵,
5 Marie Leoz¹, Sandrine Dahyot³, François Caron²

6 ¹: Normandie Univ, UNIROUEN, UNICAEN, Inserm UMR 1311 DYNAMICURE, 76000
7 Rouen, France

8 ²: Normandie Univ, UNIROUEN, UNICAEN, Inserm UMR 1311 DYNAMICURE, CHU
9 Rouen, Department of Infectious diseases, F-76000 Rouen, France

10 ³: Normandie Univ, UNIROUEN, UNICAEN, Inserm UMR 1311 DYNAMICURE, CHU
11 Rouen, Department of Microbiology, F-76000 Rouen, France

12 ⁴: Department of Microbiology, CHU Rouen, F-76000 Rouen, France

13 ⁵: Laboratoire d'analyses médicales, 76500 Elbeuf, France

14 **Abstract**

15 Recurrent cystitis is a common disease in women, mainly due to uropathogenic *Escherichia*
16 *coli* (UPEC). For decades, typing methods now considered as obsolete suggested that relapse
17 by the same clone is dominant over reinfection, most UPEC being otherwise fully susceptible
18 to antibiotics. We aimed to update these data. Thanks to a prospective study over 17 months we
19 recruited 323 women with cystitis. 251 of them had non-recurrent infection and 72 had
20 recurrence, with 2 to 9 episodes per patient for a total of 131 UPEC isolates and 145 UPEC
21 pairs at patient level. Phylogroups B2 (52.4%) and D (14.1%) were overall dominant, as
22 expected due their particular urovirulence. CH typing identified 119 distinct profiles with no
23 CH-type particularly associated with recurrence. Relapse was attested by CH-typing for only
24 30.6% (22 out of 72), with very diverse situations ranging from all episodes due to the same
25 clone to alternating reinfections and relapses. Next-Generation Sequencing confirmed the
26 clonality for all but one out the 145 UPEC pairs. Antibiotic resistance was common for recurrent
27 cystitis isolates (only 25 [17.2%] out of 145 UPEC pairs were fully susceptible), allowing to
28 predict UPEC clonality. Indeed, antibiotic susceptible profile matched with CH-typing for 104
29 (71.7%) pairs. Finally, we demonstrated a large genetic diversity among UPEC responsible for
30 cystitis in women, even in case of recurrence for which reinfection appeared dominant over
31 relapse. Recurrent cystitis appears to be a heterogeneous disease requiring tailored treatment
32 and prevention.

33 **Importance**

34 More than half of women will experience cystitis during their lifetime. Among these women,
35 25% will experience a second episode in the next six months. It is epidemiologically important
36 to discriminate relapses from reinfections. Relapse identification relies on long and laborious
37 methods and might influence treatment. Therefore, the designation of time- and cost-effective
38 strategies for this goal is of particular interest. Our work suggests using CH-typing and
39 antibiotic susceptibility profiles to type *Escherichia coli*, the main uropathogen.

40 **Introduction**

41 Cystitis is by far the most common presentation of urinary tract infections (UTI), sex-linked,
42 affecting 40% of women during their lifetime (1), from teenage years to old age. Usually
43 considered as a benign but uncomfortable condition, cystitis induces a substantial healthcare
44 burden, being one of the leading causes of antibiotic treatment in otherwise healthy women (2).
45 Cystitis is a dichotomous disease: while three quarters of affected women have only few
46 episodes (*i.e.*, sporadic cystitis, SC), the last quarter suffers from recurrent cystitis (RC),
47 commonly defined as at least 2 episodes within 6 months or 3 episodes within 12 months (3).
48 When highly frequent, RC strongly affects the quality of life as a real handicap, worsened by
49 the risk of antibiotic resistance and treatment escape (4).

50 As pointed by a recent review “*relatively little is understood about the infection cycle of RC in*
51 *humans*” (2) despite numerous *in vitro* studies focusing on bacterial virulence factors
52 particularly for uropathogenic *Escherichia coli* (UPEC), the leading species that accounts for
53 about 80% of total UTI cases. For decades, the ascending route of cystitis has been accepted,
54 considering that uropathogens usually come from the reservoir of faecal microbiota, colonise
55 the periurethral area before ascending into the bladder. Historical studies using pulse-field gel
56 electrophoresis, now recognized as insufficiently discriminant (5) has reported that a large
57 majority of RC are due to reinfection by primary strain (6–8). At that time, comparing antibiotic
58 susceptibility profiles (ASP) between UPEC was of poor interest in suspecting clonal isolates,
59 due to the low incidence of antibiotic resistance so that most lineages shared an identical wild-
60 type phenotype (9).

61 Recent data support a more multi-faceted mechanism of RC depending on patient, involving
62 faecal abundance of uropathogens, vaginal dysbiosis and persistence of UPEC within bladder
63 epithelial cells as quiescent intracellular reservoirs (QIR) (2).

64 A better understanding of RC, in particular the real proportion of relapse by the same clone
65 versus reinfection by another lineage, would be of great interest, conceptually and practically,
66 with the goal to move more frequently towards tailor-made treatment.

67 Thus, thanks to a large prospective cohort of patients with well characterized UTI and to typing
68 methods such as NGS, we describe here the current bacterial epidemiology of cystitis in women,
69 with a particular attention to UPEC recurrences.

70 **Methods**

71 ***Ethics***

72 This study was part of an epidemiological study on community acquired UTI founded by the
73 French Ministry of Health and approved by the Medical Research Ethics Committee of the
74 Rouen University Hospital (VITALE study, ClinicalTrials.gov, identifier: NCT02292160).
75 Participating patients received an information letter and provided written informed consent.

76 ***Patients***

77 Adolescent and adult women with community-acquired cystitis were prospectively included
78 over a 17-month period (September 2015 – January 2017) by a dedicated staff in a laboratory
79 deservng outpatients analysing around 4000 urine samples per year (Laboratoire d'Analyses
80 de Biologie Médicale, Elbeuf, France). Cases with negative urine culture or catheter related
81 infections were excluded. After the interview of the patient, and when necessary of her family
82 doctor (either by email or phone), UTIs were classified as uncomplicated or complicated
83 according to national guidelines (10). RC was defined as at least 2 episodes of cystitis within 6
84 months (3). Relapse was defined as a recurrence due to identical isolates by typing. Reinfection
85 was defined as a recurrence due to genetically different isolates.

86 ***Bacterial identification***

87 Identification to the species level was performed routinely using either biochemical method
88 (VITEK[®]2 system, bioMérieux, Marcy-l'Étoile, France) or matrix-assisted laser desorption
89 ionisation time-of-flight mass spectrometry (Brucker, Billerica, MA, USA). After
90 identification, isolates were stored at -80°C using cryobeads system (Technical Service
91 Consultants Ltd, Heywook, United Kingdom). Duplicate specimens defined as the same species
92 isolated from a given patient within 2 weeks from the first positive culture were excluded.

93 ***Antimicrobial susceptibility testing of UPEC***

94 ASP were determined using the VITEK[®]2 automated system. Extended spectrum beta-
95 lactamase producing isolates suspected by VITEK[®]2 were confirmed by double-disk synergy
96 test following the recommendations of the European Committee on Antimicrobials
97 Suceptibility Testing (EUCAST) (<https://www.eucast.org>).

98 ***UPEC molecular typing***

99 DNA from UPEC isolates was extracted using the InstaGene Matrix kit (Bio-Rad, Hercules,
100 CA, USA) according to the manufacturer's recommendations after overnight culture at 37°C
101 on ChromID® CPS Elite (bioMérieux, Marcy-l'Étoile, France).

102 Phylogroup assignation (A, B1, B2, C, D, E and F) was performed using the revised Clermont
103 method (11–15) with minor modifications. (Table S1 & S2).

104 CH (*fumC_fimH*) typing was performed as described by Weissman *et al.* (16) with minor
105 modifications. PCR amplification was conducted under the conditions described by Tartof *et*
106 *al.* (17, 18) for *fumC* (fumarate hydratase coding gene) and by Weissman *et al.* (16) for *fimH*
107 (type 1 fimbrin D-mannose specific adhesion protein coding gene) amplification (Table S1 &
108 S2). PCR products were purified and sequenced using Sanger technology by Eurofins GATC
109 (Konstanz, Germany). All PCR reactions for phylogroup assignation and CH typing were
110 performed on a Veriti Thermal Cycler (Applied Biosystems, Foster City, CA, USA).

111 Allelic profiles and corresponding CH types were assigned using BioNumerics software
112 (Version 7.6, Applied System, Sint-Martens-Latem, Belgium) by using the databases of the
113 online tool CHtyper 1.0 (19). All CH types were named as “(*fumC* allele number_*fimH* allele
114 number)”. Phylogeny based on concatenated sequences of *fumC* and *fimH* was used to construct
115 minimum spanning trees with BioNumerics software using unweighted pair group method with
116 arithmetic mean (UPGMA). Simpson's diversity index (DI) (20) and associated confidence
117 intervals (CI) (21) were calculated using the online tool available at
118 (<http://www.comparingpartitions.info/?link=Tool>).

119 ***NGS sequencing and bioinformatics***

120 UPEC isolates from a given RC patient assigned to the same phylogroup and the same CH type
121 were suspected not to be genetically different and selected for further analysis using whole
122 genome NGS.

123 Overnight shaking trypticase soja cultures (Bio-Rad, Hercules, CA, USA) were centrifugated
124 during 10min at 2,000 G. Bacterial pellets were resuspended in a mix of 100µL of sterile PBS,
125 100µL of MagNA Pure bacteria lysis buffer (Roche life science, Penzburg, Germany) and 40µL
126 of a proteinase K solution at 20mg/mL (Sigma-Aldrich, Saint-Louis, MO, USA). Lysis mix was
127 then heated to 65°C for 10min and to 95°C for 10min. Finally, bacterial lysates were frozen and
128 then sent to the Mutualized Plateform for Microbiology of the Pasteur institute (P2M:
129 <https://research.pasteur.fr/en/team/mutualized-plateform-for-microbiology/>) which performed

130 DNA purification on the MagNA Pure 96 system (Roche Diagnostics, Bâle, Switzerland),
131 library preparation using the Nextera XT library kit (Illumina Inc., San Diego, CA, USA). and
132 paired-end sequencing on the NextSeq 500 system (Illumina Inc.) with average read length of
133 150pb. The reads were demultiplexed using bcl2fastq v2.20.0
134 (emea.support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html), quality
135 controlled using fastq_info v2.0 (gitlab.pasteur.fr/GIPhy/fastq_info) and taxonomy controlled
136 using using Kraken v2.1.1 (github.com/DerrickWood/kraken2). Quality filtering was
137 performed using fqCleanER v21.06 (gitlab.pasteur.fr/GIPhy/fqCleanER) and contaminant
138 reads were removed using BBDMap v38.94 SEAL (sourceforge.net/projects/bbmap/).

139 Intra-patient sequential UPEC isolates belonging to the same CH type and phylogroup were
140 then pairwise compared with the first occurrence by whole genome SNPs analysis. Short reads
141 from the initial occurrences were assembled by P2M using fq2dna v21.06
142 (gitlab.pasteur.fr/GIPhy/fq2dna) and contaminant reads were removed using BBDMap v38.94
143 bbsplit (sourceforge.net/projects/bbmap/). These assemblies were then used as references for
144 mapping the short reads from corresponding subsequent occurrences and SNP analysis using
145 Snippy v4.3.6 (github.com/tseemann/snippy). A divergence rate (number of SNPs per day) was
146 determined using the dates of isolate collection. Isolates with a divergence rate inferior to 3.5
147 SNPs/day were classified as identical according to Petronella *et al* (22) and considered
148 responsible for a relapse.

149 ***Data analyses***

150 All UPEC from a given patient with recurrent cystitis were compared pair by pair. Pair numbers
151 were determined using following formula:

$$152 \quad \frac{n!}{k!(n-k)!}$$

153 Statistical analyses were performed using R (v4.1.3, R foundation, Vienna, Austria).
154 Comparisons between proportions were performed using Pearson's chi-squared test or Fisher's
155 exact test for count data when possible. Means were compared using Wilcoxon rank sum test
156 with continuity correction.

157

158 **Results**

159 ***Cohort description***

160 During the 17-month period, 445 bacterial isolates were collected from 323 women with
161 cystitis, of whom 251 were classified as having SC and 72 as patients with RC. Patients from
162 SC group were younger (mean age of 54.3 vs 58.6 years for RC group, $p = 0.048$) and had lower
163 prevalence of complicated cystitis (29.5% vs 65.3%, $p < 0.001$).

164 Overall, 265 isolates were collected among the 251 patients with SC (some patients had 2
165 reinfections more than 6 months apart during the study) and 180 isolates were collected among
166 the 72 patients with RC (2 to 9 episodes per patient). As shown in Table 1, more than three
167 quarters of isolates were identified as *E. coli* ($n = 347$) with a significantly higher representation
168 in SC group than in RC group (81.5% vs 72.8%, $p = 0.039$).

169 ***UPEC antibiotic susceptibility profile***

170 As shown in Table 2, there was among UPEC isolates an overall low rate of resistance (<5%)
171 to fosfomycin, nitrofurantoin, pivmecillinam, amoxicillin-clavulanic acid, and third generation
172 cephalosporin (3GC). For these antibiotics, no significant differences in resistance rate were
173 observed between SC and RC groups, except for 3GC (higher resistance rate in the RC group,
174 mainly due to ESBL production). In contrast, resistance was common and significantly higher
175 among RC than SC for amoxicillin, trimethoprim-sulfamethoxazole and quinolones.

176 ***UPEC isolates typing***

177 UPEC phylotyping revealed that more than half of the 347 isolates belonged to group B2
178 (182/347, 52.4%) followed by group D (49/347, 14.1%) (Table 1). There was no significant
179 difference between SC and RC when phylogroups were considered separately, except for the
180 group D (18.5% vs 6.9% respectively, $p = 0.004$). Phylogroups B2 and D, associated with
181 uropathogenic strains (23), were more frequent in SC than RC (73.6% and 55.0% respectively).

182 CH typing revealed a great genetic diversity, as 119 distinct CH types were identified for the
183 347 UPEC isolates. Five major CH types ([35_27], [14_27], [52_5], [11_54], [40_30]) were
184 shared by 104 isolates, accounting for 30.0% of the total set (Figure 2A, Table S3). In contrast,
185 72 CH types (60.5%) were identified as singletons (*i.e.*, in only one patient), in similar
186 proportions in SC group (43 singletons out of 84 CH types, 51.2 %) and in RC group (29 out
187 of 64, 45.3 %). Ninety CH types were observed in only one of the two clinical groups *i.e.*, either
188 in the SC group or in the RC group (Figure 1). The Simpson's diversity indexes (DI) attested
189 from this great diversity for the total cohort (DI = 0.976; 95% CI = 0.970-0.982), as well as for

190 the SC group (DI = 0.966; 95% CI = 0.955-0.977) and the RC group (DI = 0.979; 95% CI =
191 0.972-0.987).

192 The phylogenetic tree did not highlight any genetic links between CH types and clinical groups
193 (Figure 2 A). In contrast, results of CH typing and phylotyping were overall consistent (Figure
194 2B). Indeed, isolates belonging to the same phylogroup were assigned to genetically linked CH
195 types. The only exception was for group F isolates, which did not cluster together on the tree.

196 *Longitudinal study of RC patients*

197 As shown in Table 3, 72 patients suffered from RC during the study period, for a total of 180
198 episodes (separated by an average of 111 days). Forty patients (55.5%) presented at least two
199 episodes with UPEC, of whom 22 patients (30.6%) had at least two episodes with UPEC
200 assigned to the same phylogroup and CH-type. NGS analysis of the 35 suspected identical
201 UPEC pairs (60 UPEC isolates) from these 22 patients revealed a small number of SNPs
202 between isolates from most patients (0.39 SNPs/days on average, ranging from 0.00 to 2.15
203 SNPs/days). Of note, all UPEC suspected to be identical by phylogroup and CH type were
204 confirmed identical by SNPs analysis (Table 3), except for one isolate pair (46 SNPs/days).
205 Figure 3 shows the UPEC history for each of the 22 patients, showing very diverse situations.
206 Seven patients (# 12 and 16 to 22) had only episodes with identical UPEC, 11 patients (# 2-3,
207 6-11, 13-15) were infected by 2 different UPEC, 2 patients (# 4-5) by 3 different UPEC and 1
208 patient by 5 different UPEC.

209 As shown in Table 4, ASP matched with CH typing for 104 (71.7%) of the 145 UPEC pairs. Of
210 note, while 51 (38.5%) out of the 131 UPEC from RC displayed a wild-type phenotype, only
211 24 (16.5%) out of the 145 pairs shared this ASP.

212 **Discussion**

213 Distinguishing between reinfection and relapse helps physicians in the management of both
214 uncomplicated and complicated RC (10). Diagnosis of reinfection is easy when all subsequent
215 episodes for a given patient are due to different bacterial species, but such events are
216 uncommon: in our series of 72 patients suffering from RC, only 16 (22.2%) met this criterion.
217 Indeed, UPEC accounted for 72.8% of the episodes of our RC patients, a percentage in phase
218 with the literature data (24, 25) As reported by others (26, 27), UPEC were less prevalent in RC
219 than SC (72.8% versus 81.5 %, $p < 0.05$) and this was correlated with a higher rate of
220 complicated UTI among RC. This is consistent with the fact that urological abnormalities,
221 which are the leading cause of complicated UTI, favour colonization and infection even by
222 microbial species less urovirulent. As expected, we observed an under-representation of
223 phylogroups B2 and D when pooled together among RC compared to SC (55.0% versus 73.6%,
224 $p < 0.001$). Indeed these two groups are known to express more urinary virulence factors (23).

225 Thus, the major challenge in the diagnosis of reinfection versus relapse is to discriminate
226 between UPEC recurrences. Due to increasing of antibiotic resistance among Enterobacterales
227 over decades, wild-type ASP (*i.e.* no acquired resistance mechanism expressed) is now
228 uncommon and thus ASP are more diverse. In our series, among the 145 pairs of UPEC at a
229 patient level only 24 (16.5 %) demonstrated wild-type ASP. This dramatically illustrates the
230 deleterious effects of iterative antibiotic treatments, particularly on the rate of resistance to
231 3GC, quinolones and trimethoprim-sulfamethoxazole. Such a worrisome trend has made ASP
232 an interesting tool for suspecting clonality. In our work, ASP matched with CH-type for 71.7%
233 of the UPEC pairs at a patient level. Among the discrepancies, 61.0% (25/41) were due to
234 identical ASP corresponding in fact to different CH-types, and 39.0% (16/41) to a same clone
235 that diverged due to the expression of new antibiotic resistance over time under the treatment
236 pressure. Such evolution for urine as well as gut isolates has been recently longitudinally
237 analysed over 4 years in a patient suffering from RC due to a clonal UPEC lineage that presented
238 rapid adaptation to iterative antibiotic exposure (28).

239 Five previous studies have analysed clonality among UPEC subsequently isolated from patients
240 with RC using either pulsed-fielded gel electrophoresis or serotyping (6–8, 29, 30). All five
241 have concluded to a high rate of clonality (68 to 77%). Using more discriminant methods, we
242 describe a less than one third of patients having at least one pair of identical UPEC among their
243 different cystitis episodes: 30.6% according to CH typing and 29.2% by NGS. The high
244 performances of CH typing suggest that this method could be a reasonable alternative to more

245 time-consuming and costlier NGS technologies to distinguish relapse from reinfection by
246 UPEC. Interestingly, in in this study, the delay between clinical episodes due to genetically
247 identical UPEC was long (median time: 111 days), and thus strongly in favour of true relapse
248 rather than antibiotic failure to eradicate the initial infection. Whether *E. coli* isolate causing
249 relapse persists within urothelium as QIR and/or within digestive microbiota as dominant
250 *Enterobacterales* population needs to be prospectively studied at the patient level.

251 Our series shows an overall high diversity among UPEC responsible for either RC or SC (Figure
252 2A). These results are in line with a recent Danish study reporting a SNP based phylogeny of
253 156 UPEC which shows that RC isolates did not constitute distinct monophyletic clusters (31).

254 The strengths of our study were a recruitment of cystitis clinically well characterized and an
255 exhaustive microbiological analysis. The dominant limitation was the relatively short period of
256 time (17 months) of follow-up whereas it would be interesting to follow bacterial evolution at
257 patient level for a much longer time as performed by others (28).

258 Finally, we demonstrated a large genetic diversity among UPEC isolates from cystitis in
259 women, even for those suffering from recurrent infection, while historical studies using less
260 discriminant typing method had concluded to a high rate of clonality for these latter. Thus, the
261 physiopathology of recurrent cystitis seems more complex than previously assumed This might
262 have important consequences in terms of treatment and prevention. Due to the now high level
263 of antibiotic resistance of UPEC responsible for RC, comparison of ASP of isolates represents
264 in routine a way to approach clonality at patient level. For a more precise analysis, CH typing
265 offers in our opinion at the present time the best compromise between reliability, feasibility and
266 cost compared to NGS, in almost every laboratory.

267 **Funding**

268 **VITALE study** (NCT02292160) was funded by the French Ministry of Health (Programme
269 Hospitalier de Recherche Clinique) We are grateful to *Normandy Region* and *Rouen University*
270 for funding in part the cursus of Nicolas Vautrin.

271 **Acknowledgments**

272 We are indebted to the volunteers. We thank all the collaborators and colleagues who helped to
273 the study.

274 **References**

- 275 1. Terlizzi ME, Gribaudo G, Maffei ME. 2017. UroPathogenic *Escherichia coli* (UPEC)
276 Infections: Virulence Factors, Bladder Responses, Antibiotic, and Non-antibiotic
277 Antimicrobial Strategies. *Front Microbiol* 8:1566.
- 278 2. Murray BO, Flores C, Williams C, Flusberg DA, Marr EE, Kwiatkowska KM, Charest JL,
279 Isenberg BC, Rohn JL. 2021. Recurrent Urinary Tract Infection: A Mystery in Search of
280 Better Model Systems. *Front Cell Infect Microbiol* 11:691210.
- 281 3. Albert X, Huertas I, Pereiró II, Sanfélix J, Gosalbes V, Perrota C. 2004. Antibiotics for
282 preventing recurrent urinary tract infection in non-pregnant women. *Cochrane Database*
283 *Syst Rev* CD001209.
- 284 4. Eriksson I, Gustafson Y, Fagerström L, Olofsson B. 2010. Do urinary tract infections
285 affect morale among very old women? *Health Qual Life Outcomes* 8:73.
- 286 5. Nemoy LL, Kotetishvili M, Tigno J, Keefer-Norris A, Harris AD, Perencevich EN,
287 Johnson JA, Torpey D, Sulakvelidze A, Morris JG, Stine OC. 2005. Multilocus Sequence
288 Typing versus Pulsed-Field Gel Electrophoresis for Characterization of Extended-
289 Spectrum Beta-Lactamase-Producing *Escherichia coli* Isolates. *J Clin Microbiol* 43:1776–
290 1781.
- 291 6. Ejrnaes K, Sandvang D, Lundgren B, Ferry S, Holm S, Monsen T, Lundholm R, Frimodt-
292 Moller N. 2006. Pulsed-Field Gel Electrophoresis Typing of *Escherichia coli* Strains from
293 Samples Collected before and after Pivmecillinam or Placebo Treatment of
294 Uncomplicated Community-Acquired Urinary Tract Infection in Women. *J Clin*
295 *Microbiol* 44:1776–1781.

- 296 7. Skjøt-Rasmussen L, Hammerum AM, Jakobsen L, Lester CH, Larsen P, Frimodt-Møller
297 N. 2011. Persisting clones of *Escherichia coli* isolates from recurrent urinary tract
298 infection in men and women. *Journal of Medical Microbiology* 60:550–554.
- 299 8. Beerepoot MAJ, den Heijer CDJ, Penders J, Prins JM, Stobberingh EE, Geerlings SE.
300 2012. Predictive value of *Escherichia coli* susceptibility in strains causing asymptomatic
301 bacteriuria for women with recurrent symptomatic urinary tract infections receiving
302 prophylaxis. *Clinical Microbiology and Infection* 18:E84–E90.
- 303 9. Sanchez GV, Babiker A, Master RN, Luu T, Mathur A, Bordon J. 2016. Antibiotic
304 Resistance among Urinary Isolates from Female Outpatients in the United States in 2003
305 and 2012. *Antimicrob Agents Chemother* 60:2680–2683.
- 306 10. Caron F, Galperine T, Flateau C, Azria R, Bonacorsi S, Bruyère F, Cariou G, Clouqueur
307 E, Cohen R, Doco-Lecompte T, Elefant E, Faure K, Gauzit R, Gavazzi G, Lemaitre L,
308 Raymond J, Senneville E, Sotto A, Subtil D, Trivalle C, Merens A, Etienne M. 2018.
309 Practice guidelines for the management of adult community-acquired urinary tract
310 infections. *Médecine et Maladies Infectieuses* 48:327–358.
- 311 11. Clermont O, Christenson JK, Denamur E, Gordon DM. 2013. The Clermont *Escherichia*
312 *coli* phylo-typing method revisited: improvement of specificity and detection of new
313 phylo-groups. *Environ Microbiol Rep* 5:58–65.
- 314 12. Clermont O, Bonacorsi S, Bingen E. 2000. Rapid and simple determination of the
315 *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* 66:4555–4558.
- 316 13. Clermont O, Bonacorsi S, Bingen E. 2004. Characterization of an anonymous molecular
317 marker strongly linked to *Escherichia coli* strains causing neonatal meningitis. *J Clin*
318 *Microbiol* 42:1770–1772.

- 319 14. Clermont O, Lescat M, O'Brien CL, Gordon DM, Tenailon O, Denamur E. 2008.
320 Evidence for a human-specific *Escherichia coli* clone. *Environ Microbiol* 10:1000–1006.
- 321 15. Lescat M, Clermont O, Woerther PL, Glodt J, Dion S, Skurnik D, Djossou F, Dupont C,
322 Perroz G, Picard B, Catzeflis F, Andremont A, Denamur E. 2013. Commensal *Escherichia*
323 *coli* strains in Guiana reveal a high genetic diversity with host-dependant population
324 structure. *Environ Microbiol Rep* 5:49–57.
- 325 16. Weissman SJ, Johnson JR, Tchesnokova V, Billig M, Dykhuizen D, Riddell K, Rogers P,
326 Qin X, Butler-Wu S, Cookson BT, Fang FC, Scholes D, Chattopadhyay S, Sokurenko E.
327 2012. High-resolution two-locus clonal typing of extraintestinal pathogenic *Escherichia*
328 *coli*. *Appl Environ Microbiol* 78:1353–1360.
- 329 17. Tartof SY, Solberg OD, Manges AR, Riley LW. 2005. Analysis of a Uropathogenic
330 *Escherichia coli* Clonal Group by Multilocus Sequence Typing. *J Clin Microbiol*
331 43:5860–5864.
- 332 18. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden
333 MCJ, Ochman H, Achtman M. 2006. Sex and virulence in *Escherichia coli*: an
334 evolutionary perspective. *Mol Microbiol* 60:1136–1151.
- 335 19. Roer L, Johannesen TB, Hansen F, Stegger M, Tchesnokova V, Sokurenko E, Garibay N,
336 Allesøe R, Thomsen MCF, Lund O, Hasman H, Hammerum AM. 2018. CHTyper, a Web
337 Tool for Subtyping of Extraintestinal Pathogenic *Escherichia coli* Based on the *fumC* and
338 *fimH* Alleles. *J Clin Microbiol* 56:e00063-18.
- 339 20. Hunter PR, Gaston MA. 1988. Numerical index of the discriminatory ability of typing
340 systems: an application of Simpson's index of diversity. *J Clin Microbiol* 26:2465–2466.

- 341 21. Grundmann H, Hori S, Tanner G. 2001. Determining Confidence Intervals When
342 Measuring Genetic Diversity and the Discriminatory Abilities of Typing Methods for
343 Microorganisms. *J Clin Microbiol* 39:4190–4192.
- 344 22. Petronella N, Kundra P, Auclair O, Hébert K, Rao M, Kingsley K, De Bruyne K, Banerjee
345 S, Gill A, Pagotto F, Tamber S, Ronholm J. 2019. Changes detected in the genome
346 sequences of *Escherichia coli*, *Listeria monocytogenes*, *Vibrio parahaemolyticus*, and
347 *Salmonella enterica* after serial subculturing. *Can J Microbiol* 65:842–850.
- 348 23. Lee J h., Subhadra B, Son Y-J, Kim D h., Park H s., Kim J m., Koo S h., Oh M h., Kim
349 H-J, Choi C h. 2016. Phylogenetic group distributions, virulence factors and antimicrobial
350 resistance properties of uropathogenic *Escherichia coli* strains isolated from patients with
351 urinary tract infections in South Korea. *Letters in Applied Microbiology* 62:84–90.
- 352 24. Hisano M, Bruschini H, Nicodemo AC, Gomes CM, Lucon M, Srougi M. 2015. The
353 Bacterial Spectrum and Antimicrobial Susceptibility in Female Recurrent Urinary Tract
354 Infection: How Different They Are From Sporadic Single Episodes? *Urology* 86:492–497.
- 355 25. Warren JW. 1996. Clinical presentations and epidemiology of urinary tract infections, p.
356 3–27. *In* Urinary tract infections: molecular pathogenesis and clinical management. ASM
357 Press. Washington, DC.
- 358 26. Amna MA, Chazan B, Raz R, Edelstein H, Colodner R. 2013. Risk factors for non-
359 *Escherichia coli* community-acquired bacteriuria. *Infection* 41:473–477.
- 360 27. Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. 2015. Urinary tract infections:
361 epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*
362 13:269–284.

- 363 28. Forde BM, Roberts LW, Phan M-D, Peters KM, Fleming BA, Russell CW, Lenherr SM,
364 Myers JB, Barker AP, Fisher MA, Chong T-M, Yin W-F, Chan K-G, Schembri MA,
365 Mulvey MA, Beatson SA. 2019. Population dynamics of an *Escherichia coli* ST131
366 lineage during recurrent urinary tract infection. *Nat Commun* 10:3643.
- 367 29. Russo TA, Stapleton A, Wenderoth S, Hooton TM, Stamm WE. 1995. Chromosomal
368 restriction fragment length polymorphism analysis of *Escherichia coli* strains causing
369 recurrent urinary tract infections in young women. *J Infect Dis* 172:440–445.
- 370 30. Vosti KL. 2007. A prospective, longitudinal study of the behavior of serologically
371 classified isolates of *Escherichia coli* in women with recurrent urinary tract infections.
372 *Journal of Infection* 55:8–18.
- 373 31. Nielsen KL, Stegger M, Kiil K, Lilje B, Ejrnæs K, Leihof RF, Skjøt-Rasmussen L,
374 Godfrey P, Monsen T, Ferry S, Hammerum AM, Frimodt-Møller N. 2021. *Escherichia*
375 *coli* Causing Recurrent Urinary Tract Infections: Comparison to Non-Recurrent Isolates
376 and Genomic Adaptation in Recurrent Infections. *Microorganisms* 9:1416.

377 **Figure legends**

378 **Figure 1: Venn diagram illustrating the repartition of the CH types for 347 UPEC from**
379 **sporadic cystitis or recurrent cystitis.**

380 **Figure 2: CH-typing based phylogeny of 347 *E. coli* isolates represented as minimum**
381 **spanning trees according to clinical presentation (A) or phylogroup (B).** Each circle
382 represents a CH type and each division represent an isolate. Length of strings represent the
383 genetic distance between CH types. Major CH-types were indicated on the tree as *fimC* allele
384 number_ *fimH* allele number. SC = sporadic cystitis, RC = recurrent cystitis

385 **Figure 3: History of the 22 patients with relapse(s).** Each line represents one patient, circles
386 are for UPEC isolates, squares for *E. faecalis* and diamond for *K. pneumoniae*. Colors represents
387 identical UPEC isolates for a given patient. Numbers in circles are CH types of the isolates.

Table 1: Bacterial species identified for the 445 isolates collected from the 323 women with recurrent cystitis (RC) or sporadic cystitis (SC). Significant differences ($p < 0.05$) between groups are presented in bold.

Bacterial species n (%)	All isolates ($n = 445$)	SC isolates ($n = 265$)	RC isolates ($n = 180$)
<i>Escherichia coli</i>	347 (78%)	216 (81.5%)	131 (72.8%)
A	47 (13.5%)	26 (12%)	21 (16%)
B1	34 (9.8%)	18 (8.3%)	16 (12.2%)
B2	182 (52.4%)	119 (55.1%)	63 (48.1%)
C	16 (4.6%)	6 (2.8%)	10 (7.6%)
D	49 (14.1%)	40 (18.5%)	9 (6.9%)
E	3 (0.9%)	1 (0.5%)	2 (1.5%)
F	16 (4.6%)	6 (2.8%)	10 (7.6%)
<i>Klebsiella pneumoniae</i>	19 (4.3%)	6 (2.3%)	13 (7.2%)
<i>Proteus mirabilis</i>	14 (3.1%)	6 (2.3%)	8 (4.4%)
<i>Citrobacter koseri</i>	9 (2%)	7 (2.6%)	2 (1.1%)
Other <i>Enterobacteriales</i> species	7 (1.6%)	5 (1.9%)	2 (1.1 %)
Non <i>Enterobacteriales</i> species	49 (11%)	25 (9.4%)	24 (13.3 %)

Table 2: Rates of resistance to different antibiotics among the 347 *E. coli* isolates collected from patients with sporadic cystitis (SC) or recurrent cystitis (RC). Significant differences ($p < 0.05$) between groups are presented in bold.

Antibiotic resistance n (%)	All isolates (n = 347)	SC isolates (n = 216)	RC isolates (n = 131)
Amoxicillin	170/347 (49%)	93/216 (43.1%)	77/131 (58.8%)
Amoxicillin + Clavulanic acid	14/345 (4.1%)	6/214 (2.8%)	8/131 (6.1%)
Pivmecillinam	10/334 (3%)	5/211 (2.4%)	5/123 (4%)
Ceftriaxone	11/344 (3.2%)	3/216 (1.4%)	8/128 (6.2%)
Nalidixic Acid	56/347 (16.1%)	21/216 (9.7%)	35/131 (26.7%)
Ciprofloxacin	40/347 (11.5%)	16/216 (7.4%)	24/131 (18.3%)
Trimethoprim + sulfamethoxazole	78/343 (22.7%)	30/214 (14%)	48/129 (37.2%)
Nitrofurantoin	4/347 (1.2%)	2/347 (0.9%)	2/131 (1.5%)
Fosfomycin	3/342 (0.9%)	1/214 (0.5%)	2/128 (1.6%)

Table 3: Descriptive table of recurrent cystitis group population. UPEC: Uropathogenic *Escherichia coli*, NGS: Next Generation Sequencing.

	Recurrent cystitis			
	Total patients	Total isolates	Total UPEC	Identical UPEC isolates
Overall	72	180	131	77
No case with UPEC	14	28	0	0
1 case only with UPEC	18	36	18	0
≥ 2 cases with UPEC:				
all different by phylogroup / CH type	18	37	36	0
some identical by phylogroup / CH type	22	79	77	60
some identical by NGS	22	79	77	58

Table 4: Intra-patient pairwise comparison between recurrent cystitis isolates in terms of antibiotic susceptibility profile (ASP) and CH typing.

		Results of CH typing	
		Identical	different
Results of ASP	Identical pairs	33	25
	(WT isolates pairs)	(15)	(10)
	Different pairs	16	71

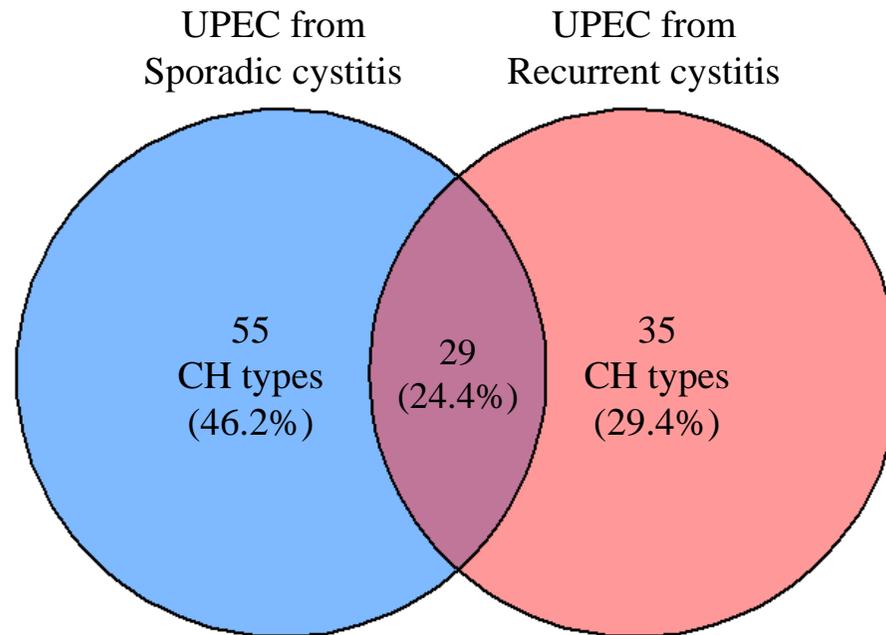


Figure 1 : Venn diagram illustrating the repartition of the CH types for 347 UPEC from sporadic cystitis or recurrent cystitis.

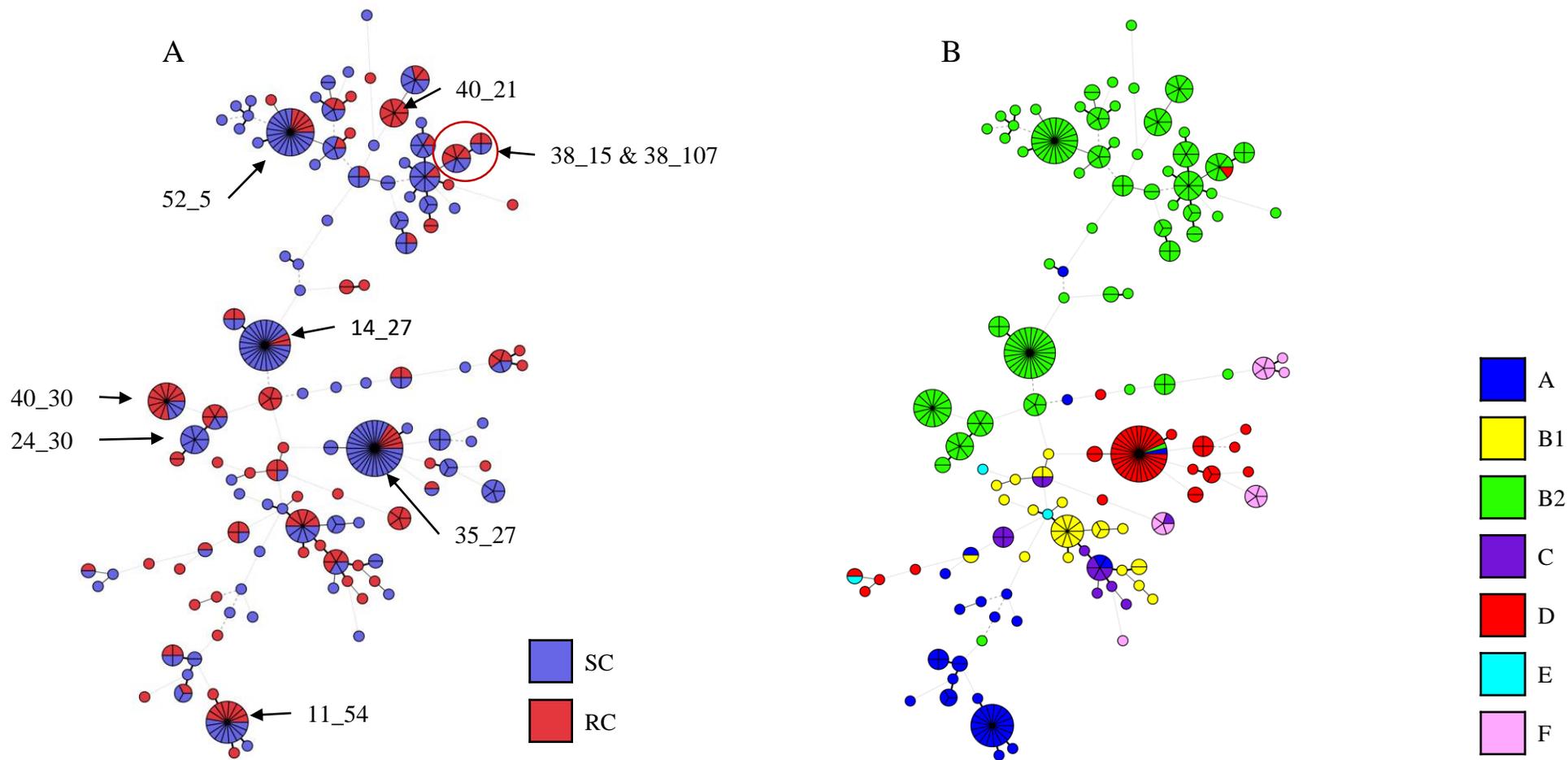


Figure 2: CH-typing based phylogeny of 347 *E. coli* isolates represented as minimum spanning trees according to clinical presentation (A) or phylogroup (B). Each circle represents a CH type and each division represent an isolate. Length of strings represent the genetic distance between CH types. Major CH-types were indicated on the tree as *fumC* allele number_*fimH* allele number. SC = sporadic cystitis, RC = recurrent cystitis

Table S1: List of primers used for phylotyping and CH typing. bp = base pair

PCR reaction	Primer ID	Target	Primer sequences	Amplicon size (bp)	Reference
Quadruplex	chuA.1b	<i>chuA</i>	5'-ATGGTACCGGACGAACCAAC-3'	288	Clermont <i>et al.</i> 2013 (11)
	chuA.2b		5'-TGCCGCCAGTACCAAAGACA-3'		Clermont <i>et al.</i> 2000 (12)
	yjaA.1b	<i>yjaA</i>	5'-CAAACGTGAAGTGTCAGGAG-3'	211	Clermont <i>et al.</i> 2013 (11)
	yjaA.2b		5'-AATGCGTTCCTCAACCTGTG-3'		Clermont <i>et al.</i> 2013 (11)
	TspE4C2.1b	TspE4.C2	5'-CACTATTCGTAAGGTCATCC-3'	152	Clermont <i>et al.</i> 2013 (11)
	TspE4C2.2b		5'-AGTTTATCGCTGCGGGTTCG-3'		Clermont <i>et al.</i> 2013 (11)
	AceK.f	<i>arpA</i>	5'-AACGCTATTCGCCAGCTTGC-3'	400	Clermont <i>et al.</i> 2013 (11)
	ArpA1.r		5'-TCTCCCCATACCGTACGCTA-3'		Clermont <i>et al.</i> 2004 (13)
Group E	ArpAgpeE.f	<i>arpA</i>	5'-GATTCCATCTTGTCAAAATATGCC-3'	301	Lescat <i>et al.</i> 2012 (15)
	ArpAgpeE.r		5'-GAAAAGAAAAAGAATTCCCAAGAG-3'		Lescat <i>et al.</i> 2012 (15)
Group C	ArpAgpeC.f	<i>arpA</i>	5'-AGTTTTATGCCAGTGCGAG-3'	219	Lescat <i>et al.</i> 2012 (15)
	ArpAgpeC.r		5'-TCTGCGCCGGTTCACGCC-3'		Lescat <i>et al.</i> 2012 (15)
Internal control	trpBA.f	<i>trpA</i>	5'-CGGCGATAAAGACATCTTCAC-3'	489	Clermont <i>et al.</i> 2008 (14)
	trpBA.r		5'-GCAACGCGGCCTGGCGGAAG-3'		Clermont <i>et al.</i> 2008 (14)
fimH typing	fimH.f	<i>fimH</i>	5'-CACTCAGGGAACCATTTCAGGCA-3'	975	Weissman <i>et al.</i> 2012 (16)
	fimH.r		5'-CTTATTGATAAACAAGATCAC-3'		Weissman <i>et al.</i> 2012 (16)
fumC typing	fumC.f	<i>fumC</i>	5'-TCACAGGTCGCCAGCGCTTC-3'	806	Wirth <i>et al.</i> 2006 (18)
	fumC.r		5'-GTACGCAGCGAAAAAGATTC-3'		Wirth <i>et al.</i> 2006 (18)

Table S2: PCRs mix compositions and reactions conditions

PCR Reaction	GoTaq® volume	Target gene	Primers amount	Extracted DNA volume	Final volume	PCR conditions
Quadruplex*	10µL	All	50pmol	2µL	20µL	Denaturation 4min at 94°C, 30 cycles (5s at 94°C and 20s at 59°C), final extension step of 5min at 72°C
Group C*	10µL	<i>arpA</i> <i>trpA</i>	20pmol 12.5pmol	2µL	20µL	Same conditions than Quadruplex
Group E*	10µL	<i>arpA</i> <i>trpA</i>	20pmol 12.5pmol	2µL	20µL	Denaturation 4min at 94°C, 30 cycles (5s at 94°C and 20s at 57°C), final extension step of 5min at 72°C
<i>fimH</i> typing	12,5µL	<i>fimH</i>	25pmol	5µL	25µL	Denaturation 2min at 95°C, 30 cycles (30s at 94°C, 30s at 57°C and 1min at 72°C), final extension step of 5min at 72°C
<i>fumC</i> typing	12,5µL	<i>fumC</i>	12,5pmol	5µL	25µL	Denaturation 2min at 95°C, 30 cycles (1min at 95°C, 1min at 52°C, 2min at 72°C), final extension step of 5min at 72°C

*PCR products were visualized by an electrophoresis performed in a 2% agarose gel with 0.005µg.mL⁻¹ of ethidium bromide. Migrations were performed at 100V during 90min.

Table S3: CH-types of the 347 UPEC isolates from sporadic cystitis (SC) or recurrent cystitis (RC). CH-types are listed by decreasing frequency. CH types identified in both clinical groups (*i.e.* SC and RC) are presented in bold.

CH type n (%)	All UPEC (n = 347)	UPEC from SC (n=216)	UPEC from RC (n=131)
35_27	29 (8%)	24 (11%)	5 (4%)
14_27	24 (7%)	22 (10%)	2 (2%)
52_5	21 (6%)	16 (7%)	5 (4%)
11_54	17 (5%)	9 (4%)	8 (6%)
40_30	13 (4%)	3 (1%)	10 (8%)
4_32	10 (3%)	5 (2%)	5 (4%)
24_9	8 (2%)	7 (3%)	1 (1%)
38_15	7 (2%)	3 (1%)	4 (3%)
40_21	7 (2%)	0 (0%)	7 (5%)
24_10	6 (2%)	5 (2%)	1 (1%)
24_30	6 (2%)	6 (3%)	0 (0%)
38_30	6 (2%)	2 (1%)	4 (3%)
4_39	6 (2%)	2 (1%)	4 (3%)
40_22	6 (2%)	4 (2%)	2 (2%)
13_5	5 (1%)	3 (1%)	2 (2%)
14_5	5 (1%)	4 (2%)	1 (1%)
231_58	5 (1%)	0 (0%)	5 (4%)
4_0	5 (1%)	2 (1%)	3 (2%)
45_97	5 (1%)	2 (1%)	3 (2%)
11_0	4 (1%)	3 (1%)	1 (1%)
11_27	4 (1%)	2 (1%)	2 (2%)
14_64	4 (1%)	2 (1%)	2 (2%)
24_120	4 (1%)	3 (1%)	1 (1%)
32_41	4 (1%)	4 (2%)	0 (0%)
35_47	4 (1%)	4 (2%)	0 (0%)
38_107	4 (1%)	2 (1%)	2 (2%)
38_5	4 (1%)	3 (1%)	1 (1%)
4_27	4 (1%)	1 (0%)	3 (2%)
4_609	4 (1%)	1 (0%)	3 (2%)
40_41	4 (1%)	2 (1%)	2 (2%)
100_96	3 (1%)	3 (1%)	0 (0%)
11_24	3 (1%)	2 (1%)	1 (1%)
24_1	3 (1%)	3 (1%)	0 (0%)
24_103	3 (1%)	3 (1%)	0 (0%)
65_32	3 (1%)	3 (1%)	0 (0%)
7_54	3 (1%)	1 (0%)	2 (2%)

11_25	2 (1%)	2 (1%)	0 (0%)
13_106	2 (1%)	2 (1%)	0 (0%)
24_2	2 (1%)	2 (1%)	0 (0%)
24_214	2 (1%)	0 (0%)	2 (2%)
24_239	2 (1%)	0 (0%)	2 (2%)
26_27	2 (1%)	2 (1%)	0 (0%)
35_30	2 (1%)	1 (0%)	1 (1%)
36_48	2 (1%)	1 (0%)	1 (1%)
38_27	2 (1%)	0 (0%)	2 (2%)
4_121	2 (1%)	2 (1%)	0 (0%)
43_197	2 (1%)	0 (0%)	2 (2%)
100_611	1 (<1%)	0 (0%)	1 (1%)
103_9	1 (<1%)	1 (<1%)	0 (0%)
108_0	1 (<1%)	1 (<1%)	0 (0%)
11_?	1 (<1%)	1 (<1%)	0 (0%)
11_215	1 (<1%)	1 (<1%)	0 (0%)
11_23	1 (<1%)	1 (<1%)	0 (0%)
11_30	1 (<1%)	1 (<1%)	0 (0%)
11_34	1 (<1%)	0 (0%)	1 (1%)
11_41	1 (<1%)	0 (0%)	1 (1%)
11_53	1 (<1%)	0 (0%)	1 (1%)
11_566	1 (<1%)	1 (<1%)	0 (0%)
11_608	1 (<1%)	0 (0%)	1 (1%)
11_613	1 (<1%)	0 (0%)	1 (1%)
11_86	1 (<1%)	1 (<1%)	0 (0%)
1154_54	1 (<1%)	0 (0%)	1 (1%)
13_201	1 (<1%)	0 (0%)	1 (1%)
13_202	1 (<1%)	1 (<1%)	0 (0%)
14_144	1 (<1%)	0 (0%)	1 (1%)
14_2	1 (<1%)	1 (<1%)	0 (0%)
14_54	1 (<1%)	1 (<1%)	0 (0%)
182_168	1 (<1%)	0 (0%)	1 (1%)
19_27	1 (<1%)	0 (0%)	1 (1%)
22_21	1 (<1%)	0 (0%)	1 (1%)
23_38	1 (<1%)	0 (0%)	1 (1%)
23_86	1 (<1%)	1 (<1%)	0 (0%)
24_102	1 (<1%)	1 (<1%)	0 (0%)
24_118	1 (<1%)	1 (<1%)	0 (0%)
24_24	1 (<1%)	1 (<1%)	0 (0%)
24_608	1 (<1%)	0 (0%)	1 (1%)
24_610	1 (<1%)	1 (<1%)	0 (0%)
24_615	1 (<1%)	1 (<1%)	0 (0%)
24_96	1 (<1%)	1 (<1%)	0 (0%)
249_274	1 (<1%)	1 (<1%)	0 (0%)

26_172	1 (<1%)	1 (<1%)	0 (0%)
26_65	1 (<1%)	0 (0%)	1 (1%)
319_197	1 (<1%)	0 (0%)	1 (1%)
35_31	1 (<1%)	1 (<1%)	0 (0%)
35_612	1 (<1%)	1 (<1%)	0 (0%)
36_54	1 (<1%)	0 (0%)	1 (1%)
36_617	1 (<1%)	1 (<1%)	0 (0%)
36_93	1 (<1%)	1 (<1%)	0 (0%)
37_27	1 (<1%)	0 (0%)	1 (1%)
372_10	1 (<1%)	1 (<1%)	0 (0%)
38_21	1 (<1%)	1 (<1%)	0 (0%)
38_41	1 (<1%)	1 (<1%)	0 (0%)
38_54	1 (<1%)	1 (<1%)	0 (0%)
38_92	1 (<1%)	1 (<1%)	0 (0%)
39_3	1 (<1%)	0 (0%)	1 (1%)
4_24	1 (<1%)	0 (0%)	1 (1%)
4_303	1 (<1%)	0 (0%)	1 (1%)
4_31	1 (<1%)	1 (<1%)	0 (0%)
4_34	1 (<1%)	0 (0%)	1 (1%)
4_35	1 (<1%)	0 (0%)	1 (1%)
4_366	1 (<1%)	0 (0%)	1 (1%)
4_38	1 (<1%)	0 (0%)	1 (1%)
4_43	1 (<1%)	0 (0%)	1 (1%)
41_30	1 (<1%)	1 (<1%)	0 (0%)
43_616	1 (<1%)	1 (<1%)	0 (0%)
45_151	1 (<1%)	0 (0%)	1 (1%)
45_189	1 (<1%)	0 (0%)	1 (1%)
52_14	1 (<1%)	1 (<1%)	0 (0%)
52_186	1 (<1%)	1 (<1%)	0 (0%)
52_414	1 (<1%)	1 (<1%)	0 (0%)
52_614	1 (<1%)	1 (<1%)	0 (0%)
52_75	1 (<1%)	1 (<1%)	0 (0%)
52_76	1 (<1%)	1 (<1%)	0 (0%)
58_24	1 (<1%)	0 (0%)	1 (1%)
6_23	1 (<1%)	1 (<1%)	0 (0%)
6_31	1 (<1%)	1 (<1%)	0 (0%)
67_222	1 (<1%)	1 (<1%)	0 (0%)
8_31	1 (<1%)	1 (<1%)	0 (0%)
95_32	1 (<1%)	1 (<1%)	0 (0%)

I.3. Principaux résultats

I. 3. 1. Description de la cohorte VITALE

Au cours des 17 mois de l'étude, 347 isolats d'UPEC ont été isolés d'urines de 323 femmes atteintes de cystites communautaires. Parmi ces patientes, 72 ont été identifiées comme souffrant de cystites récurrentes (CR) (*i.e.* ont présenté au moins deux épisodes de cystite sur une période de six mois consécutifs). Les 251 autres, ont été classées dans le groupe des patientes atteintes de cystites sporadiques (CS). Il est à noter que l'âge moyen des patientes du groupe CR était significativement plus élevé que celui du groupe CS (58,6 vs 54,3 ans, $p < 0,001$). Par ailleurs, la prévalence de cystites à risque de complications était également significativement plus élevée dans le groupe CR (65,3% vs 29,5%, $p = 0,039$).

I. 3. 2. Diversité génétique des isolats d'UPEC

Une grande diversité génétique (mesurée par l'indice de diversité de Simpson [DI]) au sein des UPEC a été observée. Parmi 347 isolats étudiés, 119 génotypes (nommés *CH types*) ont en effet été identifiés (DI = 0,976). Cinq d'entre eux ([35_27], [14_27], [52_5], [11_54] et [40_30]) ont été considérés comme majeurs car représentant 30% des isolats ($n = 104$). À l'inverse, 72 *CH types* ne représentaient qu'un seul isolat. Cette grande diversité génétique a été observée aussi bien au sein du groupe des patientes atteintes de CS (DI = 0,966), que dans celui des patientes atteintes de CR (DI = 0,979). Enfin, la phylogénie construite sur la base des données de *CH typing* n'a pas mis en évidence de lien entre génotype et caractère récurrent.

I. 3. 3. Prévalence des rechutes – apport du *CH typing* et de l'antibiotype dans l'identification des rechutes

Parmi les 72 patientes du groupe CR, 22 (30,6%) ont présenté au moins 2 épisodes de cystites causés par des isolats d'UPEC assignés au même phylogroupe, et présentant le même *CH type*. L'analyse par NGS (*wgSNP*) des 35 paires d'isolats non-distingués par le *CH typing* a confirmé l'identité de toutes les paires d'isolats sauf une, montrant que dans la quasi-totalité des cas, le *CH typing* était suffisant pour identifier les souches responsables de rechutes. Il est intéressant de noter que la simple comparaison des antibiotypes a permis de prédire la clonalité des souches dans 71,7% des cas. Enfin, de nombreuses dynamiques de récurrence ont été observées chez ces 22 patientes. Sur la période de l'analyse, 7 des 22 patientes n'ont fait des rechutes qu'à la même souche, tandis que les 15 restantes ont présenté des réinfections avec une nouvelle souche d'UPEC, entrecoupées ou non par une rechute avec un clone précédemment identifié.

Ainsi, cette première étude, a mis en évidence une grande diversité moléculaire des UPEC, indépendamment de leur caractère récurrent ou non. Les rechutes occupaient une place importante

dans le phénomène de récurrence puisqu'elles concernaient, d'après les données NGS, environ un tiers des patientes atteintes de CR. Cependant, la récurrence des cystites à UPEC est apparue clairement comme un phénomène complexe alternant chez deux tiers des patientes de la cohorte, des rechutes et des réinfections. Pour l'analyse des épisodes successifs de cystite à UPEC d'une patiente donnée, la comparaison des antibiogrammes permettait de prédire à dans 7 cas sur 10 les récurrences au même clone, une performance bien supérieure à celle jadis décrite à « l'ère pré-multirésistance ». Le *CH typing* offrait d'excellentes performances, à peine amplifiées par le NGS.

II. Second travail personnel

II.1. Contextualisation de l'étude

Ainsi, notre premier travail a identifié dans une cohorte prospective de 72 patientes, 22 femmes atteintes de CR. Le suivi longitudinal de ces patientes ayant finalement permis de distinguer – dans les limites de la période de l'étude – 24 souches d'UPEC à l'origine de rechutes et 22 de réinfection.

Très logiquement s'est posée alors la question des éventuelles différences phénotypiques ou génotypiques de ces souches d'UPEC qui pourraient rendre compte d'une physiopathologie distincte des épisodes pour des raisons purement microbiennes (les facteurs liés à l'hôte n'étant pas abordés en cette série).

D'où la mise en œuvre de différentes investigations sur notre soucier. Ainsi, les génomes de l'ensemble des souches d'UPEC responsables de rechutes ont été séquencés puis comparés avec ceux de souches responsables de CS. Ces deux groupes de souches ont également été comparés sur certains critères phénotypiques décrits comme importants dans le phénomène de récurrence (fitness et formation de biofilm en milieu mimant l'urine, internalisation dans des cellules urothéliales).

En parallèle, les isolats séquentiels d'UPEC provenant des patientes récidivantes ont été comparés par *whole genome SNP* ainsi que sur les marqueurs phénotypiques précédemment décrits, afin d'étudier les mécanismes d'adaptation à l'arbre urinaire des souches responsables de rechutes.

Pour le volet génomique, le séquençage des isolats par Illumina ainsi que l'assemblage des génomes a été sous-traité à la plateforme P2M (Pasteur, Paris). Le séquençage des isolats par Nanopore (de l'extraction des génomes au séquençage à proprement parler), a été réalisé par mes soins avec l'aide de Marie Leoz, ingénieure de recherche au laboratoire et de Maxime Grand (ATER au sein du laboratoire). Les analyses bioinformatiques (assemblage hybride, annotation des génomes, recherche de SNP...) ont été entièrement réalisées par Marie Leoz ; cependant, j'ai analysé les données génomiques en étroite collaboration avec elle.

Pour le volet phénotypique, j'ai adapté et mis en place le milieu mimant l'urine avec l'aide de Maxime Grand (à partir d'une publication datant de 1998) ainsi que les protocoles des analyses phénotypiques utilisant ce milieu (Fitness, formation de biofilm). Les tests de fitness et de formation de biofilm ont été réalisés par Audrey Feldmann (étudiante en Master 2) et Maxime Grand, sous ma supervision.

J'ai également mis au point le modèle *in vitro* d'étude de la capacité d'internalisation des souches d'UPEC avec l'aide précieuse de David Ribet, chargé de recherche à l'UMR INSERM 1073 ADEN. L'ensemble des expérimentations utilisant ce modèle ont été réalisées par mes soins.

II.2. Publication n°2

Are Uropathogenic *Escherichia coli* relapsing strains just ordinary UPEC strains?

Nicolas Vautrin, Sandrine Dahyot, Marie Leoz, François Caron, Maxime Grand, Audrey Feldmann, François Gravey, Stéphanie Legris, Seydi Thimbo, David Ribet, Kévin Alexandre, Martine Pestel-Caron

Article en préparation pour une soumission dans la revue *PLOS Pathogens*

1 Are *Escherichia coli* relapsing strains just ordinary UPEC
2 strains?

3 Nicolas Vautrin^{1*}, Sandrine Dahyot², Marie Leoz¹, François Caron³, Maxime Grand¹, Audrey
4 Feldmann¹, François Gravey⁴, Stéphanie Legris¹, Seydi Thimbo¹, David Ribet⁵, Kévin Alexandre³,
5 Martine Pestel-Caron²

6

7 ¹ Normandie Univ, UNIROUEN, UNICAEN, Inserm UMR 1311 DYNAMICURE, 76000 Rouen,
8 France

9 ² Normandie Univ, UNIROUEN, UNICAEN, Inserm UMR 1311 DYNAMICURE, CHU Rouen,
10 Department of Microbiology, F-76000 Rouen, France

11 ³ Normandie Univ, UNIROUEN, UNICAEN, Inserm UMR 1311 DYNAMICURE, CHU Rouen,
12 Department of Infectious diseases, F-76000 Rouen, France

13 ⁴ Normandie Univ, UNICAEN, UNIROUEN, Inserm UMR 1311 DYNAMICURE, 14000 Caen, France

14 ⁵ Normandie Univ, UNIROUEN, Inserm UMR 1073 ADEN, 76000 Rouen, France

15

16 *Corresponding author

17 E-mail : vautrin.nicolas@outlook.com

18 **Abstract**

19 Recurrent cystitis is a public health concern associated with an economic, societal and personal burden.
20 The specific determinants associated with recurrence of Uropathogenic Escherichia coli (UPECs) are
21 still poorly characterized. The aims of this study were (i) to describe genomic and phenotypic traits
22 associated with recurrence, using a large collection of recurrent and sporadic UPEC isolates, and (ii) to
23 explore within-host genomic adaptation using sequential UPEC isolates. Whole genome comparative
24 analyses between 24 initial recurrent isolate (iRCIs) and 24 sporadic isolates (SCIs) brought evidence
25 for loss of putative mobile genetic elements, and for the under- and over-representation of 3 and 1
26 putative genes respectively in iRCIs. The intra-patient evolution of the 24 iRCIs over time confirmed
27 that plasmid loss occurred in the course of recurrence. SNP adaptation also seemed to occur early and
28 then stabilize; it especially occurred in Open Reading Frames involved in metabolism and ABC
29 membrane transporters. Fitness and biofilm formation in artificial urinary medium were expected to be
30 optimized in RCIs compared to SCIs, but phenotypic analyses showed no difference between them. We
31 also investigated their capacity to invade bladder epithelial cells, and found that only few isolates were
32 able to efficiently infect these cells. Unexpectedly, none of those were RCIs.

33 Thus, RCIs do not seem to exhibit specific genomic nor phenotypic traits, including urothelial
34 internalization. The mechanisms driving recurrence are probably more complex and may involve
35 multiple combinations of several parameters including gene expression as well as bacterial and host
36 factors.

37 **Introduction**

38 Urinary tract infections (UTIs) are very common bacterial infections in women, as more than half of
39 them will develop at least one UTI during their lifetime (1). Up to 25% of women who have had a cystitis
40 episode will experience a second within a year (2). If a woman presents more than 2 episodes within a
41 6-month period or 3 episodes within 12 months, she will be considered as suffering from recurrent
42 cystitis (3).

43 Recurrent cystitis is a public health concern and associated with an economic, societal and personal
44 burden. First, they represent between 1% and 6% of all medical visits in the United States and a cost of
45 1.6 billion US dollars annually. Moreover, recurrent cystitis impact negatively quality of life, as they
46 are associated with anxiety and depression (4).

47 Recurrent cystitis are mainly caused by uropathogenic *Escherichia coli* (UPEC) (2). Forty-seven to 81%
48 of recurrences are caused by the primary infecting UPEC strain (3). However, this statement mainly
49 relies on typing methods such as serotyping and pulsed-field gel electrophoresis. Furthermore, there are
50 only few studies that explored genomic and phenotypic features of sequential relapse UPEC isolates
51 (5,6).

52 UPEC virulence mainly relies on their ability to survive and grow in urine, as well as to adhere and
53 invade urothelial cells (7). Despite accumulated knowledge about UPEC virulence, the physiopathology
54 of recurrent cystitis remains poorly characterized (8). Among the existing hypotheses, it has been
55 postulated that the ability of UPECs to persist in the bladder by forming intracellular bacterial
56 communities (IBC) might be involved in recurrent cystitis (9–11). However, this phenomenon hasn't
57 yet been explored in large patient cohorts. Moreover, no specific genomic or phenotypic markers have
58 been identified to distinguish recurrent cystitis associated UPECs from those causing sporadic cystitis
59 (5,12).

60 In this context, our study first aimed to identify genomic and/or phenotypic determinants associated with
61 recurrence using a large collection of 24 recurrent and sporadic cystitis isolates (RCIs and SCIs
62 respectively) sampled from patients with well-characterized cystitis (VITALE study, NTC02292160).
63 A second aim was to describe the within-host microevolution of the 24 RCIs to adapt to the urinary tract,
64 by analyzing the genomic changes of sequential isolates sampled from the patients with recurrent
65 cystitis.

66 **Results**

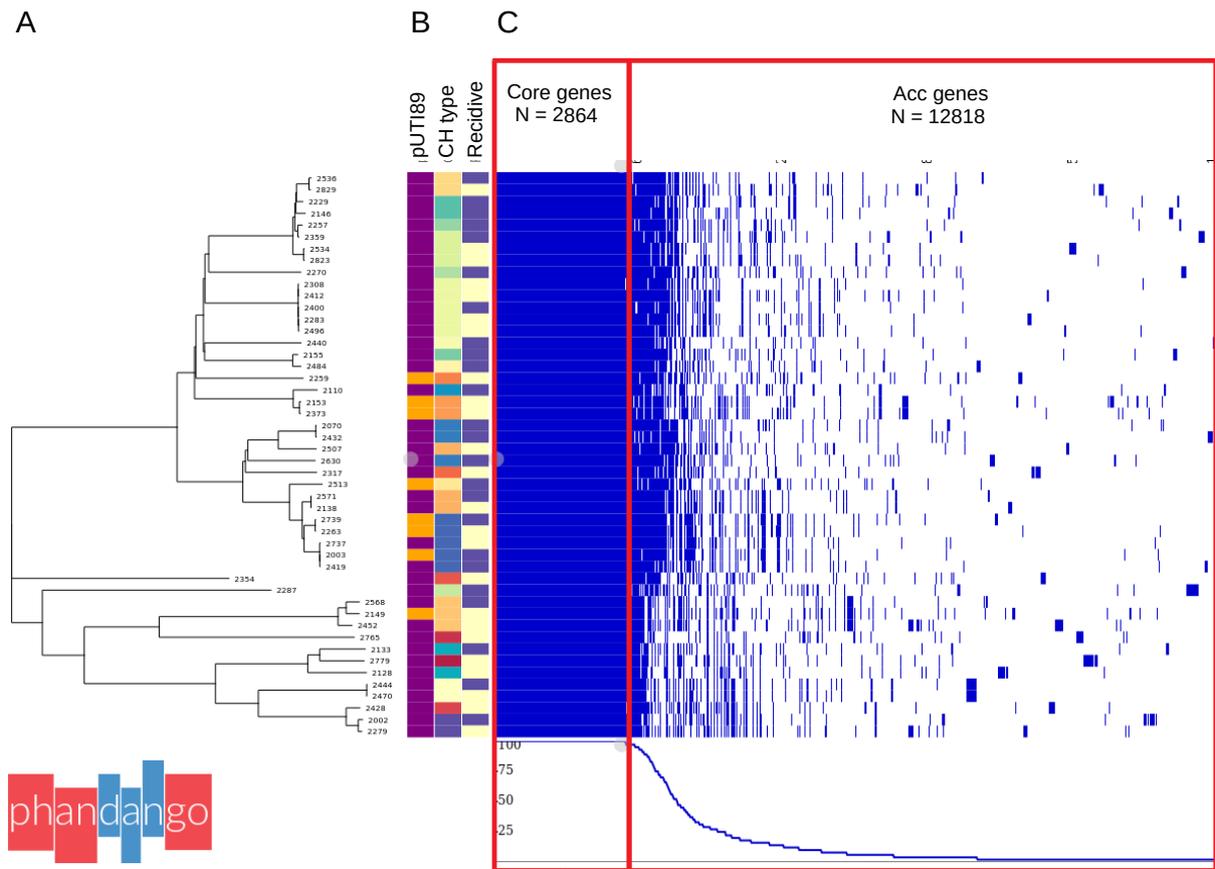
67 **Genes of interest do not distinguish RCIs from SCIs at the** 68 **interpatient scale**

69 We previously used short read whole genome sequencing (WGS) to identify 24 UPECs causing cystitis
70 relapses (13). To search for genomic traits associated with recurrence, we compared the 24 draft genome
71 assemblies of the initial RCI (iRCI) isolates with those of 24 SCIs, phylogenetically matched on the
72 basis of CH typing. The presence and protein sequences of several genes of interest were compared
73 between these two groups: adhesins (including *fim*, *csg*, *yad* and *pap* operons), toxins (*cnf*, *hly* and *sat*)
74 and nutrition genes (*chuA*, *fyuA* and *iroN*). No significant difference was observed between RCIs and
75 SCIs. We also looked for pUTI89, a plasmid that is associated with urinary pathogenicity, and found it
76 in 3 and 6 of the iRCIs and SCIs, respectively (Fig.1). The difference was not significant ($p=0.4$).

77 **Initial RCI genomes are smaller than phylogenetically paired SCIs'**

78 We observed that the mean genome size of iRCIs was significantly lower than that of SCIs (5.01Mb vs
79 5.11Mb respectively, $p < 0.05$). This was confirmed by a significantly lower mean protein count (4,658
80 vs 4,776 in iRCIs and SCIs, respectively, $p < 0.05$). A Genome Wide Association Study (GWAS) was
81 thus conducted to investigate whether some genes were specifically present in SCIs and not in iRCIs.

82 The pan genome of the 48 assemblies studied included a total of 15,682 genes, 2,864 (18.3%) of which
83 were considered as core genes. Phylogenetic analysis of the core genes confirmed that the isolates
84 clustered based on the CH type rather than on the SCI/iRCI phenotype (**Fig 1**).



85 **Fig 1. Core genome analysis of 24 iRCI and 24 phylogenetically paired SCIs. A.** Neighbor-Joining
 86 phylogenetic tree based on the 48 strains' translated core gene alignment (2,755,382 amino acid
 87 positions). **B.** Color-coded presence of pUIT89 (purple = absent, orange = present), CH Type and
 88 recidive phenotype (purple = iRCI, yellow = SCI) of the 48 strains. **C.** Presence/absence (blue/white) in
 89 each strain of the 15,682 genes from the pan genome, sorted by gene frequency among the 48 strains.
 90 The 2864 first genes to the left represent the core genome.

91 The GWAS was thus performed based on the presence or absence of genes in iRCIs vs SCIs. Only three
 92 genes were significantly more frequent in SCI genomes ($p = 0.008$, $p = 0.009$ and $p = 0.009$
 93 respectively), and one gene in iRCI genomes ($p = 0.003$). These genes were annotated as hypothetical
 94 proteins by Prokka and could not be annotated by BlastKOALA either.

95 To visualize and further explore the genomic regions that were only found in the SCI genomes, iRCI
 96 reads were mapped to their paired SCI assembly. Interestingly, very large SCI contigs (up to 40kb) were
 97 absent in paired iRCI, some of them containing genes from the *tra* operon, which suggested that
 98 plasmids may be less frequent in iRCIs than in SCIs.

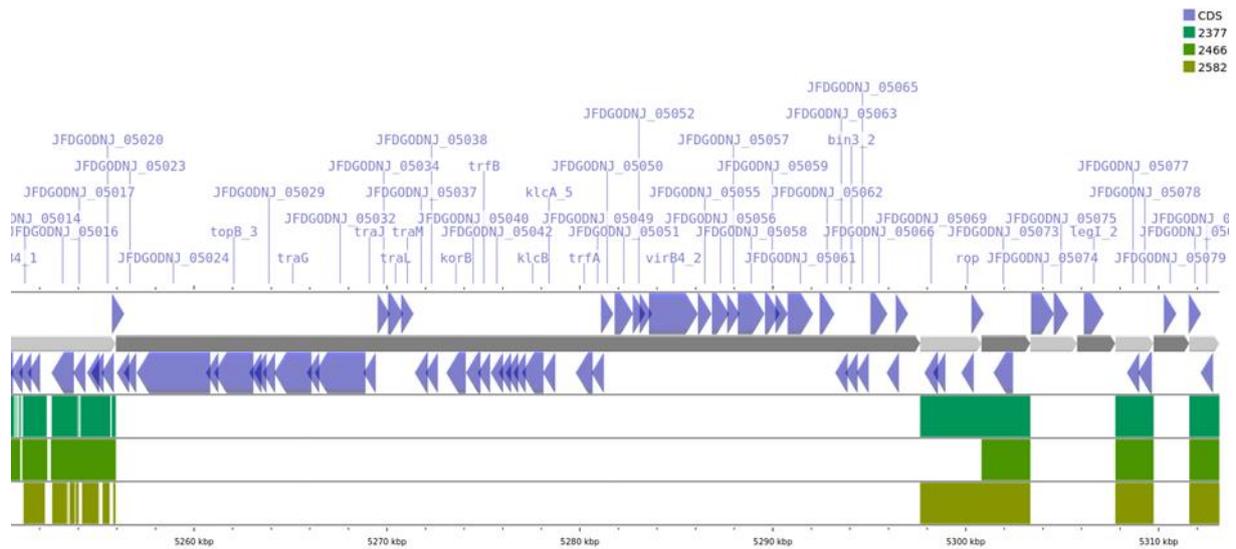
99 **Hybrid assembly improves the resolution of genome structure but**
100 **not the coverage**

101 To better describe the iRCI genomes and especially their plasmids, we performed long-read WGS and
102 hybrid assembly on the 24 iRCIs investigated so far.

103 Hybrid assembly reduced by 19-fold the mean number of contigs per genome (1 to 23 contigs by
104 genome, instead of 74 to 302 previously obtained using Illumina short read WGS). It also modestly
105 increased the total genome size (hybrid assembly mean: 5.10E+06 bases, short read assembly mean:
106 5.01E+06 bases), though mean estimated completeness remained 99.9% for hybrid as well as short read
107 assemblies according to BUSCO.

108 **Longitudinally sampled RCIs can lose plasmids over the course of**
109 **relapses**

110 To investigate longitudinally sampled RCI strain evolution over time, short reads from each RCI relapse
111 isolate (rRCI, 1 to 4 relapses per patient) were mapped to the hybrid assembly of their initial occurrence
112 (i.e. iRCI), which was used as an intra-host reference. Interestingly, in the RCI7 series, the initial isolate
113 (isolate 2359) presented several circularized plasmids including 41kb-long pRHB15-C18_3 (sequence
114 ID = [CP057780.1](#)). This plasmid contained several genes from the *tra* operon and wasn't found in the
115 successive rRCIs (relapse isolates 2377, 2466, 2582, see Fig 2), thus suggesting that RCIs may lose
116 plasmids over the course of relapses.



117

118 **Fig 2. Evidence of plasmid loss in sequential relapses in the RCI7 series.** The 8 smaller contigs of
 119 iRCI 2359 hybrid genome assembly are shown in an alternance of dark and light grey arrows. The large
 120 dark grey contig to the left corresponds to the 41kb-long pRHB15-C18_3 plasmid. Just above and below,
 121 blue arrows indicate the location and orientation of CDSs on the contigs. The corresponding gene names
 122 are indicated on the top, including *traG*, *traJ*, *traL* and *traM* on pRHB15-C18_3. Below, the coverage
 123 of these contigs by the reads from rRCIs 2377, 2466 and 2582 (relapses of iRCI 2359) is shown in 3
 124 shades of green.

125 **Longitudinally sampled RCI SNPs often occur in Open Reading** 126 **Frames, especially in metabolism and ABC membrane transporters**

127 To further investigate RCI genomic micro-evolution, we studied all the SNPs identified in rRCIs
 128 compared to their respective iRCI. A global analysis of the total 1,435 SNPs identified showed that more
 129 than half of them (N=756) occurred in Open Reading Frames (ORFs); among them, 505 were non-
 130 synonymous. Functional annotation of the 369 proteins in which SNPs occurred over time was
 131 successful for 87 proteins (23.6% of submitted proteins) as shown in **Table 1**. These were mostly
 132 involved in diverse metabolic pathways (27.6%) and membrane transport (20.7%), notably with ABC
 133 membrane transporter genes that acquired SNPs in 13 out of the 24 RCI investigated over time.

134

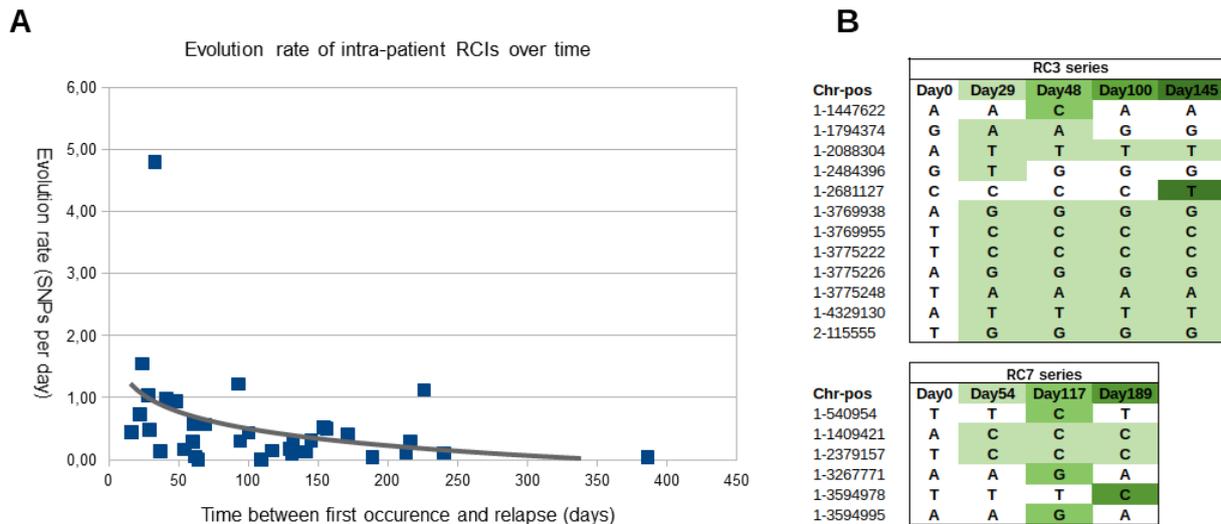
135 **Table 1. Functional annotation provided by blastKOALA for 87 proteins**

Functionnal category	Number of genes
Cellular process	10
Cell motility	1
Flagellar assembly	1
Cellular community - Prokaryotes	9
Biofilm formation - <i>Escherichia coli</i>	4
Biofilm formation - <i>Pseudomonas aeruginosa</i>	1
Biofilm formation - <i>Vibrio cholerae</i>	1
Quorum sensing	3
Environmental information processing	21
Membrane transport	18
ABC transporters	15
Bacterial secretion system	2
Phosphotransferase system (PTS)	1
Signal transduction	3
Two-component system	3
Genetic information processing	11
Folding, sorting and degradation	2
RNA degradation	2
Replication and repair	6
Base excision repair	1
Homologous recombination	5
Translation	5
Aminoacyl-tRNA biosynthesis	2
Ribosome	1
Human diseases	18
Cancer: overview	1
MicroRNAs in cancer	1
Drug resistance: antimicrobial	2
beta-Lactam resistance	1
Cationic antimicrobial peptide (CAMP) resistance	1
Drug resistance: antineoplastic	1
Platinum drug resistance	1
Endocrine and metabolic disease	3
Type I diabetes mellitus	2
Type II diabetes mellitus	1
Infectious disease: bacterial	10
Bacterial invasion of epithelial cells	1
Pertussis	8
<i>Salmonella</i> infection	1
Infectious disease: parasitic	1
Amoebiasis	1
Metabolism	24
Amino acid metabolism	2
Glycine, serine and threonine metabolism	1

Histidine metabolism	1
Biosynthesis of other secondary metabolites	1
Tropane, piperidine and pyridine alkaloid biosynthesis	1
Carbohydrate metabolism	4
Amino sugar and nucleotide sugar metabolism	2
Starch and sucrose metabolism	2
Glycan biosynthesis and metabolism	6
Lipopolysaccharide biosynthesis	4
O-Antigen nucleotide sugar biosynthesis	1
Other glycan degradation	1
Metabolism of cofactors and vitamins	4
Biotin metabolism	2
Porphyrin metabolism	1
Thiamine metabolism	1
Metabolism of other amino acids	2
Glutathione metabolism	1
Taurine and hypotaurine metabolism	1
Metabolism of terpenoids and polyketides	2
Biosynthesis of siderophore group nonribosomal peptides	2
Nucleotide metabolism	1
Purine metabolism	1
Xenobiotics biodegradation and metabolism	2
Benzoate degradation	1
Nitrotoluene degradation	1
Organismal systems	3
Environmental adaptation	3
Plant-pathogen interaction	3
Total	87

137 **RCI genomic evolution rate decrease over time**

138 Evolution rates between each rRCI and its corresponding iRCI ranged from 0 to 4,79 SNPs/day (mean:
 139 0,55 SNPs/day) and tended to decrease over time (Fig 3A). This was confirmed at the individual scale
 140 with the largest RCI series (RCI3 and RCI7, 4 and 3 relapses respectively), in which most of the long-
 141 term selected SNPs occurred between the iRCI and the first rRCI. Then, we rather observed a dynamic
 142 of emergence and clearance of various SNPs, as shown in Fig 3B.

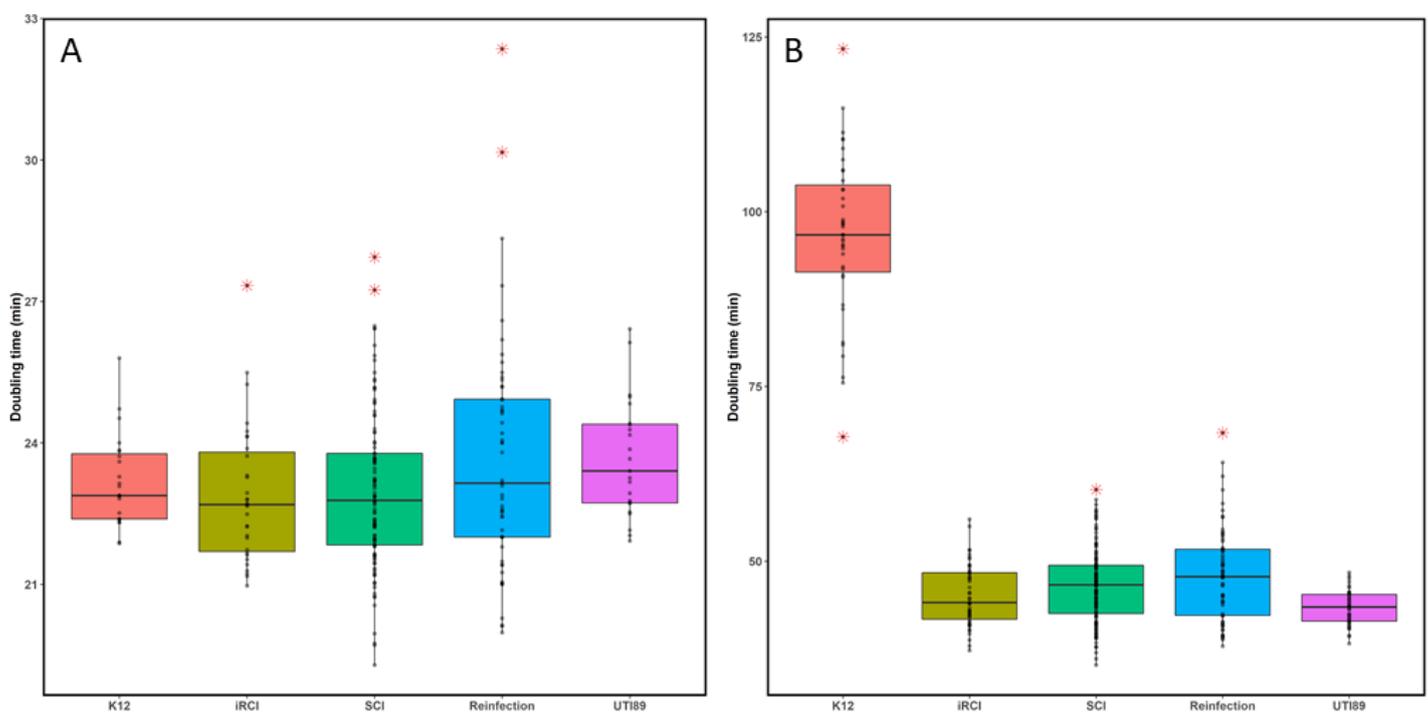


143 **Fig 3. Intra patient evolution of RCIs.** **A.** Global representation of the evolution rates (SNPs per day)
 144 from the 24 series of intra-patient RCIs depending on the time elapsed between the first occurrence and
 145 the relapse (days). **B.** Individual representation of the core SNPs observed over time in the two largest
 146 intra-patient RC series (RC3 and RC7). SNPs are color-coded depending of the time of emergence (the
 147 darker, the latter).

148 **Fitness in AUM does not discriminate initial RCIs from SCIs**

149 The fitness of iRCI, reinfection isolates (isolates responsible for infections in a given patient but
 150 genetically different from her RCIs) and SCIs was evaluated in rich medium (LB) and in artificial
 151 urinary medium (AUM). Mean doubling time in LB was similar between RCIs, SCIs and reinfection
 152 isolates (22.9 min, 23.0 min and 23.60min, resp) as well as with the laboratory strain *E. coli str.* K12
 153 *substr.* MG1655 and the well characterized UPEC strain *E. coli* UTI89, as well as for isolates from 10
 154 first and last RCI (S1 table). Isolates from the three groups grew significantly slower in AUM than in

155 LB (mean doubling times: 43.6 min, 46.2min and 47.9min resp, $p < 0.01$). Even though the doubling
 156 times were more variable in AUM (RCIs: 39.8 to 45.8min, SCIs: 36,6 to 55.8min, reinfection isolates:
 157 39.9 to 63.6min) than in LB (RCIs: 21.2 to 27.4min, SCIs: 20,1 to 26.8min, reinfection isolates: 20.4 to
 158 29.3min), no significant difference was observed between the three groups in this medium either. The
 159 growth rates of RCIs, SCIs and reinfection isolates were similar to that of the well characterized UPEC
 160 strain *E. coli* UTI89, while the laboratory strain *E. coli* str. K12 substr. MG1655 grew significantly
 161 slower than urinary isolates in AUM ($G_{K12-AUM} = 96.6$, $p < 0.01$) (Fig 4).

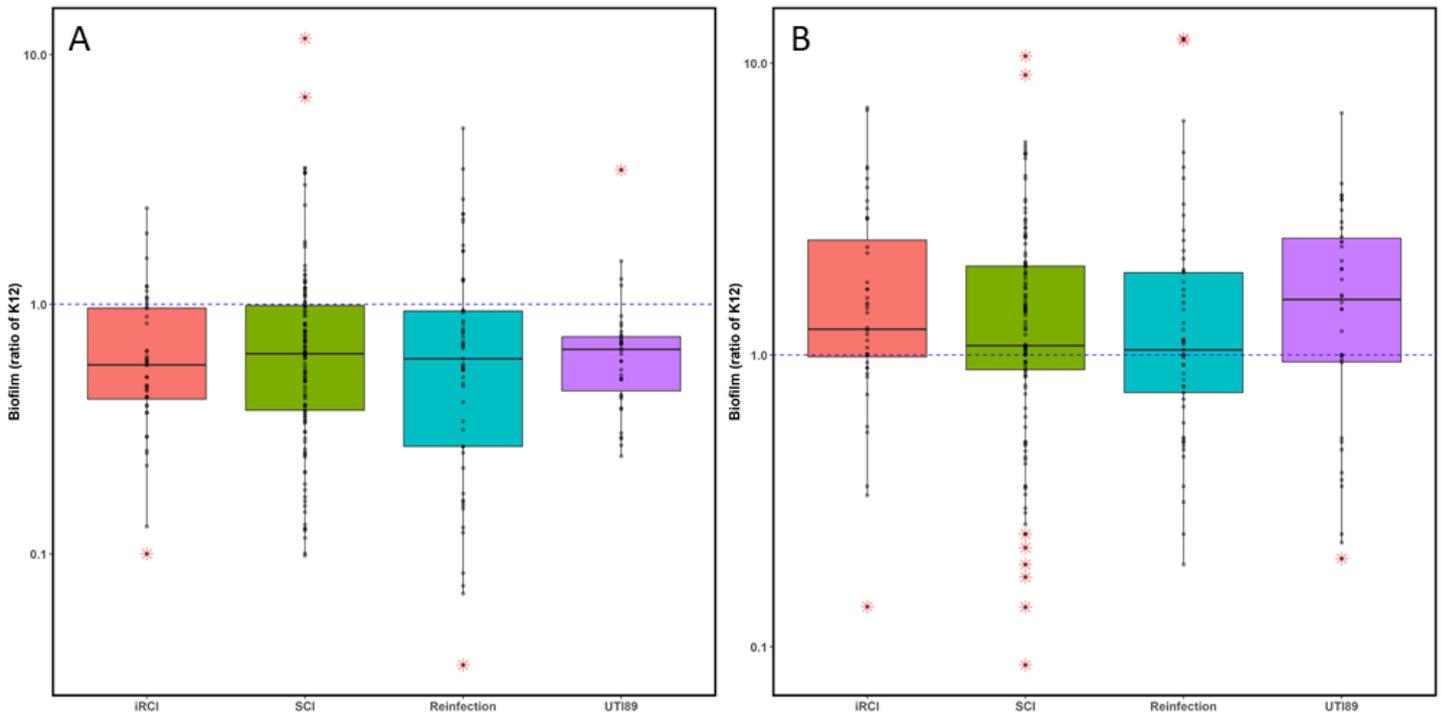


162 **Fig 4. Boxplots representing the doubling time (in min) of each isolate group in LB (A) or in AUM**
 163 **(B).** Each dot represents a mean doubling time for one isolate in the group. Red asterisks represent
 164 extreme phenotypes for a given group (outliers).

165 **RCIs and SCIs product less biofilm in AUM than in rich-medium**

166 To evaluate whether biofilm production could be a specific trait associated with recurrence, we studied
 167 biofilm formation capacity of 10 RCI pairs (10 iRCIs and their 10 respective last sampled rRCIs), the
 168 22 reinfection isolates and the 35 SCIs, in LB and in AUM. Globally, all isolates tested exhibited low
 169 levels of biofilm formation in LB ($A_{590nm} < 1$) and even lower levels in AUM ($A_{590nm} < 0.4$). Clinical

170 isolates didn't exhibit higher levels of biofilm compared to the negative control *E. coli* str. K12 in LB
171 (Fig 5).

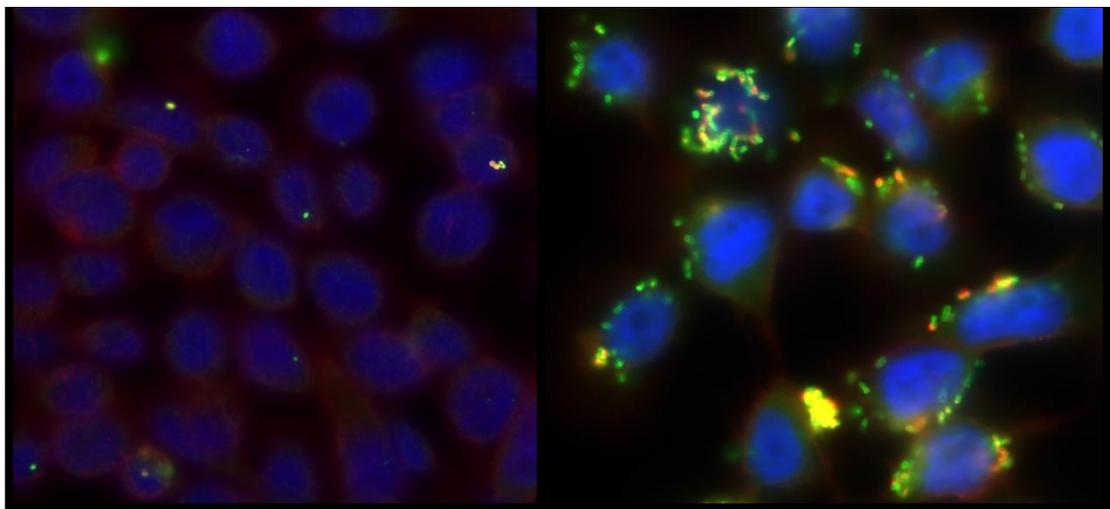
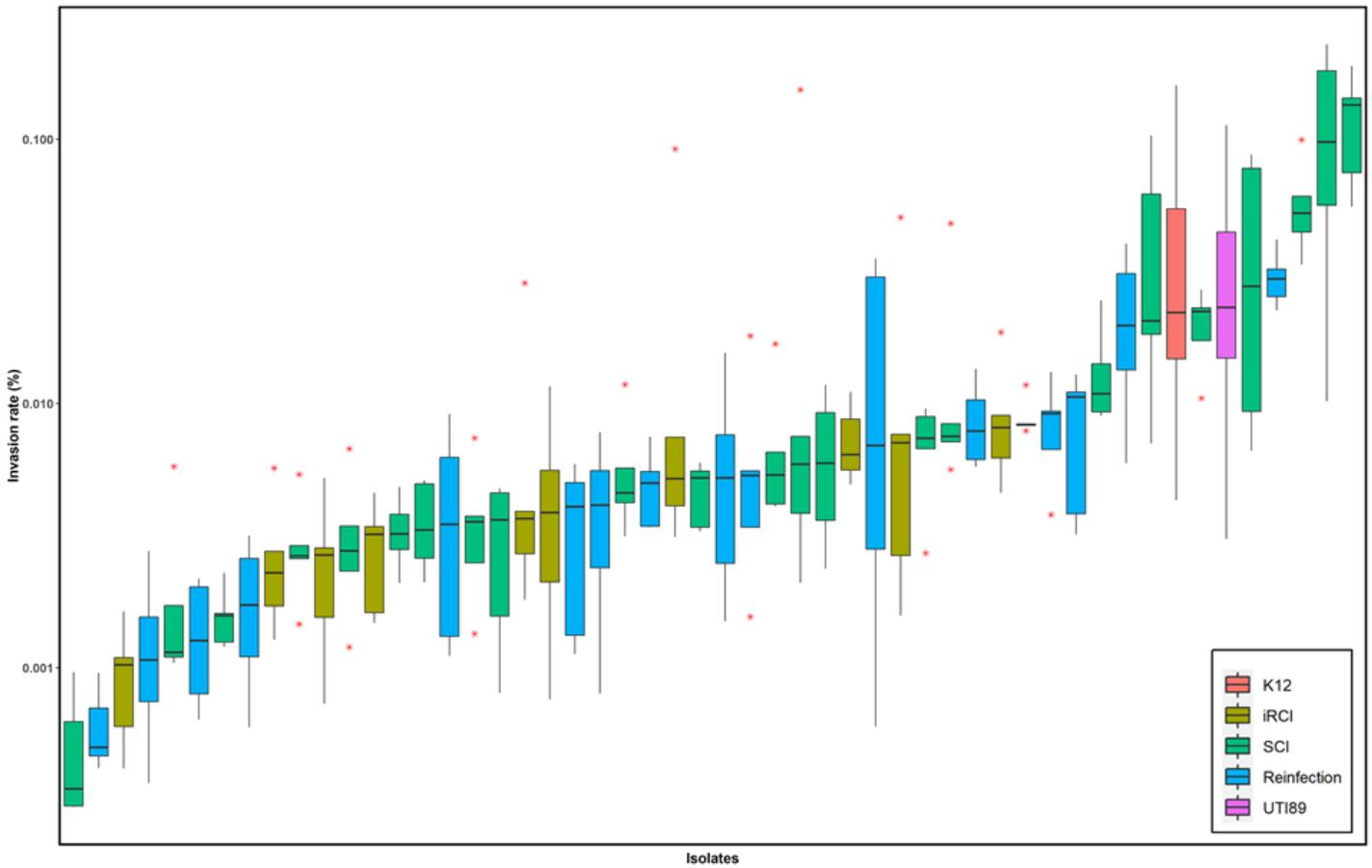


172 **Fig 5. Biofilm formation in LB (A) and in AUM (B) expressed as a ratio of K12 biofilm production**

173 **RCI did not invade BEC significantly more than SCI**

174 To determine if bladder epithelial cell (BEC) invasion was a common feature of RCI, we performed
175 gentamicin protection assay on 10 RCI pairs (10 iRCIs and their 10 respective last sampled rRCIs) and
176 35 SCIs. We observed globally very low invasion rates (between 0.0005 and 0.1194 %). Only 7 isolates
177 out of 55 (12.7%) exhibited similar or greater invasion capacity than the positive control UTI89 (Fig
178 6A). There was no correlation between the invasion capacity and recurrence, since the 7 isolates with
179 BEC invasion capacity were 2 reinfection isolates and 5 SCIs. Of note, similar invasion capacity was
180 observed between the 10 iRCIs studied and their respective last rRCIs (S1 Fig.). Interestingly, *E. coli*
181 str. K12 exhibited a similar invasion rate as *E. coli* UTI89.

182 We were able to further test 6 of the 7 isolates with the highest invasion rates by fluorescence microscopy
183 and compare them to the positive control UTI89 (Fig 6B, left panel). Intracellular bacteria were observed
184 for the 6 tested isolates (Fig 6B, right panel), thus confirming their invasion capacity.



185

186 **Fig 6. Invasion capacity of UPEC isolates in HTB-9 bladder epithelial cells after gentamicin**
 187 **protection assay. A** Invasion rates of 10 iRCIs, 16 reinfection isolates, 24 SCIs, K12 and positive
 188 control UTI89. **B.** Observation of intracellular bacteria (in green) by fluorescence microscopy. On the
 189 left panel: bacterial strain = positive control *E. coli* UTI89, on the right panel: bacterial strain = a clinical

190 SCI with high invasion rate. Red bacteria are extracellular. Yellow fluorescence is due to red and green
191 fluorescence superposition. Nucleus of BECs are colored in blue.

192 **Discussion**

193 The main objectives of our study were (i) to compare RCI and SCI to identify genomic and/or phenotypic
194 characteristics associated with recurrence, and (ii) to analyze sequential RCI isolates collected from
195 patients with recurrent cystitis, in order to characterize within patient adaptation to the urinary tract. Our
196 main findings were (i) an absence of genotypic and phenotypic characteristics significantly associated
197 with recurrence (ii) an intra-patient genomic adaptation involving occurrence of SNPs in metabolism
198 and membrane transport genes and/or large genetic loss and (iii) a rare but significant invasion rate for
199 12.7% of the UPEC tested, unrelated with recurrence.

200 To our knowledge, there is only one similar previous comparative genomic study which didn't highlight
201 significant differences between 45 recurrent and 43 non-recurrent isolates (5). Consistent with this study,
202 we found no difference in phylogenetically paired RCI and SCI genomes regarding the presence and
203 sequence of diverse genes of interest (adhesins, toxins...). Pan genome analysis and GWAS only found
204 4 genes whose presence/absence may be associated with recurrence, but their function is to be
205 investigated because they could not be annotated using Prokka nor BlastKoala.

206 However, we observed that RCIs presented a significant reduction of genome size and protein count
207 compared to phylogenetically paired SCIs. We suspected that large plasmids may be lost in RCIs, based
208 on the size of missing contigs and the fact that these contained genes from the Tra operon. Indeed, *tra*
209 genes are involved in conjugative bridges that allow for plasmid transfer (14).

210 Thus, we used long read sequencing and hybrid assembly on the 24 iRCIs, to better characterize their
211 genomes and plasmids. Even though short read assembly had already reached 99,9% completeness
212 according to BUSCO, hybrid genome structure was better resolved, with almost 20 times less contigs
213 per genome. Mapping reads from the subsequent occurrences of each rRCI to its initial hybrid reference
214 assembly (iRCI) thus allowed us to investigate intra-patient evolution and confirm that plasmids were
215 lost over time. *E. coli* is known for his great genomic plasticity mediated by gain and loss of mobile

216 genetic elements (15). As recently suggested by Thänert et al, loss of mobile genetic elements such as
217 plasmids could be associated with UPEC fitness and recurrence (6).

218 Longitudinal intra-patient evolution was also investigated at the SNP level, and indicated that SNPs
219 mostly occurred in genes of unknown function; but when successfully annotated, most of the genes with
220 SNPs were involved in metabolism and membrane transport. Nielsen et al. already observed that SNP
221 appearance in metabolism involved genes was frequent in longitudinal relapse isolates (5). Interestingly,
222 we found that evolution rates between iRCIs and corresponding rRCIs decreased over time: SNPs that
223 were long term conserved – thus possible adaptative SNPs - mostly occurred early, between the iRCI
224 and first rRCI. Then, less SNPs emerged and/or cleared in later rRCIs, suggesting micro-evolution with
225 less selection pressure. The plasmid loss observed in RCI7 also occurred between the iRCI and first
226 rRCI; taken together, these results suggest that RCIs may rapidly adapt to the urinary tract, followed by
227 a phase of micro-evolution of the adapted forms.

228 One of the main characteristics of UPECs is their ability to efficiently grow in urine (16). A previous
229 study estimated a 36min doubling time for UPECs in pooled human urine (17). To our knowledge, our
230 work is the first one to use AUM instead of human urine on a large isolate collection. This synthetic
231 medium provides more reproducible data that can be more easily compared to other studies than pooled
232 human urine, as pointed by a recent study (18). However, doubling times that we observed in AUM
233 were higher than those described for human urine ($G = 36\text{min}$) (17). This difference might be explained
234 by the incredible complexity of human urine, which probably contains more utilizable substrates than
235 the AUM (19).

236 Even though no mean fitness difference was observed between RCIs and SCIs, we found a notable
237 within-group diversity. It might be interesting to further investigate genomic differences between
238 isolates of extreme doubling times, in order to explore the mechanisms that lead to a better fitness in
239 AUM. Further studies might also focus on describing RCI growth in the urine of its corresponding
240 patient to see if RCIs undergo specific adaptations to patient urine composition.

241 Growth and biofilm formation are two closely related phenomena. Biofilm has also been long known as
242 a persistent source of infection in patients (20). To date, the association of biofilm formation with
243 recurrent UTI is controversial, as some studies found it to be associated to these infections (21) and
244 other ones didn't. (12). The majority of the studies described biofilm formation in commercial minimal
245 media (12,21), as nutrient stress seems to be an important factor for *E. coli* biofilm induction (22). To
246 our knowledge, this is the first study that investigates biofilm formation in AUM. As for fitness
247 experiments, we didn't find significant differences between RCIs and SCIs, confirming the results
248 obtained by Erjnaes et al (12). We observed very low biofilm production rates, especially in AUM
249 medium, preventing us from drawing conclusions about the RCI and SCI comparison and RCI intra-
250 patient evolution. However, testing conditions may have been suboptimal for biofilm formation as other
251 studies suggest that the combination of mechanical and biological stresses encountered in urinary tract
252 triggers biofilm formation (22,23). Next studies might consider performing biofilm formation assays in
253 dynamic models that mimic urine flux.

254 BEC invasion by UPECs is considered as one of the main mechanisms explaining relapses (9–11). Since
255 its discovery, this mechanism was extensively studied in murine UTI models using cystitis causing
256 reference strains as UTI89 (24–26). However, data concerning clinical isolates are lacking. Using UTI89
257 as a positive control to threshold significant internalization, we found that massive internalization of
258 clinical isolates is a rare but existing phenomenon that we also confirmed by microscopy data. No
259 evident correlation between relapse and invasion capacity was supported by our data, suggesting that
260 this factor alone can't explain the onset of recurrence. However, more complex models such as
261 organoids or *in vivo* experiments should be used to confirm the results from our screening model, and
262 further investigate the existence of quiescent intracellular reservoir. However, urothelial internalization
263 is not an exclusive mode of UPEC persistence, gastrointestinal and/or vaginal reservoir have also to be
264 considered to explain recurrence (6,27,28).

265 Overall, despite the use of a large collection of recurrent and non-recurrent UPEC, this study did not
266 allow to identify any genomic nor phenotypic traits significantly associated with recurrence. However,
267 we described here genomic within-patient adaptations that deserve to be further explored as well as the

268 quiescent intracellular reservoir as one of the mechanisms implicated in the pathophysiology of recurrent
269 UTI.

270 **Materials and methods**

271 **Ethics statements**

272 This study is based on a UPEC collection from an epidemiological study on community acquired UTI,
273 founded by the French Ministry of Health and approved by the Medical Research Ethics Committee of
274 the Rouen University hospital (VITALE study, Clinicaltrials.gov, identifier: NCT02292160).
275 Participating patients received an information letter and provided written informed consent.

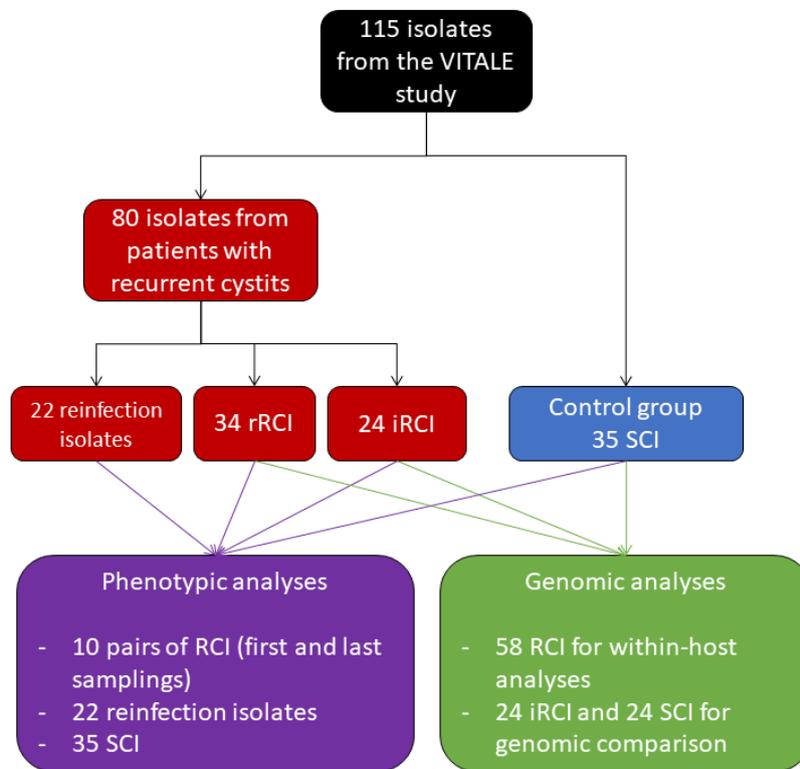
276 **Bacterial isolates**

277 One hundred and fifteen isolates from the VITALE UPEC collection were included in this study (Fig
278 4). Among them, 80 were isolated from patients with recurrent cystitis (*i.e.* with ≥ 2 episodes over a
279 study period of 6 months or ≥ 4 episodes over 12 months). Based on a previous comparative genomic
280 analysis, 58 isolates were responsible for 1 to 4 cystitis relapses (13) and were then defined as “relapse
281 causing isolates (RCIs)”. The 24 initial occurrences of these RCIs (iRCIs) together led to 34 relapses
282 (rRCIs), *i.e.* same-strain isolates responsible for later cystitis episodes of the patient. The remaining 22
283 out of the 80 isolates were responsible for a single cystitis episode during the 18-month follow-up of
284 these patients, and were defined as “reinfection isolates”.

285 A control group ($n = 35$) was formed with VITALE UPEC “sporadic isolates” (SCIs) collected from
286 patients experiencing only one cystitis episode over the 18-month period of the VITALE study. The
287 SCIs were selected based on their genetic proximity with the included RCIs, according to a phylogeny
288 constructed with CH typing data (13).

289 Comparative genomic studies were performed on 48 isolates: the 24 iRCIs and 24 of the 35
290 phylogenetically paired SCIs from the control group. Intra-host evolution studies were based on the 58
291 RCIs (24 iRCIs and 34 rRCIs). Phenotypic studies were performed on 77 isolates (Fig 4).

292 *E. coli* str. UTI89 and *E. coli* str. K12 were used as control strains for phenotypic experiments. *E. coli*
293 str. K12 was provided by Pasteur institute (CIP 106782).



294

295 **Fig 7 Bacterial isolates description.** iRCI = initial Relapse Causing Isolate, rRCI = relapse of
296 Relapse causing isolate, SCI = Sporadic Cystitis Isolate

297 Whole genome sequencing and assembly

298 Short read whole genome sequencing (WGS) and assembly were performed on SCIs and RCIs as
299 previously described (13), using the Nextera XT library kit (Illumina Inc., San Diego, CA, USA) and
300 fq2dna v21.06 (gitlab.pasteur.fr/GIPhy/fq2dna).

301 Long read WGS was also performed for the initial occurrence of each RCI (iRCI). These were amplified
302 by overnight incubation in 10mL of lysogenic broth (LB, MP biomedical, Santa-Ana, USA), at 37°C
303 and under agitation. Up to 2.10^{-9} cells were used for genomic DNA extraction using the Dneasy Blood
304 and Tissue minikit (QIAGEN, Venlo, Netherlands) according to the manufacturer recommendations for
305 isolating GRAM-negative bacteria DNA. Alternatively, the MONARCH HMW DNA Extraction kit for
306 Tissue (NEB, Ipswich, MA, USA) was used to improve DNA fragment length. DNA concentration and

307 quality were checked using a Qubit (Thermo Fisher Scientific, Waltham, MA, USA) and a Nanodrop
308 2000 (Thermo Fisher Scientific) instruments.

309 Up to 1µg genomic DNA was used for library preparation according to the « Native barcoding genomic
310 DNA protocol with EXP-NBD104, EXP-NBD114 and SQK-LSK109 » (ONT, Oxford, UK). Libraries
311 were sequenced using R9.4.1 flowcells on a minION Mk1B device. Real-time basecalling,
312 demultiplexing and filtering were performed using MinKNOW v21.11.8 and Guppy v5.1.13. Quality
313 control was then performed using Nanoplot v1.38.1 (github.com/wdecoster/NanoPlot).

314 Short and long reads from Illumina and ONT sequencing, resp, were then used for hybrid assembly
315 using Unicycler 0.4.9. Short read assemblies from the previous study (Notre ref) and newly obtained
316 hybrid assemblies were compared using Bandage v0.8.1 (rrwick.github.io/Bandage/) for visualization
317 and BUSCO v5.2.2 (busco.ezlab.org) with the enterobacterales_odb10 lineage dataset for genome
318 completeness evaluation.

319 **SCIs and RCIs comparative genomic study**

320 Short read assemblies of phylogenetically paired iRCIs and SCIs (24 each) were annotated using Prokka.
321 The protein sequences of genes of interest were retrieved from the annotated genomes and compared.
322 RCI and SCI short reads were also mapped to the sequence of pUTI89 (NC_007941.1) using Snippy
323 v4.3.6 (github.com/tseemann/snippy). More than 75% coverage of the plasmid sequence was interpreted
324 as positive match.

325 Pan-genome analysis of the 48 RCIs and SCIs was performed using Roary 3.12.0 ([sanger-
326 pathogens.github.io/Roary/](https://sanger-pathogens.github.io/Roary/)). Neighbor Joining tree inference was performed using MEGA X (29) based
327 on the core gene alignment produced by Roary. The phylogenetic tree, the core genome allele profiles
328 obtained from the raw short reads using cgMLSTFinder v1.1.5
329 (bitbucket.org/genomicepidemiology/cgmlstfinder/) and the presence/absence table produced by Roary
330 were visualized using Phandango v1.3.0 (jameshadfield.github.io/phandango/).

331 The presence/absence table produced by Roary was also used with Scoary v1.6.16
332 (github.com/AdmiralenOla/Scoary) for Genome Wide Association Study (GWAS), in order to identify

333 the genes that were significantly associated with recurrence. The protein sequences deriving from these
334 genes were submitted to BlastKOALA and Kegg Mapper Reconstruct (www.genome.jp/kegg/) to
335 classify them according to functional pathways.

336 **Intra-patient analysis of RCI gene evolution**

337 Hybrid genome assembly of each initial RCI occurrence (iRCI) was annotated using Prokka v1.14.5
338 (github.com/tseemann/prokka). The annotated genomes were used as references for SNP analysis using
339 Snippy v4.3.6 (github.com/tseemann/snippy), by mapping short reads from subsequent RCI occurrences
340 (rRCIs) to their respective iRCI hybrid genome assembly.

341 The alignment produced by Snippy for each iRCI and its mapped rRCIs was visualized using Proksee
342 (proksee.ca) to identify plasmids potentially lost over time. The corresponding contigs were blasted
343 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) to identify whether they were previously described as
344 plasmids. The genes in which snippy identified SNPs between an iRCI and its rRCIs were listed and
345 their sequences were submitted to BlastKOALA and Kegg Mapper Reconstruct
346 (www.genome.jp/kegg/) for functional classification.

347 **Growth assays**

348 AUM used in phenotypic experiments was prepared as described by Brooks and Keevil (30) with minor
349 modifications. Briefly, 100mM of 2-(N-morpholino)ethanesulfonic acid (MES) was added to improve
350 buffer capacity of previously described medium and then limit precipitate formation.

351 Briefly, isolates were grown overnight in lysogenic broth (LB) or AUM at 37°C with shaking at 150
352 rpm. These cultures were then normalized to optical density (OD_{600nm}) of 0.05 and 200µL of each
353 standardized suspension was placed in triplicate in a microtiter plate. Plate was then incubated under
354 continuous double orbital shaking conditions (108 rpm) at 37°C during 24h in a microplate reader
355 (Spark®, Tecan, Männedorf, Switzerland) and OD₆₀₀ was measured every 15 min. All experiments were
356 performed at least three times. The *E. coli* K12 and UTI89 strains were used as internal control and fresh
357 medium as negative control for each experiment. Statistical analyses of the growth curves were
358 performed using R software (v 4.2.1, <https://www.R-project.org/>).

359 **Biofilm assays**

360 Isolates were grown to late stationary phase, normalized as described above and incubated in a
361 microplate in static conditions at 37°C for 24h. Nonadherent bacteria were removed by washing with
362 H₂O before crystal violet 0.5% addition. After 10 min of incubation under gentle shaking at room
363 temperature, the excess of dye was discarded and each well was washed 3 times with H₂O. Ethanol 95%
364 was then added and the plate was incubated under gentle shaking at room temperature for 10 min.
365 Absorbance at 590nm (A_{590n}) was measured using a microplate reader. All experiments were performed
366 at least three times. The *E. coli* K12 strain was used as a positive control and fresh medium as a negative
367 control for each experiment. Data were expressed as percent of biofilm formation relative to the positive
368 control.

369 **Cell line and growth conditions**

370 Human bladder epithelial cell line 5637 (ATCC HTB-9) was maintained at 37°C and 5% CO₂ in RPMI
371 1640 media (Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum and 2mM
372 Glutamin. Supplemented RPMI will be called complete RPMI.

373 **Gentamicin protection assay**

374 Bacterial isolates were grown in LB at 37°C for 48h in static conditions to promote type I pilus
375 expression (31). Bacteria were then washed three times in PBS to eliminate eventual secreted toxins.
376 Resulting bacterial suspension was enumerated by serial dilution plating method on LB agar plates and
377 used as inoculum for infections. Confluent BEC monolayers in 12-well tissue culture plates were
378 infected by clinical UPEC isolates at a multiplicity of infection of around 15 bacteria per host cell. To
379 synchronize bacterial contact to host cells, plates were centrifugated at 750rpm for 5min at room
380 temperature. After 2 hours of incubation at 37°C, cells were washed two times using phosphate buffer
381 salt (PBS) and were incubated 1 hour at 37°C in complete RPMI containing gentamicin (50µg/mL) to
382 kill any extracellular bacteria. Gentamicin solution was then removed and cell monolayers were washed
383 three times before being lysed by Triton-X-100 0.2% solution for 10min at 37°C. BEC lysates were
384 enumerated by serial dilution plating method on LB agar plates. Invasion percentage was determined by

385 dividing the number of bacteria recovered from the lysate by the number of bacteria inoculated in the
386 plate.

387 **Fluorescence microscopy**

388 First intracellular bacteria were obtained by gentamicin protection assay experiment. However, BEC
389 didn't underwent lysis but were fixed into 500 μ L of a PBS – paraformaldehyde 4% solution. After that
390 cells were washed and stored at 4°C overnight. Then, double immunostaining was performed to
391 differentially label extracellular and intracellular bacteria. First Cells were washed with PBS and
392 incubated 1 hour with mouse anti-*E. coli* primary antibody (1:200; Abcam, Cambridge, UK) at RT in
393 antibody buffer (PBS – BSA 1% solution). Cells were washed to remove primary antibody in excess
394 and they were incubated for 1 hour in antibody buffer with Alexa Fluor546-conjugated goat anti-mouse
395 secondary antibody (1:500; Molecular Probes, Eugene, OR, USA). Cells were then washed with PBS
396 and permeabilized by a 15min treatment with a PBS-Triton-X-100 (0.3%) solution at RT. Permeabilized
397 cells were then treated again with primary antibody for 1h and with antibody buffer containing Hoechst
398 (1:1,000), Alexa Fluor488-conjugated goat anti-mouse secondary antibody (1:500, Molecular Probes)
399 and Alexa Fluor647-conjugated phalloidin (1:100; Molecular probes). Cell monolayers were mounted
400 in Fluoromount reagent (Invitrogen, Waltham, MA, USA) and images were acquired with a Leica
401 Thunder tissue 3D microscope and processed with ImageJ software
402 (<https://imagej.nih.gov/ij/index.html>).

403 **References**

- 404 1. Foxman B. Epidemiology of Urinary Tract Infections: Incidence, Morbidity, and Economic Costs.
405 2002;9.
- 406 2. Foxman B. The epidemiology of urinary tract infection. *Nat Rev Urol.* déc 2010;7(12):653-60.
- 407 3. Glover M, Moreira CG, Sperandio V, Zimmern P. Recurrent urinary tract infections in healthy and
408 nonpregnant women. *Urol Sci.* mars 2014;25(1):1-8.
- 409 4. Medina M, Castillo-Pino E. An introduction to the epidemiology and burden of urinary tract
410 infections. *Ther Adv Urol.* déc 2019;11:1756287219832172.
- 411 5. Nielsen KL, Stegger M, Kiil K, Lilje B, Ejrnæs K, Leihof RF, et al. *Escherichia coli* Causing
412 Recurrent Urinary Tract Infections: Comparison to Non-Recurrent Isolates and Genomic
413 Adaptation in Recurrent Infections. *Microorganisms.* 30 juin 2021;9(7):1416.
- 414 6. Thänert R, Choi J, Reske KA, Hink T, Thänert A, Wallace MA, et al. Persisting uropathogenic
415 *Escherichia coli* lineages show signatures of niche-specific within-host adaptation mediated by
416 mobile genetic elements. *Cell Host Microbe.* juill 2022;30(7):1034-1047.e6.
- 417 7. Lüthje P, Brauner A. Virulence Factors of Uropathogenic *E. coli* and Their Interaction with the
418 Host. In: *Advances in Microbial Physiology* [Internet]. Elsevier; 2014 [cité 18 mai 2022]. p. 337-72.
419 Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0065291114000071>
- 420 8. Murray BO, Flores C, Williams C, Flusberg DA, Marr EE, Kwiatkowska KM, et al. Recurrent
421 Urinary Tract Infection: A Mystery in Search of Better Model Systems. *Front Cell Infect Microbiol.*
422 2021;11:691210.
- 423 9. Robino L, Scavone P, Araujo L, Algorta G, Zunino P, Vignoli R. Detection of intracellular bacterial
424 communities in a child with *Escherichia coli* recurrent urinary tract infections. *Pathog Dis.* août
425 2013;68(3):78-81.
- 426 10. Robino L, Scavone P, Araujo L, Algorta G, Zunino P, Pérez MC, et al. Intracellular bacteria in the
427 pathogenesis of *Escherichia coli* urinary tract infection in children. *Clin Infect Dis Off Publ Infect
428 Dis Soc Am.* 1 déc 2014;59(11):e158-164.
- 429 11. Liu S cheng, Han X min, Shi M, Pang Z li. Persistence of uropathogenic *Escherichia coli* in the
430 bladders of female patients with sterile urine after antibiotic therapies. *J Huazhong Univ Sci
431 Technolog Med Sci.* oct 2016;36(5):710-5.
- 432 12. Ejrnæs K, Stegger M, Reisner A, Ferry S, Monsen T, Holm SE, et al. Characteristics of *Escherichia
433 coli* causing persistence or relapse of urinary tract infections: Phylogenetic groups, virulence factors
434 and biofilm formation. *Virulence.* nov 2011;2(6):528-37.
- 435 13. Vautrin N, Alexandre K, Pestel-Caron M, Bernard E, Fabre E, Leoz M, Dahyot S, Caron F.
436 Contribution of antibiotic susceptibility testing and CH typing compared to Next Generation
437 Sequencing for the diagnosis of recurrent urinary tract infections due to genetically identical
438 *Escherichia coli*: a prospective cohort study of cystitis in women. Manuscript in preparation.
- 439 14. Lu J, Peng Y, Wan S, Frost LS, Raivio T, Glover JNM. Cooperative Function of TraJ and ArcA in
440 Regulating the F Plasmid *tra* Operon. DiRita VJ, éditeur. *J Bacteriol* [Internet]. janv 2019 [cité 20
441 oct 2022];201(1). Disponible sur: <https://journals.asm.org/doi/10.1128/JB.00448-18>

- 442 15. Kaper JB, Nataro JP, Mobley HLT. Pathogenic *Escherichia coli*. Nat Rev Microbiol. févr
443 2004;2(2):123-40.
- 444 16. Reitzer L, Zimmern P. Rapid Growth and Metabolism of Uropathogenic *Escherichia coli* in Relation
445 to Urine Composition. Clin Microbiol Rev. 18 déc 2019;33(1):e00101-19.
- 446 17. Forsyth VS, Armbruster CE, Smith SN, Pirani A, Springman AC, Walters MS, et al. Rapid Growth
447 of Uropathogenic *Escherichia coli* during Human Urinary Tract Infection. mBio. 6 mars
448 2018;9(2):e00186-18.
- 449 18. Hogins J, Fan E, Seyan Z, Kusin S, Christie AL, Zimmern PE, et al. Bacterial Growth of
450 Uropathogenic *Escherichia coli* in Pooled Urine Is Much Higher than Predicted from the Average
451 Growth in Individual Urine Samples. Prokesh BC, éditeur. Microbiol Spectr. 26 sept 2022;e02016-
452 22.
- 453 19. Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, et al. The Human Urine Metabolome.
454 Dzeja P, éditeur. PLoS ONE. 4 sept 2013;8(9):e73076.
- 455 20. Donlan RM. Biofilm Formation: A Clinically Relevant Microbiological Process. Clin Infect Dis.
456 15 oct 2001;33(8):1387-92.
- 457 21. Soto SM, Smithson A, Horcajada JP, Martinez JA, Mensa JP, Vila J. Implication of biofilm
458 formation in the persistence of urinary tract infection caused by uropathogenic *Escherichia coli*.
459 Clin Microbiol Infect. oct 2006;12(10):1034-6.
- 460 22. Chu EK, Kilic O, Cho H, Groisman A, Levchenko A. Self-induced mechanical stress can trigger
461 biofilm formation in uropathogenic *Escherichia coli*. Nat Commun. déc 2018;9(1):4087.
- 462 23. Eberly AR, Floyd KA, Beebout CJ, Colling SJ, Fitzgerald MJ, Stratton CW, et al. Biofilm
463 Formation by Uropathogenic *Escherichia coli* Is Favored under Oxygen Conditions That Mimic the
464 Bladder Environment. Int J Mol Sci. 30 sept 2017;18(10):E2077.
- 465 24. Mulvey MA, Schilling JD, Hultgren SJ. Establishment of a persistent *Escherichia coli* reservoir
466 during the acute phase of a bladder infection. Infect Immun. juill 2001;69(7):4572-9.
- 467 25. Eto DS, Sundsbak JL, Mulvey MA. Actin-gated intracellular growth and resurgence of
468 uropathogenic *Escherichia coli*. Cell Microbiol. avr 2006;8(4):704-17.
- 469 26. Schwartz DJ, Chen SL, Hultgren SJ, Seed PC. Population dynamics and niche distribution of
470 uropathogenic *Escherichia coli* during acute and chronic urinary tract infection. Infect Immun. oct
471 2011;79(10):4250-9.
- 472 27. Worby CJ, Schreiber HL, Straub TJ, van Dijk LR, Bronson RA, Olson BS, et al. Longitudinal multi-
473 omics analyses link gut microbiome dysbiosis with recurrent urinary tract infections in women. Nat
474 Microbiol. mai 2022;7(5):630-9.
- 475 28. Salazar AM, Neugent ML, De Nisco NJ, Mysorekar IU. Gut-bladder axis enters the stage:
476 Implication for recurrent urinary tract infections. Cell Host Microbe. août 2022;30(8):1066-9.
- 477 29. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics
478 Analysis across Computing Platforms. Mol Biol Evol. 1 juin 2018;35(6):1547-9.
- 479 30. Brooks T, Keevil CW. A simple artificial urine for the growth of urinary pathogens. Lett Appl
480 Microbiol. mars 1997;24(3):203-6.

481 31. Blango MG, Ott EM, Erman A, Veranic P, Mulvey MA. Forced Resurgence and Targeting of
482 Intracellular Uropathogenic Escherichia coli Reservoirs. Beloin C, éditeur. PLoS ONE. 25 mars
483 2014;9(3):e93327.

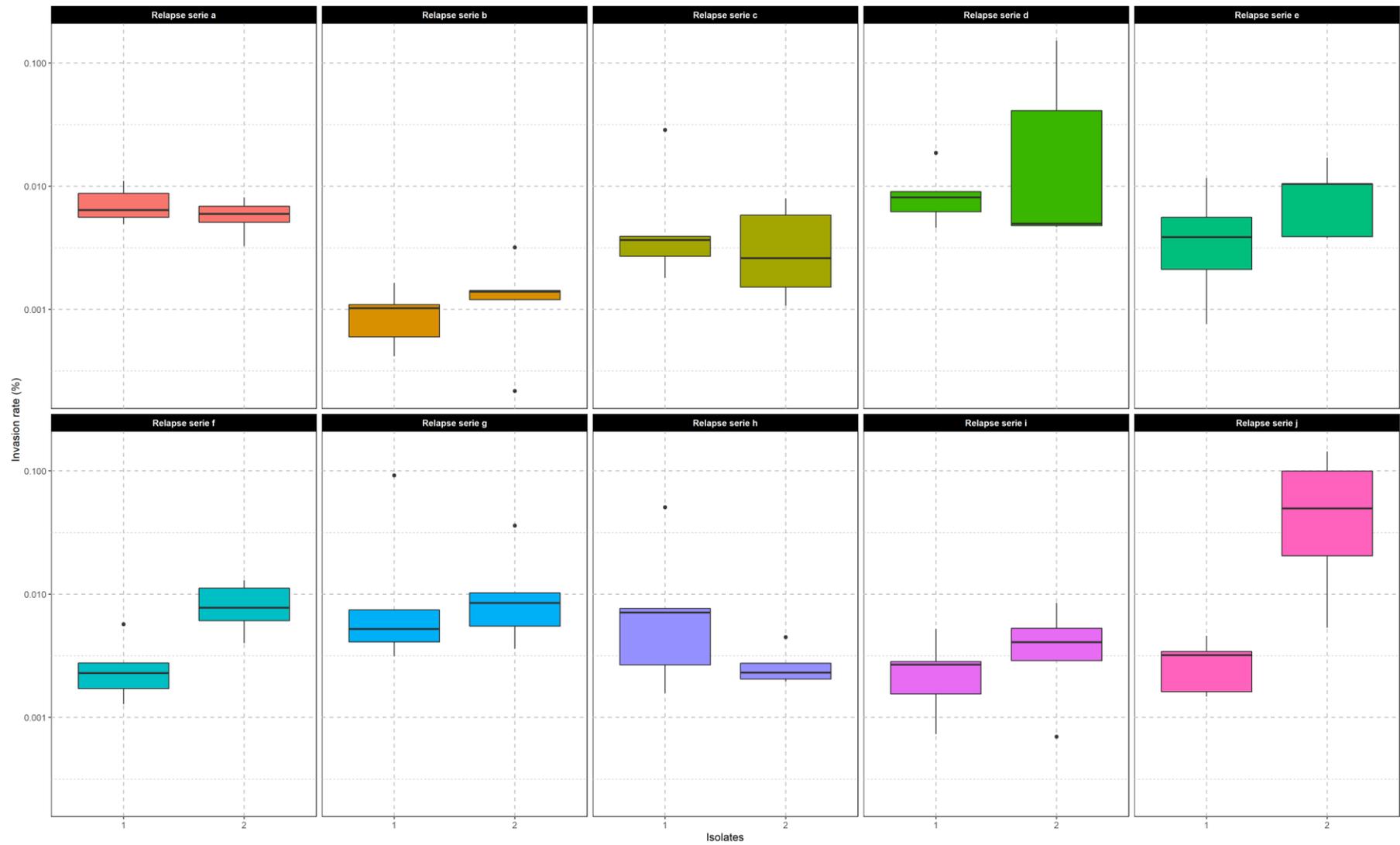
484 **Supporting information**

485 **S1 Table. Doubling times of first and last isolates for each relapse series.** LB = lysogenic broth,
486 AUM = Artificial urine medium, G = mean doubling time (in minutes)

487 **S1 Fig. Comparison of invasions rates between iRCI (isolate 1) and last rRCI (isolate 2)**

S1 Table. Doubling times of first and last isolates for each relapse series. LB = lysogenic broth, AUM = Artificial urine medium, G = mean doubling time (in minutes)

Relapse Series	LB		AUM	
	G iRCI (min)	G rRCI (min)	G iRCI (min)	G rRCI (min)
1	22,18	22,15	42,76	39,83
2	22,93	24,28	43,46	45,59
3	22,13	23,44	43,24	40,49
4	22,67	21,22	49,77	44,55
5	24,63	27,36	42,54	44,30
6	21,79	21,46	44,68	45,76
7	22,42	23,85	42,03	41,66
8	25,65	22,55	49,69	45,18
9	22,09	21,98	51,27	45,24
10	22,47	23,32	43,74	43,56



S1 Fig. Comparison of invasion rates between iRCI (isolate 1) and last rRCI (isolate 2)

II.3. Principaux résultats

II. 3. 1. [Caractéristiques génomiques des souches d'UPEC à l'origine de rechutes de cystites](#)

Caractéristiques génomiques des souches d'UPEC responsables de rechutes

La taille moyenne des génomes des souches ayant causé des rechutes était significativement plus faible que celle observée pour les souches du groupe contrôle (5,01 Mb vs 5,11 Mb). Cette observation a été corrélée à un nombre moyen de séquences codant des protéines (CDS) significativement plus faible chez les souches responsables de rechutes (4 658 vs 4 776 pour le groupe contrôle). La comparaison par paires de chaque souche responsable de rechute avec une souche de CS proche génétiquement (appariée sur le *CH type*) a mis en évidence que de larges *contigs* contenant l'opéron *tra* étaient absents chez certaines souches responsables de rechutes, pouvant indiquer la perte d'un plasmide conjugatif.

Recherche de marqueurs génomiques associés aux souches responsables de rechutes

La prévalence de gènes de virulence déjà décrits chez les UPEC était similaire entre le groupe « rechute » et le groupe contrôle. C'est pourquoi, une analyse sans *a priori* a été entreprise pour tenter d'identifier des marqueurs génomiques, dont la prévalence varierait significativement entre les deux groupes. Cette analyse a identifié trois CDS significativement plus fréquents dans le groupe contrôle ($p < 0,01$) et un marqueur qui, à l'inverse, était plus fréquent dans le groupe « rechute » ($p = 0,003$). La fonction des protéines correspondant à ces quatre marqueurs est inconnue à ce jour.

Adaptation génomique des souches responsables de rechute au cours du temps

Au total, 1 435 SNPs ont été identifiés en comparant les isolats de rechute avec leur souche index. Parmi ces SNPs, 756 étaient localisés dans des CDS, la plupart (66,8%) étaient non synonymes. Parmi les 369 CDS touchés par au moins un SNP, seuls 87 ont pu être annotés fonctionnellement. Les groupes fonctionnels les plus représentés étaient les protéines impliquées dans le métabolisme (27,6%), dans des voies très diverses, ainsi que les protéines impliquées dans le transport membranaire (20,7%).

II. 3. 1. [Caractéristiques phénotypiques des souches d'UPEC à l'origine de rechutes de cystites](#)

Comparaison du fitness et de la formation de biofilm de souches causant des rechutes et de souches sporadiques

Les souches du groupe « rechute » ne présentaient pas un meilleur fitness en milieu riche ou en milieu mimant l'urine (MMU) que les souches du groupe contrôle. Une très grande hétérogénéité a été observée au sein des deux groupes en termes de vitesse de croissance. Globalement, l'ensemble

de ces souches d'UPEC se développaient significativement plus vite que la souche intestinale témoin *E. coli* K12 en milieu mimant l'urine.

La production de biofilm de toutes les souches testées était très faible en milieu riche comme dans le MMU. Aucune différence n'a pu être mise en évidence entre le groupe « rechute » et le groupe contrôle.

Comparaison de la capacité d'invasion *in vitro* de souches responsables de rechutes et de souches sporadiques

Le phénomène d'internalisation semble être un évènement peu fréquent puisque les taux d'internalisation observés étaient globalement faibles (entre 0,0005 et 0,1194 %). Sur les 52 souches testées, 7 (13,5%) ont présenté un taux d'invasion équivalent ou supérieur à celui du témoin positif choisi pour cette expérience : *E. coli* UTI89. La présence de bactéries intracellulaires a été confirmée par microscopie à fluorescence pour 6 de ces 7 souches souches. Aucun lien n'a pu être mis en évidence entre le caractère récidivant de la souche et la capacité à infecter les cellules urothéliales. La souche commensale intestinale (*E. coli* K12), présentait une capacité d'internalisation similaire à celle de UTI89.

Évolution du phénotype des souches responsables de rechute au cours du temps

Quel que soit le caractère phénotypique étudié, il n'a pas été observé d'évolution significative du phénotype des souches responsables de rechute au cours du temps (rechutes espacées de 16 à 386 jours, moyenne de 111 jours).

À notre connaissance, ce travail constitue une des premières analyses génomiques et phénotypiques sur une aussi large collection prospective d'isolats cliniques d'UPEC responsables de rechutes ou de cystites sporadiques.

Pris dans leur ensemble, nos résultats suggèrent que les souches de *E. coli* responsables de rechutes de cystite sont finalement des UPEC « ordinaires », c'est-à-dire des clones particulièrement bien adaptées à l'environnement urinaire, mais ne se distinguant pas des souches de réinfection ou de cystites sporadiques. En effet notre étude longitudinale des rechutes n'a pas permis à ce stade d'identifier de mécanismes précis d'adaptation. Elle a toutefois mis en évidence une évolution génomique, affectant plus particulièrement le métabolisme bactérien ainsi que les systèmes de transport membranaire. Elle a aussi identifié un phénomène de perte de matériel plasmidique chez certains isolats dont l'impact sur l'aptitude à la rechute reste à investiguer.

III. Discussion générale

L'objectif de ce doctorat était d'étudier les facteurs microbiens associés aux CR de la femme. Pour ce faire, nous avons travaillé sur une collection d'isolats d'UPEC, issue d'une cohorte prospective de 323 femmes atteintes de cystites communautaires, sporadiques ou récidivantes.

Ce travail de doctorat a comporté deux volets dont les objectifs étaient les suivants :

- 1- Étudier la diversité génotypique des souches d'UPEC responsables de CR, et définir la place des rechutes dans le phénomène de récurrence ;
- 2- Rechercher des caractères génomiques et/ou phénotypiques associés aux souches d'UPEC responsables de rechutes, et étudier l'adaptation de ces dernières au sein du tractus urinaire au cours du temps.

Nous allons ici discuter les traits saillants de ces travaux ainsi que leurs limites au regard des données de la littérature.

III.1. Constitution de la cohorte

La cohorte de patientes de cette étude a été constituée au cours d'un protocole de recherche sur la performance du système de soins (VITALE NCT02292160). Dans le cadre de ce doctorat, nous avons sélectionné les patientes recrutées dans un laboratoire d'analyses de biologie médicale de proximité (LABM d'Elbeuf), de façon à limiter la perte de vue des patientes et avoir le meilleur suivi longitudinal possible. Il s'agit, à notre connaissance, du plus long travail de suivi longitudinal de patientes atteintes d'infections urinaires qui existe à ce jour (*Liu et al., 2016, Thänaert et al. 2022*, suivi de patientes sur 6 mois et *Nielsen et al., 2021*, suivi sur 1 mois et demi).

Cependant notre approche n'est pas dénuée de biais. Tout d'abord, bien que les cystites aient été cliniquement bien caractérisées (inclusions contrôlées par un staff d'infectiologues), les facteurs liés à l'hôte, et notamment ceux décrits dans la littérature comme étant impliqués dans le phénomène de récurrence (ménopause, activité sexuelle, facteurs anatomiques...), n'ont pas été relevés (*Hooton, 2001*). Ensuite, la période d'inclusion de 17 mois reste relativement courte au regard de l'histoire des patientes présentant des CR. Il aurait ainsi été intéressant de suivre la microévolution intra-hôte des souches récidivantes sur une période beaucoup plus longue. Par exemple, *Forde et al* ont étudié la dynamique d'une population de *E. coli* ST131 sur une période de 5 ans chez une femme âgée souffrant de CR (*Forde et al., 2019*). Enfin, l'assignation des patientes au groupe CR ou CS a reposé uniquement sur le nombre d'épisodes d'IU inclus durant la période de l'étude. Les antécédents d'IU de ces patientes n'ayant pas été recherchés par exemple via un questionnaire, il n'est pas exclu que certaines patientes

classées en CS souffraient en fait de CR. De plus, certaines patientes ont possiblement été perdues de vue au cours de l'étude.

Il serait intéressant d'étendre cette étude à un plus grand nombre de patientes en constituant une nouvelle cohorte prospective dédiée.

III.2. Evaluation de la diversité génotypique des souches d'UPEC

Les UPEC sont un groupe hétérogène de souches sur le plan génétique. Plusieurs études ont observé cette diversité, que ce soit avec une méthode de référence comme le typage par MLST (*Yamaji et al., 2018*) ou des méthodes alternatives comme le *CH typing* (*García-Meniño et al., 2022*). Nos travaux ont également été le témoin de cette diversité, dans la mesure où sur les 347 isolats d'UPEC que nous avons collectés, 119 *CH types* différents ont été identifiés. Nous n'avons par ailleurs pas mis en évidence de lignée de souches associée à la récurrence. Quatre des cinq *CH types* les plus fréquemment retrouvés dans notre collection étaient aussi parmi les plus représentés au sein des 100 souches de cystites simples génotypées par García-Meniño *et al.* (*García-Meniño et al., 2022*).

III.3. Identification des rechutes et évaluation de leur place dans le phénomène de récurrence

Au sein des CR, une distinction est faite entre les rechutes (épisodes liés à une même souche) et les réinfections (épisodes causés par des souches distinctes) (*Hooton, 2001*). Lorsqu'une même espèce est identifiée au cours du temps, ces définitions sont conditionnées par le pouvoir discriminant des méthodes de typage utilisées pour distinguer les souches. Plusieurs études cliniques utilisant la PFGE, ont établi que les rechutes constituent une part très importante des épisodes de récurrence à *E. coli* (entre 70 et 77%) (*Beerepoot et al., 2012; Ejrnaes et al., 2006; Skjøt-Rasmussen et al., 2011*). En utilisant le *CH typing*, méthode basée sur le polymorphisme de séquence de deux gènes (*Weissman et al., 2012*), nous avons mis en évidence une proportion de rechute moins importante (30,6%). Grâce à une analyse SNP des génomes, la proximité phylogénétique a été confirmée pour chaque paire de souches sauf une, montrant ainsi que la part des rechutes dans la récurrence était probablement surestimée par les « anciennes » méthodes de typage, probablement du fait de leur pouvoir discriminant inférieur.

Nos travaux soulignent également l'intérêt du *CH typing* dans l'identification des rechutes. Cette méthode semble en effet représenter le meilleur compromis coût, complexité et performance, comparativement au typage par SNP qui reste une méthode laborieuse et difficile à standardiser d'un laboratoire à l'autre (*Besser et al., 2018*).

Du fait de la progression importante de la résistance aux antibiotiques chez les Entérobactérales impliquées dans des IU ces dernières années (*Thänert et al., 2019*), la comparaison des antibiotypes a

montré dans nos travaux son intérêt en tant que première approximation de la clonalité d'isolats récidivants d'UPEC.

III.4. Comparaison génomique et phénotypique de souches récidivantes et sporadiques

III. 4. 1. Recherche de déterminants génomiques associés à la récidence

L'un des objectifs de ce travail de thèse était d'identifier des déterminants génomiques associés à la récidence, et plus particulièrement au phénomène de rechute. Pour ce faire, nous avons comparé les génomes des 24 souches d'UPEC responsables de rechutes (sélection des isolats responsables des épisodes dits « initiaux » sur la période de l'étude) à ceux de 24 souches responsables de CS ; ces souches ont été appariées sur la base des données de *CH typing* afin d'obtenir un fond phylogénétique le plus proche possible.

Sur la base des facteurs de virulence déjà décrits pour le pathovar des UPEC (*Johnson and Russo, 2018*), nous n'avons pas mis en évidence de profil de virulence associé aux souches responsables de rechute. Il n'y a que très peu d'autres exemples d'études similaires dans la littérature. L'une d'entre elle, comparant la présence de 29 gènes de virulence chez 78 souches de *E. coli* responsables de rechute à 77 responsables de réinfection, n'a pas mis en évidence de combinaison spécifique de facteurs de virulence chez les souches responsables de rechutes (*Ejrnæs et al., 2011*).

Pour compléter cette approche avec *a priori*, nous avons entrepris un analyse génomique comparative sans *a priori*. Ainsi, nous avons identifié trois protéines hypothétiques dont la prévalence était significativement plus importante chez les souches de CS et une autre protéine, elle aussi de fonction inconnue, qui était plus représentée dans le groupe des souches récidivantes. Il serait intéressant de caractériser plus précisément ces protéines en tentant dans un premier temps d'identifier des domaines fonctionnels connus dans leur séquence protéique. Nos résultats sont comparables à ceux de Nielsen *et al.*, qui ont également utilisé une approche de génomique comparative sans *a priori* de souches de *E. coli* responsables de récurrences et d'isolats non récidivants, sans mettre en évidence de marqueurs distinctifs en les deux (*Nielsen et al., 2021*).

De manière intéressante, nous avons pu constater que la taille moyenne du génome ainsi que le nombre de CDS moyen des souches responsables de rechutes était significativement plus faible que ceux des souches responsables de CS. L'alignement des génomes produits par Illumina des souches de rechutes sur celui de leur souche sporadique appariée a mis en évidence l'absence de *contigs* d'environ 40kb portant fréquemment l'opéron *tra* (codant un pont de conjugaison) (*Lu et al., 2019*) chez les souches de rechute. Cet opéron étant principalement associé à des plasmides (*Lu et al., 2019*), cela suggère que les souches de rechute s'adaptent à l'arbre urinaire en curant leur plasmide,

probablement pour éviter le coût physiologique lié à sa présence (Thänert et al., 2022). Ces résultats méritent d'être approfondis en séquençant le génome des souches sporadiques par une technologie de séquençage dite *long-reads*, plus adaptées aux analyses d'éléments génétiques mobiles, afin de confirmer la présence de ces plasmides.

III. 4. 2. Recherche de déterminants phénotypiques associés à la récurrence

Nous avons ensuite cherché à mettre en évidence des déterminants phénotypiques associés aux souches de rechute, notre hypothèse de travail étant que ces souches sont mieux adaptées à l'arbre urinaire, ce qui leur permet de s'y maintenir et s'y développer efficacement et ainsi être à l'origine de multiples épisodes de cystites. Il s'agit à notre connaissance de la première étude utilisant cette approche comparative phénotypique.

Comparaison du fitness en milieu urinaire synthétique

L'une des principales caractéristiques des UPEC étant leur capacité à croître efficacement dans l'urine (Reitzer and Zimmern, 2019), nous nous sommes tout d'abord intéressés au fitness, en comparant la croissance des souches de rechute et des souches sporadiques dans un milieu urinaire synthétique (MMU). Notre choix s'est porté sur l'utilisation de ce milieu artificiel plutôt que sur celui d'une *pool* d'urines humaines comme adopté dans certaines études (Alteri and Mobley, 2007) car il permet d'une part une grande reproductibilité des résultats, et d'autre part facilite la comparaison avec d'autres études (Hogins et al., 2022). Cependant, une des limites de ce milieu est qu'il peut sous-estimer la complexité incroyable de l'urine humaine (Bouatra et al., 2013), qui contient vraisemblablement plus de substrats utilisables que le milieu artificiel. D'ailleurs, les temps de génération obtenus avec ce milieu synthétique (G = 45.8min) étaient plus élevés que ceux observés dans la littérature dans de l'urine humaine (G de la souche de *E. coli* CFT073 de 36min en moyenne) (Forsyth et al., 2018).

Comparaison de la capacité à former du biofilm en milieu urinaire synthétique

Nos résultats n'ont pas mis en évidence de différence significative de croissance en milieu mimant l'urine entre les souches de rechute et les souches sporadiques. Cependant, une grande diversité de comportement a pu être observée dans chacun des groupes, certaines souches se développant très rapidement dans le milieu contrairement à d'autres. Explorer plus finement le génome (chromosome et éléments génétiques mobiles) de ces souches présentant des phénotypes « extrêmes » pourrait permettre de mieux appréhender les mécanismes à l'origine d'un meilleur fitness dans l'urine.

La seconde caractéristique phénotypique sur laquelle nous nous sommes concentrés est la formation de biofilm. En effet, il est communément admis dans la littérature que les biofilms bactériens sont l'une des principales sources d'infections récurrentes (Donlan, 2001). Cependant, leur

association avec la récurrence des cystites est encore discutée. À titre d'exemple, une étude de 2006 a mis en évidence une association entre le caractère récurrent des souches d'UPEC et une capacité supérieure à former du biofilm comparativement à un groupe contrôle de souches responsables de réinfections (*Soto et al., 2006*) tandis qu'une autre étude plus récente a montré l'absence de lien entre ces deux paramètres (*Ejrnæs et al., 2011*).

Nous avons choisi d'étudier la formation de biofilm en milieu riche mais également en MMU afin de s'approcher davantage des conditions rencontrées par les souches d'UPEC au sein de l'arbre urinaire. Nos résultats ne nous permettent pas de conclure à l'existence d'un lien entre le caractère récurrent des souches et une production plus importante de biofilm. Cependant, les productions de biofilm au cours de nos expérimentations étaient particulièrement faibles pour l'ensemble des souches testées en MMU, rendant l'interprétation des résultats délicate au regard de la faible sensibilité de la méthode utilisée (coloration au crystal violet). Ces taux faibles peuvent également être la conséquence de conditions inadaptées à la formation de biofilm. Certaines études utilisant des modèles plus complexes ont en effet mis en évidence que la combinaison de stress mécaniques (flux urinaire) et biologiques (microaérophilie) auxquels la bactérie est confrontée au sein de l'arbre urinaire pourrait stimuler la production de biofilm (*Chu et al., 2018 ; Eberly et al., 2017*). Ainsi, il serait intéressant de compléter ces résultats par l'étude de la formation de biofilm en MMU dans un modèle dynamique mimant le flux urinaire.

Comparaison de la capacité d'internalisation des souches d'UPEC

De nombreuses études ont investigué le mécanisme d'internalisation des UPEC en modèle murin (*Eto et al., 2006 ; Justice et al., 2004 ; Mulvey et al., 2001*) et pour certaines ont même été jusqu'à tester différentes stratégies thérapeutiques pour éradiquer les bactéries intracellulaires (*Blango et al., 2014; Blango and Mulvey, 2010*).

Le point commun de toutes ces études antérieures est qu'elles utilisent *E. coli* UTI89, une souche de référence pour le pathovar des UPEC. Cependant, peu d'études se sont employées à caractériser le pouvoir invasif d'isolats cliniques.

Notre travail poursuivi lui sur des isolats cliniques, et nombreux, n'a pas mis en évidence de corrélation entre la capacité d'internalisation et le caractère récurrent des UPEC. Les taux d'internalisation moyens que nous avons observés sont cohérents avec l'étude de Schwartz *et al.* (entre 0,01% et 0,001% de l'inoculum s'internalisant) et montrent que l'internalisation est phénomène rare à l'échelle d'un inoculum bactérien (*Schwartz et al., 2011*). Pris dans leur ensemble, ces résultats suggèrent que si l'internalisation est un phénomène pouvant être à l'origine de rechute, ce dernier ne peut à lui seul expliquer le caractère récurrent des souches de cystite. Il a été montré qu'une rechute

n'est pas obligatoirement causée par la persistance intra-vésicale d'une souche d'UPEC, mais qu'elle peut être provoquée par la ré-inoculation vésicale à partir d'un réservoir digestif (Thänert *et al.* 2019).

La capacité des UPEC à infecter les cellules urothéliales et former des réservoirs intracellulaires quiescents est aujourd'hui le scénario privilégié pour expliquer la survenue de rechutes (Glover *et al.*, 2014; Klein and Hulgren, 2020). Il serait intéressant de poursuivre l'étude du phénomène d'internalisation en complexifiant notre modèle. En effet, si l'infection de monocouches de cellules urothéliales constitue une approche intéressante en *screening*, il est impossible d'étudier la formation de réservoirs quiescents intracellulaires dans ce type de modèle, ces derniers se formant uniquement dans les cellules non différenciées (Eto *et al.*, 2007). Parmi les modèles envisagés, des organoïdes (Sharma *et al.*, 2021) ou un modèle pluricellulaire (Suzuki *et al.*, 2019) pourraient compléter notre approche en permettant d'objectiver la formation d'éventuels réservoirs quiescents intracellulaires chez les souches responsables de rechute, ayant démontré leur capacité d'internalisation.

Malgré tout, la place de l'internalisation reste à évaluer dans un modèle physiopathologique qui n'a eu de cesse de se complexifier. Thänert *et al.* ont montré que l'intestin constitue un réservoir d'UPEC qui peut proliférer dans l'intestin à la faveur d'une altération du microbiote, puis aller coloniser la région périnéale et la vessie (Thänert *et al.*, 2019). L'équipe de Worby *et al.* a, quant à elle, mis en évidence un lien entre la composition du microbiote intestinal et la récurrence. Leurs travaux ont en effet montré que la diminution de la part d'espèces productrices d'acides gras à chaîne courte dans l'intestin crée un environnement inflammatoire dans la vessie favorisant les infections récurrentes (Worby, Schreiber, *et al.*, 2022).

A défaut ou en complément d'une internalisation dans les cellules urothéliales, les souches d'UPEC responsables de rechute pourraient également former des réservoirs intracellulaires dans la muqueuse vaginale comme cela a été démontré par Brannon *et al.* dans la lignée de cellules épithéliales vaginales VK2 E6/E7 (Brannon *et al.*, 2020).

III. 4. 3. Suivi longitudinal des souches responsables de rechute

Au sein de la cohorte VITALE, nous avons identifié 23 patientes ayant présenté deux à cinq épisodes causés par une même souche de *E. coli*. L'analyse *whole genome SNP* longitudinale a mis en évidence 369 CDS touchés par un ou plusieurs SNPs. Parmi ces CDS identifiées, seules 23,6% ont pu être annotées malgré l'utilisation de deux banques de données différentes (Prokka basé sur la banque de données NCBI et blastKOALA basé sur la banque de données KEGG).

Parmi les protéines annotées, un grand nombre étaient impliquées dans diverses voies métaboliques, comme l'ont observé par Nielsen *et al.* sans qu'aucune voie précise ne soit

particulièrement touchée (Nielsen et al., 2021). Ce résultat suggère néanmoins l'importance du métabolisme dans l'adaptation au milieu urinaire. D'autres études ont montré que l'apparition de SNP ou la perte de certains gènes pourtant centraux dans le métabolisme peuvent constituer des adaptations aux conditions particulières rencontrées au sein de l'arbre urinaire. Ainsi, une étude de 2019 a montré que la perte de *pykA* (pyruvate kinase II) pouvait se traduire par une amélioration du fitness de la bactérie dans un modèle murin d'IU (Alteri et al., 2019). Une autre équipe a également montré l'existence d'un allèle du gène *argI* (enzyme terminale de la voie de biosynthèse de l'arginine) spécifique des UPEC et non fonctionnel, les rendant donc auxotrophe des vis-à-vis de l'arginine. Cependant, le remplacement de cet allèle par un autre qui lui est fonctionnel provoque une altération de la capacité des UPEC à causer des bactériuries persistantes, démontrant ainsi que l'inactivation du gène *argI* constitue une adaptation des souches d'UPEC à l'environnement urinaire (Hibbing et al., 2020).

Il est intéressant de noter que la deuxième fonction la plus représentée chez les gènes touchés par les SNPs était le transport membranaire et plus particulièrement des gènes impliqués dans la production de transporteurs dits *ATP binding cassette* (transporteurs ABC). Ces transporteurs sont majoritairement impliqués dans l'import de substrats du milieu extérieur vers le cytoplasme chez *E. coli* et jouent un rôle important dans l'utilisation de substrats variés pour la croissance (Moussatova et al., 2008). L'apparition de SNPs au cours du temps dans ces gènes constitue donc une piste intéressante pour l'étude de l'adaptation des souches responsables de rechutes à leur environnement.

Sur le plan phénotypique, peu de variations ont été observées entre le premier et le dernier isolat de chaque série de rechute en termes de fitness et de formation de biofilm. De la même façon, la capacité d'invasion des souches responsables de rechutes n'a pas évolué au fil des épisodes. Aucune autre étude comparable n'a été publiée à ce jour, portant sur de nombreuses patientes et une telle succession d'épisodes cliniques. Cependant, notre approche mériterait d'être reprise avec à la fois un suivi davantage prolongé et des données cliniques permettant d'approcher aussi les facteurs de rechute ou de réinfection liés à l'hôte.

IV. Conclusion générale

Ce travail de doctorat souligne la complexité de la physiopathologie des cystites récidivantes UPEC. Les rechutes surviennent chez près de 30% des patientes montrant que les réinfections occupent une place au moins aussi importante que les rechutes dans le phénomène de récurrence.

En comparant sur le plan génomique les souches de rechutes et les souches sporadiques, nous avons mis en évidence la perte d'éléments génétiques mobiles putatifs chez les souches de rechutes, ayant pour conséquence une réduction significative de la taille du génome de ces souches, comparativement aux souches sporadiques. Sur le plan phénotypique, aucune différence significative n'a été mise en évidence entre les souches de rechutes et les souches sporadiques.

Le suivi longitudinal des isolats séquentiels de rechutes a pointé une possible évolution des souches de rechutes par accumulation de SNPs, notamment dans des gènes impliqués dans le métabolisme et les transporteurs membranaires. Cependant, aucune évolution significative n'a été mise en évidence sur le plan phénotypique.

Il reste enfin à mieux comprendre l'origine anatomique des rechutes et des réinfections de cystite à UPEC, d'où pourraient découler des approches thérapeutiques différentes. La physiopathologie classique de « cystites récidivantes à partir du réservoir digestif » s'est trouvée enrichie par la découverte des bactéries internalisées dans la vessie entre deux épisodes cliniques (concept d'IBC). Cependant, il n'y a pas de dichotomie entre les sites anatomiques et la clonalité des UPEC source de récurrence. Une authentique rechute peut en effet provenir du réservoir digestif si ce clone y est très dominant pour des raisons diverses ; une authentique réinfection peut correspondre à une réactivation d'un IBC anciennement archivé, resté quiescent au cours de plusieurs épisodes cliniques avant de ressurgir.

C'est dire si les travaux analysant les interactions entre microbiotes vésical, intestinal et vaginal dans la genèse des IU de la femme ont encore bien d'inconnues à lever.

Bibliographie

- Abelson, B., Sun, D., Que, L., Nebel, R.A., Baker, D., Popiel, P., Amundsen, C.L., Chai, T., Close, C., DiSanto, M., Fraser, M.O., Kielb, S.J., Kuchel, G., Mueller, E.R., Palmer, M.H., Parker-Autry, C., Wolfe, A.J. and Damaser, M.S. 2018. Sex differences in lower urinary tract biology and physiology. *Biol Sex Differ*, **9**: 45.
- Agence Nationale de Sécurité du Médicament et des produits de santé. 2017, La consommation d'antibiotiques en France en 2016. Disponible à l'URL : <https://ansm.sante.fr/>.
- Alonso-Caballero, A., Schönfelder, J., Poly, S., Corsetti, F., De Sancho, D., Artacho, E. and Perez-Jimenez, R. 2018. Mechanical architecture and folding of *E. coli* type 1 pilus domains. *Nat Commun*, **9**: 2758.
- Alteri, C.J., Himpfl, S.D., Shea, A.E. and Mobley, H.L.T. 2019. Flexible Metabolism and Suppression of Latent Enzymes Are Important for *Escherichia coli* Adaptation to Diverse Environments within the Host. *J Bacteriol*, **201**: e00181-19.
- Alteri, C.J. and Mobley, H.L.T. 2007. Quantitative profile of the uropathogenic *Escherichia coli* outer membrane proteome during growth in human urine. *Infect Immun*, **75**: 2679–2688.
- Beerepoot, M.A.J., den Heijer, C.D.J., Penders, J., Prins, J.M., Stobberingh, E.E. and Geerlings, S.E. 2012. Predictive value of *Escherichia coli* susceptibility in strains causing asymptomatic bacteriuria for women with recurrent symptomatic urinary tract infections receiving prophylaxis. *Clinical Microbiology and Infection*, **18**: E84–E90.
- Besser, J., Carleton, H.A., Gerner-Smidt, P., Lindsey, R.L. and Trees, E. 2018. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin Microbiol Infect*, **24**: 335–341.
- Billips, B.K., Schaeffer, A.J. and Klumpp, D.J. 2008. Molecular Basis of Uropathogenic *Escherichia coli* Evasion of the Innate Immune Response in the Bladder. *Infect Immun*, **76**: 3891–3900.
- Blango, M.G. and Mulvey, M.A. 2010. Persistence of uropathogenic *Escherichia coli* in the face of multiple antibiotics. *Antimicrob Agents Chemother*, **54**: 1855–1863.
- Blango, M.G., Ott, E.M., Erman, A., Veranic, P. and Mulvey, M.A. 2014. Forced Resurgence and Targeting of Intracellular Uropathogenic *Escherichia coli* Reservoirs. *PLoS ONE*, **9**: e93327.
- Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**: 1453–1462.
- Blount, Z.D. 2015. The unexhausted potential of *E. coli*. *eLife*, **4**: e05826.
- Bolla, S.R., Odeluga, N. and Jetti, R. 2022. Histology, Bladder. In: *StatPearls*.
- Bonacorsi, S.P., Clermont, O., Tinsley, C., Le Gall, I., Beaudoin, J.C., Elion, J., Nassif, X. and Bingen, E. 2000. Identification of regions of the *Escherichia coli* chromosome specific for neonatal meningitis-associated strains. *Infect Immun*, **68**: 2096–2101.

- Borges, S., Silva, J. and Teixeira, P. 2014. The role of lactobacilli and probiotics in maintaining vaginal health. *Arch Gynecol Obstet*, **289**: 479–489.
- Bossa, L., Kline, K., McDougald, D., Lee, B.B. and Rice, S.A. 2017. Urinary catheter-associated microbiota change in accordance with treatment and infection status. *PLoS ONE*, **12**: e0177633.
- Bouatra, S., Aziat, F., Mandal, R., Guo, A.C., Wilson, M.R., Knox, C., Bjorndahl, T.C., Krishnamurthy, R., Saleem, F., Liu, P., Dame, Z.T., Poelzer, J., Huynh, J., Yallou, F.S., Psychogios, N., Dong, E., Bogumil, R., Roehring, C. and Wishart, D.S. 2013. The Human Urine Metabolome. *PLoS ONE*, **8**: e73076.
- Boyd, E.F. and Hartl, D.L. 1998. Chromosomal Regions Specific to Pathogenic Isolates of *Escherichia coli* Have a Phylogenetically Clustered Distribution. *J. BACTERIOL.*, **180**: 7.
- Brannon, J.R., Dunigan, T.L., Beebout, C.J., Ross, T., Wiebe, M.A., Reynolds, W.S. and Hadjifrangiskou, M. 2020. Invasion of vaginal epithelial cells by uropathogenic *Escherichia coli*. *Nat Commun*, **11**: 2803.
- Brooks, T. and Keevil, C.W. 1997. A simple artificial urine for the growth of urinary pathogens. *Lett Appl Microbiol*, **24**: 203–206.
- Bučević Popović, V., Šitum, M., Chow, C.-E.T., Chan, L.S., Roje, B. and Terzić, J. 2018. The urinary microbiome associated with bladder cancer. *Sci Rep*, **8**: 12157.
- Buckles, E.L., Wang, X., Lane, M.C., Lockett, C.V., Johnson, D.E., Rasko, D.A., Mobley, H.L.T. and Donnenberg, M.S. 2009. Role of the K2 capsule in *Escherichia coli* urinary tract infection and serum resistance. *J Infect Dis*, **199**: 1689–1697.
- Burns, S.M. and Hull, S.I. 1999. Loss of resistance to ingestion and phagocytic killing by O(-) and K(-) mutants of a uropathogenic *Escherichia coli* O75:K5 strain. *Infect Immun*, **67**: 3757–3762.
- Caron, F. 2014, Traitement des infections à *E. coli* multi-résistants en 2014, RICAI 2014
- Caron, F., Galperine, T., Fleteau, C., Azria, R., Bonacorsi, S., Bruyère, F., Cariou, G., Clouqueur, E., Cohen, R., Doco-Lecompte, T., Elefant, E., Faure, K., Gauzit, R., Gavazzi, G., Lemaitre, L., Raymond, J., Senneville, E., Sotto, A., Subtil, D., *et al.* 2018. Practice guidelines for the management of adult community-acquired urinary tract infections. *Médecine et Maladies Infectieuses*, **48**: 327–358.
- Casadevall, A. and Pirofski, L.A. 2000. Host-pathogen interactions: basic concepts of microbial commensalism, colonization, infection, and disease. *Infect Immun*, **68**: 6511–6518.
- Chan, Y., Sandlin, S., Kurzrock, E. and Osborn, S. 2017. The Current Use of Stem Cells in Bladder Tissue Regeneration and Bioengineering. *Biomedicines*, **5**: 4.
- Chaudhuri, R.R. and Henderson, I.R. 2012. The evolution of the *Escherichia coli* phylogeny. *Infection, Genetics and Evolution*, **12**: 214–226.
- Chu, E.K., Kilic, O., Cho, H., Groisman, A. and Levchenko, A. 2018. Self-induced mechanical stress can trigger biofilm formation in uropathogenic *Escherichia coli*. *Nat Commun*, **9**: 4087.
- Clermont, O., Bonacorsi, S. and Bingen, E. 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol*, **66**: 4555–4558.

- Clermont, O., Christenson, J.K., Denamur, E. and Gordon, D.M. 2013. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep*, **5**: 58–65.
- Croxen, M.A. and Finlay, B.B. 2010. Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol*, **8**: 26–38.
- Denamur, E., Clermont, O., Bonacorsi, S. and Gordon, D. 2021. The population genetics of pathogenic *Escherichia coli*. *Nat Rev Microbiol*, **19**: 37–54.
- Desvaux, M., Dalmasso, G., Beyrouthy, R., Barnich, N., Delmas, J. and Bonnet, R. 2020. Pathogenicity Factors of Genomic Islands in Intestinal and Extraintestinal *Escherichia coli*. *Front. Microbiol.*, **11**: 2065.
- Deurenberg, R.H., Bathoorn, E., Chlebowicz, M.A., Couto, N., Ferdous, M., García-Cobos, S., Kooistra-Smid, A.M.D., Raangs, E.C., Rosema, S., Veloo, A.C.M., Zhou, K., Friedrich, A.W. and Rossen, J.W.A. 2017. Application of next generation sequencing in clinical microbiology and infection prevention. *Journal of Biotechnology*, **243**: 16–24.
- Dhakal, B.K., Kulesus, R.R. and Mulvey, M.A. 2008. Mechanisms and consequences of bladder cell invasion by uropathogenic *Escherichia coli*. *Eur J Clin Invest*, **38 Suppl 2**: 2–11.
- Dhakal, B.K. and Mulvey, M.A. 2012. The UPEC pore-forming toxin α -hemolysin triggers proteolysis of host proteins to disrupt cell adhesion, inflammatory, and survival pathways. *Cell Host Microbe*, **11**: 58–69.
- Dias, R.C.S., Moreira, B.M. and Riley, L.W. 2010. Use of *fimH* single-nucleotide polymorphisms for strain typing of clinical isolates of *Escherichia coli* for epidemiologic investigation. *J Clin Microbiol*, **48**: 483–488.
- Domingos, M. de O., da Silva Junior, S.M., Milanello, W., Nakano, S.S.N., Franzolin, M.R., Dos Santos, L.F., Nunes, K.O., Marques, V.D., Elias, W.P., Silva, H.G. de S., Caetano, B.D.L. and Piazza, R.M.F. 2022. *Escherichia coli* Strains Responsible for Cystitis in Female Pediatric Patients with Normal and Abnormal Urinary Tracts Have Different Virulence Profiles. *Pathogens*, **11**: 231.
- Donlan, R.M. 2001. Biofilm Formation: A Clinically Relevant Microbiological Process. *CLIN INFECT DIS*, **33**: 1387–1392.
- Eberly, A.R., Floyd, K.A., Beebout, C.J., Colling, S.J., Fitzgerald, M.J., Stratton, C.W., Schmitz, J.E. and Hadjifrangiskou, M. 2017. Biofilm Formation by Uropathogenic *Escherichia coli* Is Favored under Oxygen Conditions That Mimic the Bladder Environment. *Int J Mol Sci*, **18**: E2077.
- European Centre for Disease prevention and Control. 2013. Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals 2011-2012. Stockholm, available at <https://ecdc.europa.eu>
- Ejrnaes, K., Sandvang, D., Lundgren, B., Ferry, S., Holm, S., Monsen, T., Lundholm, R. and Frimodt-Møller, N. 2006. Pulsed-Field Gel Electrophoresis Typing of *Escherichia coli* Strains from Samples Collected before and after Pivmecillinam or Placebo Treatment of Uncomplicated Community-Acquired Urinary Tract Infection in Women. *J Clin Microbiol*, **44**: 1776–1781.

- Ejrnæs, K., Stegger, M., Reisner, A., Ferry, S., Mønsen, T., Holm, S.E., Lundgren, B. and Frimodt-Møller, N. 2011. Characteristics of *Escherichia coli* causing persistence or relapse of urinary tract infections: Phylogenetic groups, virulence factors and biofilm formation. *Virulence*, **2**: 528–537.
- Escherich, Th. 1988. The Intestinal Bacteria of the Neonate and Breast-Fed Infant. *Clinical Infectious Diseases*, **10**: 1220–1225.
- Escobar-P, P., Vaury, C., Denamur, E. and Lecomte, G. 2004. Decreasing the effects of horizontal gene transfer on bacterial phylogeny: the *Escherichia coli* case study. *Molecular Phylogenetics and Evolution*, **8**.
- Eto, D.S., Jones, T.A., Sundsbak, J.L. and Mulvey, M.A. 2007. Integrin-Mediated Host Cell Invasion by Type 1–Piliated Uropathogenic *Escherichia coli*. *PLoS Pathog*, **3**: e100.
- Eto, D.S., Sundsbak, J.L. and Mulvey, M.A. 2006. Actin-gated intracellular growth and resurgence of uropathogenic *Escherichia coli*. *Cell Microbiol*, **8**: 704–717.
- Flores-Mireles, A.L., Walker, J.N., Caparon, M. and Hultgren, S.J. 2015. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*, **13**: 269–284.
- Forde, B.M., Roberts, L.W., Phan, M.-D., Peters, K.M., Fleming, B.A., Russell, C.W., Lenherr, S.M., Myers, J.B., Barker, A.P., Fisher, M.A., Chong, T.-M., Yin, W.-F., Chan, K.-G., Schembri, M.A., Mulvey, M.A. and Beatson, S.A. 2019. Population dynamics of an *Escherichia coli* ST131 lineage during recurrent urinary tract infection. *Nat Commun*, **10**: 3643.
- Forsyth, V.S., Armbruster, C.E., Smith, S.N., Pirani, A., Springman, A.C., Walters, M.S., Nielubowicz, G.R., Himpsl, S.D., Snitkin, E.S. and Mobley, H.L.T. 2018. Rapid Growth of Uropathogenic *Escherichia coli* during Human Urinary Tract Infection. *mBio*, **9**: e00186-18.
- Foster-Nyarko, E. and Pallen, M.J. 2022. The microbial ecology of *Escherichia coli* in the vertebrate gut. *FEMS Microbiol Rev*, **46**: fuac008.
- Foxman, B. 2002. Epidemiology of Urinary Tract Infections: Incidence, Morbidity, and Economic Costs. **9**.
- Foxman, B. 2010. The epidemiology of urinary tract infection. *Nat Rev Urol*, **7**: 653–660.
- Foxman, B. and Brown, P. 2003. Epidemiology of urinary tract infections. *Infectious Disease Clinics of North America*, **17**: 227–241.
- Fratamico, P.M., DebRoy, C., Liu, Y., Needleman, D.S., Baranzoni, G.M. and Feng, P. 2016. Advances in Molecular Serotyping and Subtyping of *Escherichia coli*. *Front. Microbiol.*, **7**.
- Fukushi, Y., Orikasa, S. and Kagayama, M. 1979. An electron microscopic study of the interaction between vesical epithelium and *E. coli*. *Invest Urol*, **17**: 61–68.
- García-Meniño, I., Lumberras, P., Lestón, L., Álvarez-Álvarez, M., García, V., Hammerl, J.A., Fernández, J. and Mora, A. 2022. Occurrence and Genomic Characterization of Clone ST1193 Clonotype 14-64 in Uncomplicated Urinary Tract Infections Caused by *Escherichia coli* in Spain. *Microbiol Spectr*, **10**: e0004122.

- Garretto, A., Miller-Ensminger, T., Ene, A., Merchant, Z., Shah, A., Gerodias, A., Biancofiori, A., Canchola, Stacey, Canchola, Stephanie, Castillo, E., Chowdhury, T., Gandhi, N., Hamilton, S., Hatton, K., Hyder, S., Krull, K., Lagios, D., Lam, T., Mitchell, K., *et al.* 2020. Genomic Survey of *E. coli* From the Bladders of Women With and Without Lower Urinary Tract Symptoms. *Front Microbiol*, **11**: 2094.
- Geerlings, S.E. 2016. Clinical Presentations and Epidemiology of Urinary Tract Infections. *Microbiol Spectr*, **4**: 4.5.03.
- Geibel, S. and Waksman, G. 2014. The molecular dissection of the chaperone–usher pathway. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, **1843**: 1559–1567.
- Gilbert, N.M., O’Brien, V.P. and Lewis, A.L. 2017. Transient microbiota exposures activate dormant *Escherichia coli* infection in the bladder and drive severe outcomes of recurrent disease. *PLoS Pathog*, **13**: e1006238.
- Glover, M., Moreira, C.G., Sperandio, V. and Zimmern, P. 2014. Recurrent urinary tract infections in healthy and nonpregnant women. *Urol Sci*, **25**: 1–8.
- Goodwin, S. 2016. Coming of age: ten years of next-generation sequencing technologies. 19.
- Gordon, D.M., Clermont, O., Tolley, H. and Denamur, E. 2008. Assigning *Escherichia coli* strains to phylogenetic groups: multi-locus sequence typing versus the PCR triplex method. *Environ Microbiol*, **10**: 2484–2496.
- Grad, Y.H., Lipsitch, M., Feldgarden, M., Arachchi, H.M., Cerqueira, G.C., FitzGerald, M., Godfrey, P., Haas, B.J., Murphy, C.I., Russ, C., Sykes, S., Walker, B.J., Wortman, J.R., Young, S., Zeng, Q., Abouelleil, A., Bochicchio, J., Chauvin, S., DeSmet, T., *et al.* 2012. Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. *Proc. Natl. Acad. Sci. U.S.A.*, **109**: 3065–3070.
- Gupta, K., Hooton, T.M., Naber, K.G., Wullt, B., Colgan, R., Miller, L.G., Moran, G.J., Nicolle, L.E., Raz, R., Schaeffer, A.J. and Soper, D.E. 2011. International Clinical Practice Guidelines for the Treatment of Acute Uncomplicated Cystitis and Pyelonephritis in Women: A 2010 Update by the Infectious Diseases Society of America and the European Society for Microbiology and Infectious Diseases. *Clinical Infectious Diseases*, **52**: e103–e120.
- Gupta, K., Stapleton, A.E., Hooton, T.M., Roberts, P.L., Fennell, C.L. and Stamm, W.E. 1998. Inverse Association of H₂O₂-Producing Lactobacilli and Vaginal *Escherichia coli* Colonization in Women with Recurrent Urinary Tract Infections. *Journal of Infectious Diseases*, **178**: 446–450.
- Henderson, J.P., Crowley, J.R., Pinkner, J.S., Walker, J.N., Tsukayama, P., Stamm, W.E., Hooton, T.M. and Hultgren, S.J. 2009. Quantitative metabolomics reveals an epigenetic blueprint for iron acquisition in uropathogenic *Escherichia coli*. *PLoS Pathog*, **5**: e1000305.
- Henry, R. 2015. Etymologia: *Escherichia coli*. *Emerg. Infect. Dis.*, **21**: 1310–1310.
- Herzer, P.J., Inouye, S., Inouye, M. and Whittam, T.S. 1990. Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J Bacteriol*, **172**: 6175–6181.

- Hibbing, M.E., Dodson, K.W., Kalas, V., Chen, S.L. and Hultgren, S.J. 2020. Adaptation of Arginine Synthesis among Uropathogenic Branches of the *Escherichia coli* Phylogeny Reveals Adjustment to the Urinary Tract Habitat. *mBio*, **11**: e02318-20.
- Hickling, D.R., Sun, T.-T. and Wu, X.-R. 2015. Anatomy and Physiology of the Urinary Tract: Relation to Host Defense and Microbial Infection. 17.
- Hilbert, D.W., Pascal, K.E., Libby, E.K., Mordechai, E., Adelson, M.E. and Trama, J.P. 2008. Uropathogenic *Escherichia coli* dominantly suppress the innate immune response of bladder epithelial cells by a lipopolysaccharide- and Toll-like receptor 4-independent pathway. *Microbes Infect*, **10**: 114–121.
- Hilt, E.E., McKinley, K., Pearce, M.M., Rosenfeld, A.B., Zilliox, M.J., Mueller, E.R., Brubaker, L., Gai, X., Wolfe, A.J. and Schreckenberger, P.C. 2014. Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. *J Clin Microbiol*, **52**: 871–876.
- Hisano, M., Bruschini, H., Nicodemo, A.C., Gomes, C.M., Lucon, M. and Srougi, M. 2015. The Bacterial Spectrum and Antimicrobial Susceptibility in Female Recurrent Urinary Tract Infection: How Different They Are From Sporadic Single Episodes? *Urology*, **86**: 492–497.
- Hogins, J., Fan, E., Seyan, Z., Kusin, S., Christie, A.L., Zimmern, P.E. and Reitzer, L. 2022. Bacterial Growth of Uropathogenic *Escherichia coli* in Pooled Urine Is Much Higher than Predicted from the Average Growth in Individual Urine Samples. *Microbiol Spectr*, e02016-22.
- Hooton, T.M. 2001. Recurrent urinary tract infection in women. *International Journal of Antimicrobial Agents*, **10**.
- Hooton, T.M., Stapleton, A.E., Roberts, P.L., Winter, C., Scholes, D., Bavendam, T. and Stamm, W.E. 1999. Perineal anatomy and urine-voiding characteristics of young women with and without recurrent urinary tract infections. *Clin Infect Dis*, **29**: 1600–1601.
- Huang, L., Li, X., Zheng, B., Li, P., Wei, D., Huang, C., Sun, L. and Li, H. 2022. Differential Urinary Microbiota Composition Between Women With and Without Recurrent Urinary Tract Infection. *Front. Microbiol.*, **13**: 888681.
- Hudson, P.L., Hung, K.J., Bergerat, A. and Mitchell, C. 2020. Effect of Vaginal *Lactobacillus* Species on *Escherichia coli* Growth. *Female Pelvic Med Reconstr Surg*, **26**: 146–151.
- Ikram, R., Psutka, R., Carter, A. and Priest, P. 2015. An outbreak of multi-drug resistant *Escherichia coli* urinary tract infection in an elderly population: a case-control study of risk factors. *BMC Infect Dis*, **15**: 224.
- Ipe, D.S., Horton, E. and Ulett, G.C. 2016. The Basics of Bacteriuria: Strategies of Microbes for Persistence in Urine. *Front. Cell. Infect. Microbiol.*, **6**.
- Jaureguy, F., Landraud, L., Passet, V., Diancourt, L., Frapy, E., Guigon, G., Carbonnelle, E., Lortholary, O., Clermont, O., Denamur, E., Picard, B., Nassif, X. and Brisse, S. 2008. Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics*, **9**: 560.
- Johnson, J.R. and Russo, T.A. 2018. Molecular Epidemiology of Extraintestinal Pathogenic *Escherichia coli*. *EcoSal Plus*, **8**: ecosalplus.ESP-0004-2017.

- Justice, S.S., Hung, C., Theriot, J.A., Fletcher, D.A., Anderson, G.G., Footer, M.J. and Hultgren, S.J. 2004. Differentiation and developmental pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. *Proc Natl Acad Sci U S A*, **101**: 1333–1338.
- Justice, S.S., Hunstad, D.A., Seed, P.C. and Hultgren, S.J. 2006. Filamentation by *Escherichia coli* subverts innate defenses during urinary tract infection. *Proc. Natl. Acad. Sci. U.S.A.*, **103**: 19884–19889.
- Kahlmeter, G., Brown, D.F.J., Goldstein, F.W., MacGowan, A.P., Mouton, J.W., Osterlund, A., Rodloff, A., Steinbakk, M., Urbaskova, P. and Vatopoulos, A. 2003. European harmonization of MIC breakpoints for antimicrobial susceptibility testing of bacteria. *J Antimicrob Chemother*, **52**: 145–148.
- Kai-Larsen, Y., Lüthje, P., Chromek, M., Peters, V., Wang, X., Holm, Å., Kádas, L., Hedlund, K.-O., Johansson, J., Chapman, M.R., Jacobson, S.H., Römling, U., Agerberth, B. and Brauner, A. 2010. Uropathogenic *Escherichia coli* Modulates Immune Responses and Its Curli Fimbriae Interact with the Antimicrobial Peptide LL-37. *PLoS Pathog*, **6**: e1001010.
- Kaper, J.B., Nataro, J.P. and Mobley, H.L.T. 2004. Pathogenic *Escherichia coli*. *Nat Rev Microbiol*, **2**: 123–140.
- Kauffmann, F. 1947. The serology of the coli group. *J Immunol*, **57**: 71–100.
- Klein, R.D. and Hultgren, S.J. 2020. Urinary tract infections: microbial pathogenesis, host–pathogen interactions and new treatment strategies. *Nat Rev Microbiol*, **18**: 211–226.
- Komesu, Y.M., Dinwiddie, D.L., Richter, H.E., Lukacz, E.S., Sung, V.W., Siddiqui, N.Y., Zyczynski, H.M., Ridgeway, B., Rogers, R.G., Arya, L.A., Mazloomdoost, D., Levy, J., Carper, B., Gantz, M.G., and Eunice Kennedy Shriver National Institute of Child Health and Human Development Pelvic Floor Disorders Network. 2020. Defining the relationship between vaginal and urinary microbiomes. *Am J Obstet Gynecol*, **222**: 154.e1-154.e10.
- Kudinha, T. and Kong, F. 2022. Distribution of papG alleles among uropathogenic *Escherichia coli* from reproductive age women. *J Biomed Sci*, **29**: 66.
- Lebeaux, D., Ghigo, J.-M. and Beloin, C. 2014. Biofilm-related infections: bridging the gap between clinical management and fundamental aspects of recalcitrance toward antibiotics. *Microbiol Mol Biol Rev*, **78**: 510–543.
- Lecointre, G., Rachdi, L., Darlu, P. and Denamur, E. 1998. *Escherichia coli* molecular phylogeny using the incongruence length difference test. *Mol Biol Evol*, **15**: 1685–1695.
- Lee, G., Cho, Y.-H., Shim, B.S. and Lee, S.D. 2010. Risk Factors for Antimicrobial Resistance Among the *Escherichia coli* Strains Isolated from Korean Patients with Acute Uncomplicated Cystitis: A Prospective and Nationwide Study. *J Korean Med Sci*, **25**: 1205.
- Lee, J., Hiibel, S.R., Reardon, K.F. and Wood, T.K. 2010. Identification of stress-related proteins in *Escherichia coli* using the pollutant cis-dichloroethylene. *J Appl Microbiol*, **108**: 2088–2102.
- Lescat, M., Clermont, O., Woerther, P.L., Glodt, J., Dion, S., Skurnik, D., Djossou, F., Dupont, C., Perroz, G., Picard, B., Catzeflis, F., Andremont, A. and Denamur, E. 2013. Commensal *Escherichia coli* strains in Guiana reveal a high genetic diversity with host-dependant population structure. *Environ Microbiol Rep*, **5**: 49–57.

- Lewis, A.J., Richards, A.C. and Mulvey, M.A. 2016. Invasion of Host Cells and Tissues by Uropathogenic Bacteria. *Microbiol Spectr*, **4**: 4.6.38.
- Lewis, A.L. and Gilbert, N.M. 2020. Roles of the vagina and the vaginal microbiota in urinary tract infection: evidence from clinical correlations and experimental models. *GMS Infect Dis*, **8**: Doc02.
- Lindblom, A., Kiszakiewicz, C., Kristiansson, E., Yazdanshenas, S., Kamenska, N., Karami, N. and Åhrén, C. 2022. The impact of the ST131 clone on recurrent ESBL-producing *E. coli* urinary tract infection: a prospective comparative study. *Sci Rep*, **12**: 10048.
- Liu, S., Han, X., Shi, M. and Pang, Z. 2016. Persistence of uropathogenic *Escherichia coli* in the bladders of female patients with sterile urine after antibiotic therapies. *J. Huazhong Univ. Sci. Technol. [Med. Sci.]*, **36**: 710–715.
- Lopez, S.R. and Mangır, N. 2021. Current standard of care in treatment of bladder pain syndrome/interstitial cystitis. *Therapeutic Advances in Urology*, **13**: 175628722110224.
- Lu, J., Peng, Y., Wan, S., Frost, L.S., Raivio, T. and Glover, J.N.M. 2019. Cooperative Function of TraJ and ArcA in Regulating the F Plasmid *tra* Operon. *J Bacteriol*, **201**.
- Lund, B., Lindberg, F., Marklund, B.-I. and Normark, S. 1987. The PapG protein is the alpha-D-galactopyranosyl-(1-4)-beta-D-galactopyranose-binding adhesin of uropathogenic *Escherichia coli*. *Medical Sciences*, **84**: 5.
- Luo, C., Walk, S.T., Gordon, D.M., Feldgarden, M., Tiedje, J.M. and Konstantinidis, K.T. 2011. Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. *Proc Natl Acad Sci U S A*, **108**: 7200–7205.
- Lüthje, P. and Brauner, A. 2014. Virulence Factors of Uropathogenic *E. coli* and Their Interaction with the Host. In: *Advances in Microbial Physiology*, pp. 337–372.
- Magruder, M., Sholi, A.N., Gong, C., Zhang, L., Edusei, E., Huang, J., Albakry, S., Satlin, M.J., Westblade, L.F., Crawford, C., Dadhania, D.M., Lubetzky, M., Taur, Y., Littman, E., Ling, L., Burnham, P., De Vlaminc, I., Pamer, E., Suthanthiran, M., *et al.* 2019. Gut uropathogen abundance is a risk factor for development of bacteriuria and urinary tract infection. *Nat Commun*, **10**: 5521.
- Maiden, M.C.J., Jansen van Rensburg, M.J., Bray, J.E., Earle, S.G., Ford, S.A., Jolley, K.A. and McCarthy, N.D. 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol*, **11**: 728–736.
- Manges, A.R., Geum, H.M., Guo, A., Edens, T.J., Fibke, C.D. and Pitout, J.D.D. 2019. Global Extraintestinal Pathogenic *Escherichia coli* (ExPEC) Lineages. *Clin Microbiol Rev*, **32**: e00135-18.
- Martinez, J.J. 2000. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *The EMBO Journal*, **19**: 2803–2812.
- Martinez, J.J. and Hultgren, S.J. 2002. Requirement of Rho-family GTPases in the invasion of Type 1-piliated uropathogenic *Escherichia coli*. *Cell Microbiol*, **4**: 19–28.
- Matuszewski, M.A., Tupikowski, K., Dołowy, Ł., Szymańska, B., Dembowski, J. and Zdrojowy, R. 2016. Uroplakins and their potential applications in urology. *Cent European J Urol*, **69**: 252–257.

- Mctaggart, L.A., Rigby, R.C. and Elliott, T.S.J. 1990. The pathogenesis of urinary tract infections associated with *Escherichia coli*, *Staphylococcus saprophyticus* and *S. epidermidis*. *Journal of Medical Microbiology*, **32**: 135–141.
- Meštrović, T., Matijašić, M., Perić, M., Čipčić Paljetak, H., Barešić, A. and Verbanac, D. 2020. The Role of Gut, Vaginal, and Urinary Microbiome in Urinary Tract Infections: From Bench to Bedside. *Diagnostics*, **11**: 7.
- Miao, Y., Li, G., Zhang, X., Xu, H. and Abraham, S.N. 2015. A TRP Channel Senses Lysosome Neutralization by Pathogens to Trigger Their Expulsion. *Cell*, **161**: 1306–1319.
- Moissenet, D., Salauze, B., Clermont, O., Bingen, E., Arlet, G., Denamur, E., Mérens, A., Mitanchez, D. and Vu-Thien, H. 2010. Meningitis caused by *Escherichia coli* producing TEM-52 extended-spectrum beta-lactamase within an extensive outbreak in a neonatal ward: epidemiological investigation and characterization of the strain. *J Clin Microbiol*, **48**: 2459–2463.
- Moussatova, A., Kandt, C., O'Mara, M.L. and Tieleman, D.P. 2008. ATP-binding cassette transporters in *Escherichia coli*. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **1778**: 1757–1771.
- Mueller, N.T., Bakacs, E., Combellick, J., Grigoryan, Z. and Dominguez-Bello, M.G. 2015. The infant microbiome development: mom matters. *Trends Mol Med*, **21**: 109–117.
- Mulvey, M.A., Lopez-Boado, Y.S., Wilson, C.L., Roth, R., Parks, W.C., Heuser, J. and Hultgren, S.J. 1998. Induction and Evasion of Host Defenses by Type 1-Piliated Uropathogenic *Escherichia coli*. *Science*, **282**: 1494–1497.
- Mulvey, M.A., Schilling, J.D. and Hultgren, S.J. 2001. Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect Immun*, **69**: 4572–4579.
- Mysorekar, I.U. and Hultgren, S.J. 2006. Mechanisms of uropathogenic *Escherichia coli* persistence and eradication from the urinary tract. *Proc Natl Acad Sci U S A*, **103**: 14170–14175.
- Nakamura, A., Takahashi, H., Arai, M., Tsuchiya, T., Wada, S., Fujimoto, Y., Shimabara, Y., Kuda, T. and Kimura, B. 2021. Molecular subtyping for source tracking of *Escherichia coli* using core genome multilocus sequence typing at a food manufacturing plant. *PLoS One*, **16**: e0261352.
- Nicolas-Chanoine, M.-H., Bertrand, X. and Madec, J.-Y. 2014. *Escherichia coli* ST131, an intriguing clonal group. *Clin Microbiol Rev*, **27**: 543–574.
- Nielsen, D.W., Klimavicz, J.S., Cavender, T., Wannemuehler, Y., Barbieri, N.L., Nolan, L.K. and Logue, C.M. 2018. The Impact of Media, Phylogenetic Classification, and *E. coli* Pathotypes on Biofilm Formation in Extraintestinal and Commensal *E. coli* From Humans and Animals. *Front Microbiol*, **9**: 902.
- Nielsen, K.L., Stegger, M., Kiil, K., Lilje, B., Ejrnæs, K., Leihof, R.F., Skjøt-Rasmussen, L., Godfrey, P., Mønsen, T., Ferry, S., Hammerum, A.M. and Frimodt-Møller, N. 2021. *Escherichia coli* Causing Recurrent Urinary Tract Infections: Comparison to Non-Recurrent Isolates and Genomic Adaptation in Recurrent Infections. *Microorganisms*, **9**: 1416.
- Oelschlaeger, T.A., Dobrindt, U. and Hacker, J. 2002. Pathogenicity islands of uropathogenic *E. coli* and the evolution of virulence. *International Journal of Antimicrobial Agents*, **5**.

- ONERBA (Observatoire National de l'épidémiologie de la Résistance Bactérienne aux Antibiotiques) : rapport d'activité 2015–2016. n.d.<http://www.onerba.org/spip.php?article137> (published online, Nov 2016).
- Oren, A. and Garrity, G.M. 2021. Valid publication of the names of forty-two phyla of prokaryotes. *Int J Syst Evol Microbiol*, **71**.
- Pang, Y., Cheng, Z., Zhang, S., Li, S., Li, Xueping, Li, Xiaodan, Zhang, X., Li, Xiaoxiao, Feng, Y., Cui, H., Chen, Z., Liu, L., Li, Q., Huang, J., Zhang, M., Zhu, S., Wang, L. and Feng, L. 2022. Bladder epithelial cell phosphate transporter inhibition protects mice against uropathogenic *Escherichia coli* infection. *Cell Rep*, **39**: 110698.
- Pearce, M.M., Hilt, E.E., Rosenfeld, A.B., Zilliox, M.J., Thomas-White, K., Fok, C., Kliethermes, S., Schreckenberger, P.C., Brubaker, L., Gai, X. and Wolfe, A.J. 2014. The female urinary microbiome: a comparison of women with and without urgency urinary incontinence. *mBio*, **5**: e01283-01214.
- Pellatt, G.C. 2007. Anatomy and physiology of urinary elimination. Part 1. *British Journal of Nursing*, **16**: 5.
- Perez-Carrasco, V., Soriano-Lerma, A., Soriano, M., Gutiérrez-Fernández, J. and Garcia-Salcedo, J.A. 2021. Urinary Microbiome: Yin and Yang of the Urinary Tract. *Front Cell Infect Microbiol*, **11**: 617002.
- Petronella, N., Kundra, P., Auclair, O., Hébert, K., Rao, M., Kingsley, K., De Bruyne, K., Banerjee, S., Gill, A., Pagotto, F., Tamber, S. and Ronholm, J. 2019. Changes detected in the genome sequences of *Escherichia coli*, *Listeria monocytogenes*, *Vibrio parahaemolyticus*, and *Salmonella enterica* after serial subculturing. *Can. J. Microbiol.*, **65**: 842–850.
- Pitout, J.D.D., Peirano, G., Chen, L., DeVinney, R. and Matsumura, Y. 2022. *Escherichia coli* ST1193: Following in the Footsteps of *E. coli* ST131. *Antimicrob Agents Chemother*, **66**: e0051122.
- Poirel, L., Madec, J.-Y., Lupo, A., Schink, A.-K., Kieffer, N., Nordmann, P. and Schwarz, S. 2018. Antimicrobial Resistance in *Escherichia coli*. *Microbiol Spectr*, **6**.
- Rangel, D.E., Marín-Medina, N., Castro, J.E., González-Mancera, A. and Forero-Shelton, M. 2013. Observation of Bacterial Type I Pili Extension and Contraction under Fluid Flow. *PLoS ONE*, **8**: e65563.
- Rasko, D.A., Rosovitz, M.J., Myers, G.S.A., Mongodin, E.F., Fricke, W.F., Gajer, P., Crabtree, J., Sebaihia, M., Thomson, N.R., Chaudhuri, R., Henderson, I.R., Sperandio, V. and Ravel, J. 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J Bacteriol*, **190**: 6881–6893.
- Raz, R. 2011. Urinary tract infection in postmenopausal women. *Korean J Urol*, **52**: 801–808.
- Raz, R., Gennesin, Y., Wasser, J., Stoler, Z., Rosenfeld, S., Rottensterich, E. and Stamm, W.E. 2000. Recurrent Urinary Tract Infections in Postmenopausal Women. *Clinical Infectious Diseases*, **30**: 152–156.
- Reid, S.D., Herbelin, C.J., Bumbaugh, A.C., Selander, R.K. and Whittam, T.S. 2000. Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature*, **406**: 64–67.

- Reis, N.M., Alves, I., Pereira, F., Jegouic, S. and Edwards, A.D. 2021. 14 - Point-of-need detection with smartphone. In: *Smartphone-Based Detection Devices* (C. Hussain, ed), pp. 311–362.
- Reitzer, L. and Zimmern, P. 2019. Rapid Growth and Metabolism of Uropathogenic *Escherichia coli* in Relation to Urine Composition. *Clin Microbiol Rev*, **33**: e00101-19.
- Renard, J., Ballarini, S., Mascarenhas, T., Zahran, M., Quimper, E., Choucair, J. and Iselin, C.E. 2014. Recurrent Lower Urinary Tract Infections Have a Detrimental Effect on Patient Quality of Life: a Prospective, Observational Study. *Infect Dis Ther*, doi: 10.1007/s40121-014-0054-6.
- Riley, L.W. 2020. Distinguishing Pathovars from Nonpathovars: *Escherichia coli*. *Microbiol Spectr*, **8**: 8.4.1.
- Robino, L., Scavone, P., Araujo, L., Algorta, G., Zunino, P. and Vignoli, R. 2013. Detection of intracellular bacterial communities in a child with *Escherichia coli* recurrent urinary tract infections. *Pathog Dis*, **68**: 78–81.
- Robins-Browne, R.M., Holt, K.E., Ingle, D.J., Hocking, D.M., Yang, J. and Tauschek, M. 2016. Are *Escherichia coli* Pathotypes Still Relevant in the Era of Whole-Genome Sequencing? *Front Cell Infect Microbiol*, **6**: 141.
- Rosen, D.A., Hooton, T.M., Stamm, W.E., Humphrey, P.A. and Hultgren, S.J. 2007. Detection of Intracellular Bacterial Communities in Human Urinary Tract Infection. *PLoS Med*, **4**: e329.
- Russo, T. 2003. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection*, **5**: 449–456.
- Sanchez, G.V., Babiker, A., Master, R.N., Luu, T., Mathur, A. and Bordon, J. 2016. Antibiotic Resistance among Urinary Isolates from Female Outpatients in the United States in 2003 and 2012. *Antimicrob Agents Chemother*, **60**: 2680–2683.
- Sande, C. and Whitfield, C. 2021. Capsules and Extracellular Polysaccharides in *Escherichia coli* and *Salmonella*. *EcoSal Plus*, **9**: eESP-0033-2020.
- Santé publique France, 2017. Enquête nationale de prévalence des infections nosocomiales et des traitements anti-infectieux en établissements de santé, France, mai-juin 2017. Disponible à partir de l'URL : www.santepubliquefrance.fr
- Santé publique France, 2020. Surveillance des toxi-infection alimentaires collectives. Données de la déclaration obligatoire de 2020. Disponible à partir de l'URL : www.santepubliquefrance.fr
- Santos, A.C. de M., Santos, F.F., Silva, R.M. and Gomes, T.A.T. 2020. Diversity of Hybrid- and Hetero-Pathogenic *Escherichia coli* and Their Potential Implication in More Severe Diseases. *Front Cell Infect Microbiol*, **10**: 339.
- Scholes, D., Hawn, T.R., Roberts, P.L., Li, S.S., Stapleton, A.E., Zhao, L.-P., Stamm, W.E. and Hooton, T.M. 2010. Family history and risk of recurrent cystitis and pyelonephritis in women. *J Urol*, **184**: 564–569.
- Schwartz, D.J., Chen, S.L., Hultgren, S.J. and Seed, P.C. 2011. Population dynamics and niche distribution of uropathogenic *Escherichia coli* during acute and chronic urinary tract infection. *Infect Immun*, **79**: 4250–4259.

- Scott, V.C.S., Haake, D.A., Churchill, B.M., Justice, S.S. and Kim, J.-H. 2015. Intracellular Bacterial Communities: A Potential Etiology for Chronic Lower Urinary Tract Symptoms. *Urology*, **86**: 425–431.
- Sender, R., Fuchs, S. and Milo, R. 2016. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol*, **14**: e1002533.
- Sharma, K., Dhar, N., Thacker, V.V., Simonet, T.M., Signorino-Gelo, F., Knott, G.W. and McKinney, J.D. 2021. Dynamic persistence of UPEC intracellular bacterial communities in a human bladder-chip model of urinary tract infection. *eLife*, **10**: e66481.
- Silva, M.T. 2012. Classical labeling of bacterial pathogens according to their lifestyle in the host: inconsistencies and alternatives. *Front Microbiol*, **3**: 71.
- Skjøt-Rasmussen, L., Hammerum, A.M., Jakobsen, L., Lester, C.H., Larsen, P. and Frimodt-Møller, N. 2011. Persisting clones of *Escherichia coli* isolates from recurrent urinary tract infection in men and women. *Journal of Medical Microbiology*, **60**: 550–554.
- Snyder, J.A., Haugen, B.J., Lockett, C.V., Maroncle, N., Hagan, E.C., Johnson, D.E., Welch, R.A. and Mobley, H.L.T. 2005. Coordinate expression of fimbriae in uropathogenic *Escherichia coli*. *Infect Immun*, **73**: 7588–7596.
- Soto, S.M., Smithson, A., Horcajada, J.P., Martinez, J.A., Mensa, J.P. and Vila, J. 2006. Implication of biofilm formation in the persistence of urinary tract infection caused by uropathogenic *Escherichia coli*. *Clinical Microbiology and Infection*, **12**: 1034–1036.
- Spaulding, C.N., Schreiber, H.L., Zheng, W., Dodson, K.W., Hazen, J.E., Conover, M.S., Wang, F., Svenmarker, P., Luna-Rico, A., Francetic, O., Andersson, M., Hultgren, S. and Egelman, E.H. 2018. Functional role of the type 1 pilus rod structure in mediating host-pathogen interactions. *eLife*, **7**: e31662.
- Spurbeck, R.R., Stapleton, A.E., Johnson, J.R., Walk, S.T., Hooton, T.M. and Mobley, H.L.T. 2011. Fimbrial profiles predict virulence of uropathogenic *Escherichia coli* strains: contribution of *ygi* and *yad* fimbriae. *Infect Immun*, **79**: 4753–4763.
- Stanley, T. and Wilson, I.G. 2003. Multilocus Enzyme Electrophoresis: A Practical Guide. *MB*, **24**: 203–220.
- Stapleton, A.E., Au-Yeung, M., Hooton, T.M., Fredricks, D.N., Roberts, P.L., Czaja, C.A., Yarova-Yarovaya, Y., Fiedler, T., Cox, M. and Stamm, W.E. 2011. Randomized, Placebo-Controlled Phase 2 Trial of a *Lactobacillus crispatus* Probiotic Given Intravaginally for Prevention of Recurrent Urinary Tract Infection. *Clinical Infectious Diseases*, **52**: 1212–1217.
- Stoffler, D., Stolz, M., Schoenenberger, C.-A., Huang, N. and Aebi, U. 2005. Imaging, measuring, and manipulating biological matter across dimensions. *Nanobiotechnol*, **1**: 263–266.
- Subashchandrabose, S. and Mobley, H.L.T. 2015. Virulence and Fitness Determinants of Uropathogenic *Escherichia coli*. *Microbiol Spectr*, **3**.
- Sumati, A.H. and Saritha, N.K. 2009. Association of urinary tract infection in women with bacterial vaginosis. *J Glob Infect Dis*, **1**: 151–152.

- Suzuki, K., Koyanagi-Aoi, M., Uehara, K., Hinata, N., Fujisawa, M. and Aoi, T. 2019. Directed differentiation of human induced pluripotent stem cells into mature stratified bladder urothelium. *Sci Rep*, **9**: 10506.
- Tandogdu, Z. and Wagenlehner, F.M.E. 2016. Global epidemiology of urinary tract infections. *Urinary tract infections*, **29**: 7.
- Tariq, R., Pardi, D.S., Tosh, P.K., Walker, R.C., Razonable, R.R. and Khanna, S. 2017. Fecal Microbiota Transplantation for Recurrent *Clostridium difficile* Infection Reduces Recurrent Urinary Tract Infection Frequency. *Clinical Infectious Diseases*, **65**: 1745–1747.
- Tchesnokova, V.L., Ottley, L.L., Sakamoto, K., Fierer, J., Sokurenko, E. and Liss, M.A. 2015. Rapid Identification of Rectal Multidrug-resistant *Escherichia coli* Before Transrectal Prostate Biopsy. *Urology*, **86**: 1200–1205.
- Tenaillon, O., Skurnik, D., Picard, B. and Denamur, E. 2010. The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol*, **8**: 207–217.
- Thänert, R., Choi, J., Reske, K.A., Hink, T., Thänert, A., Wallace, M.A., Wang, B., Seiler, S., Cass, C., Bost, M.H., Struttman, E.L., Iqbal, Z.H., Sax, S.R., Fraser, V.J., Baker, A.W., Foy, K.R., Williams, B., Xu, B., Capocci-Tolomeo, P., *et al.* 2022. Persisting uropathogenic *Escherichia coli* lineages show signatures of niche-specific within-host adaptation mediated by mobile genetic elements. *Cell Host & Microbe*, **30**: 1034-1047.e6.
- Thänert, R., Reske, K.A., Hink, T., Wallace, M.A., Wang, B., Schwartz, D.J., Seiler, S., Cass, C., Burnham, C.-A.D., Dubberke, E.R., Kwon, J.H. and Dantas, G. 2019. Comparative Genomics of Antibiotic-Resistant Uropathogens Implicates Three Routes for Recurrence of Urinary Tract Infections. *mBio*, **10**: e01977-19.
- Thomas-White, K., Forster, S.C., Kumar, N., Van Kuiken, M., Putonti, C., Stares, M.D., Hilt, E.E., Price, T.K., Wolfe, A.J. and Lawley, T.D. 2018. Culturing of female bladder bacteria reveals an interconnected urogenital microbiota. *Nat Commun*, **9**: 1557.
- Thumbikat, P., Berry, R.E., Zhou, G., Billips, B.K., Yaggie, R.E., Zaichuk, T., Sun, T.-T., Schaeffer, A.J. and Klumpp, D.J. 2009. Bacteria-induced uroplakin signaling mediates bladder response to infection. *PLoS Pathog*, **5**: e1000415.
- Torres, A.G. and Payne, S.M. 1997. Haem iron-transport system in enterohaemorrhagic *Escherichia coli* O157:H7. *Mol Microbiol*, **23**: 825–833.
- Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, M.E., Frapy, E., *et al.* 2009. Organised Genome Dynamics in the *Escherichia coli* Species Results in Highly Diverse Adaptive Paths. *PLOS Genetics*, **5**: e1000344.
- Trompette, A., Gollwitzer, E.S., Yadava, K., Sichelstiel, A.K., Sprenger, N., Ngom-Bru, C., Blanchard, C., Junt, T., Nicod, L.P., Harris, N.L. and Marsland, B.J. 2014. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat Med*, **20**: 159–166.
- van Elsas, J.D., Semenov, A.V., Costa, R. and Trevors, J.T. 2011. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *ISME J*, **5**: 173–183.

- Vincent, A.T., Derome, N., Boyle, B., Culley, A.I. and Charette, S.J. 2017. Next-generation sequencing (NGS) in the microbiological world: How to make the most of your money. *Journal of Microbiological Methods*, **12**.
- Weissman, S.J., Johnson, J.R., Tchesnokova, V., Billig, M., Dykhuizen, D., Riddell, K., Rogers, P., Qin, X., Butler-Wu, S., Cookson, B.T., Fang, F.C., Scholes, D., Chattopadhyay, S. and Sokurenko, E. 2012. High-resolution two-locus clonal typing of extraintestinal pathogenic *Escherichia coli*. *Appl Environ Microbiol*, **78**: 1353–1360.
- Whiteside, S.A., Razvi, H., Dave, S., Reid, G. and Burton, J.P. 2015. The microbiome of the urinary tract—a role beyond infection. *Nat Rev Urol*, **12**: 81–90.
- Wick, R.R., Judd, L.M., Gorrie, C.L. and Holt, K.E. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol*, **13**: e1005595.
- Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L.H., Karch, H., Reeves, P.R., Maiden, M.C.J., Ochman, H. and Achtman, M. 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol*, **60**: 1136–1151.
- Wolfe, A.J., Toh, E., Shibata, N., Rong, R., Kenton, K., FitzGerald, M., Mueller, E.R., Schreckenberger, P., Dong, Q., Nelson, D.E. and Brubaker, L. 2012. Evidence of Uncultivated Bacteria in the Adult Female Bladder. *J Clin Microbiol*, **50**: 1376–1383.
- Worby, C.J., Olson, B.S., Dodson, K.W., Earl, A.M. and Hultgren, S.J. 2022. Establishing the role of the gut microbiota in susceptibility to recurrent urinary tract infections. *J Clin Invest*, **132**: e158497.
- Worby, C.J., Schreiber, H.L., Straub, T.J., van Dijk, L.R., Bronson, R.A., Olson, B.S., Pinkner, J.S., Obernuefemann, C.L.P., Muñoz, V.L., Paharik, A.E., Azimzadeh, P.N., Walker, B.J., Desjardins, C.A., Chou, W.-C., Bergeron, K., Chapman, S.B., Klim, A., Manson, A.L., Hannan, T.J., *et al.* 2022. Longitudinal multi-omics analyses link gut microbiome dysbiosis with recurrent urinary tract infections in women. *Nat Microbiol*, **7**: 630–639.
- Yamaji, R., Rubin, J., Thys, E., Friedman, C.R. and Riley, L.W. 2018. Persistent Pandemic Lineages of Uropathogenic *Escherichia coli* in a College Community from 1999 to 2017. *J Clin Microbiol*, **56**: e01834-17.
- Yang, Z.-K., Luo, H., Zhang, Y., Wang, B. and Gao, F. 2019. Pan-genomic analysis provides novel insights into the association of *E. coli* with human host and its minimal genome. *Bioinformatics*, **35**: 1987–1991.
- Yoo, J.-J., Shin, H., Song, J., Kim, M., Yun, J., Kim, Z., Lee, Y., Lee, S., Lee, K., Kim, W., Ryu, C., Park, S.-W., Park, S., Song, H.-Y., Kim, Y., and on the behalf of SMS (Soonchunhyang microbiome multi-disciplinary study group). 2021. Urinary Microbiome Characteristics in Female Patients with Acute Uncomplicated Cystitis and Recurrent Cystitis. *JCM*, **10**: 1097.
- Zare, M., Vehreschild, M.J.G.T. and Wagenlehner, F. 2022. Management of uncomplicated recurrent urinary tract infections. *BJU International*, **129**: 668–678.
- Zhou, Z., Alikhan, N.-F., Mohamed, K., Fan, Y., Agama Study Group and Achtman, M. 2020. The Enterobase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res*, **30**: 138–152.

Cystites récidivantes à *Escherichia coli* : Facteurs microbiens de la persistance

Nos travaux ont dans un premier temps proposé une approche rapide pour distinguer de façon fiable les rechutes (épisodes de cystites causées par une même souche) des réinfections (épisodes causés par deux souches distinctes), et évaluer la place de chacune dans le phénomène de récurrence. Pour ce faire, nous avons typé par diverses approches (antibiotype, phylogroupage, *CH typing*, *NGS*) une collection prospective de 347 isolats de *E. coli* recueillis chez une 323 patientes présentant une cystite, dont 72, atteintes de CR. Ce faisant, nous avons mis en évidence que 30,6% des patientes atteintes de CR, ont rechuté au moins une fois avec la même souche de *E. coli*. Dans la majorité des cas, la comparaison des antibiogrammes des isolats séquentiels et/ou le *CH typing* étaient suffisants pour monter l'identité de deux isolats.

À partir des 24 isolats responsables de rechutes identifiés dans la première partie de notre travail, nous avons mené une large étude génomique et phénotypique dans le but d'identifier des marqueurs prédictifs du caractère récidivant d'une souche de *E. coli* et d'étudier l'évolution des souches responsables de rechutes au cours du temps. Sur le plan génomique, nous avons observé que le génome des souches de rechutes est plus petit que celui des souches sporadiques, cette différence de taille étant supposément due à la perte d'un plasmide. Sur le plan phénotypique (fitness et biofilm), aucune différence significative n'a pu être notée entre les deux groupes, suggérant que les souches de rechute ne sont « que » des souches *E. coli* uropathogènes « ordinaires ». Enfin, les souches de rechute semblent s'adapter à leur hôte au cours du temps via l'accumulation de SNPs, principalement dans des gènes impliqués dans le métabolisme ou le transport membranaire.

Mots clés : *Escherichia coli* uropathogène, infections urinaires, antibiotype, *CH typing*, séquençage haut-débit, adaptation, fitness, biofilm, internalisation dans des cellules urothéliales

Recurrent cystitis caused by *Escherichia coli*, microbial factors associated with persistence

Our study first proposed a rapid approach to reliably distinguish relapses (episodes of cystitis caused by the same strain) from reinfections (episodes caused by two distinct strains), and to evaluate the role of each in recurrence. We typed by various approaches (antibiotype, phylogrouping, *CH typing*, *NGS*) a prospective collection of 347 *E. coli* isolates collected from 323 patients with cystitis, including 72 with RC. We demonstrated that 30.6% of patients with RC relapsed at least once with the same *E. coli* strain. In most cases, comparison of antibiograms of sequential isolates and/or *CH* types were sufficient to show the identity of two isolates.

From the 24 relapsing isolates identified in the first part of our work, we conducted a large genomic and phenotypic study in order to identify predictive markers of relapse. We also studied the evolution of relapsing strains over time. We observed that the genome of relapsing strains is smaller than that of sporadic strains, this size difference is probably due to the loss of a plasmid. Phenotypically (fitness and biofilm assays), no significant difference has been observed, suggesting that the relapsing strains are "just ordinary" uropathogenic *E. coli*. Finally, the relapsing strains seem to adapt to their host over time via the accumulation of SNPs, mainly in genes involved in metabolism or membrane transport.

Keywords: Uropathogenic *Escherichia coli*, urinary tract infections, antibiotype, *CH typing*, high-throughput sequencing, adaptation, fitness, biofilm, invasion of bladder epithelial cells