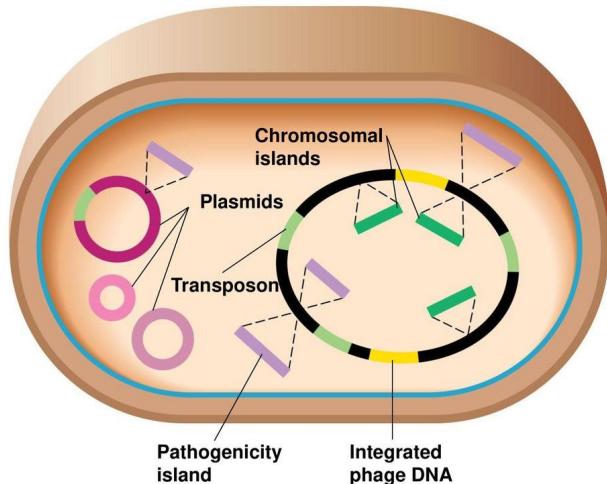


# BASES GÉNÉTIQUES DE LA VIRULENCE BACTÉRIENNE

2022-2023



**Dr Lise CREMET - MCU-PH**

Service de Bactériologie - CHU de Nantes

Faculté de Pharmacie de Nantes

CR2TI Equipe 6, IRS2 Nantes Biotech

[lise.cremet@univ-nantes.fr](mailto:lise.cremet@univ-nantes.fr)

[lise.cremet@chu-nantes.fr](mailto:lise.cremet@chu-nantes.fr)

# Notions générales de virulence



- **Bactéries saprophytes :**

Une bactérie est saprophyte lorsqu'elle vit et se nourrit dans l'environnement (sol, eaux, surfaces).

- **Bactéries commensales :**

Une bactérie est commensale lorsqu'elle vit au contact du revêtement cutanéomuqueux d'un hôte sans entraîner de désordres. Les bactéries commensales proviennent soit de l'environnement, soit d'autres hôtes.

- **Bactéries pathogènes :**

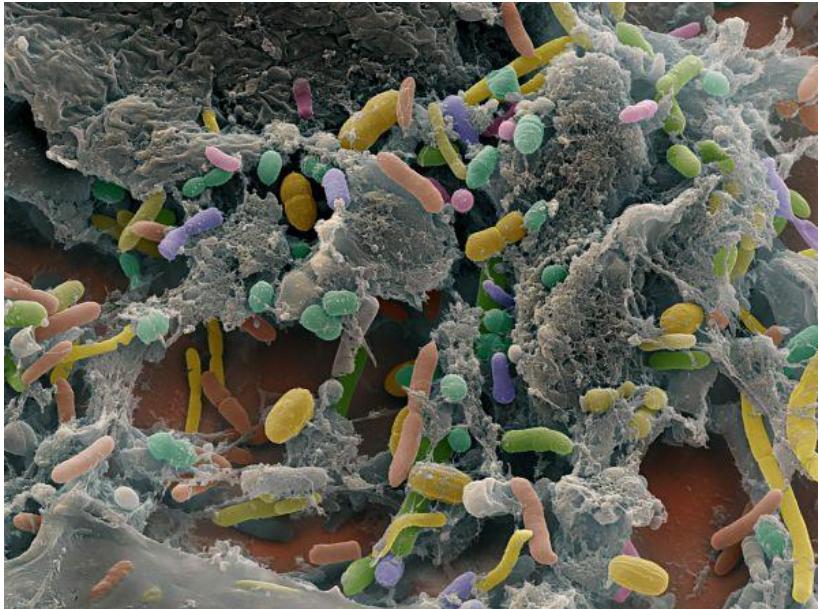
Bactéries capables de provoquer une maladie chez un sujet dont les mécanismes de défense sont normaux (ex: coqueluche, fièvre typhoïde, choléra)

- **Bactéries opportunistes :**

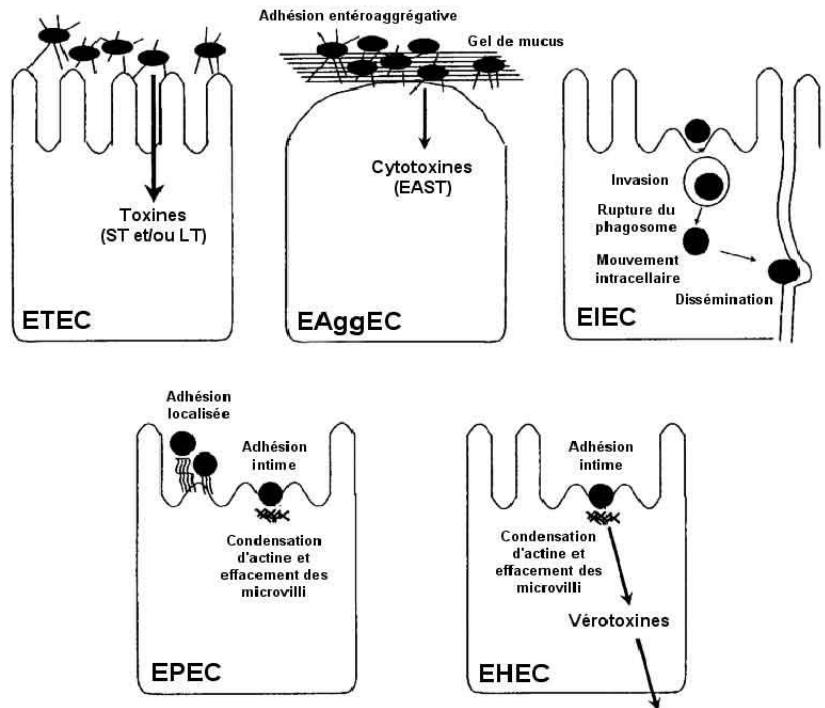
Certaines bactéries peuvent devenir pathogènes lorsque les défenses de l'hôte sont affaiblies, mais ne donnent pas habituellement de maladie chez le sujet sain (ex: staphylocoques blancs, corynébactéries)

# *E. coli* : une bactérie commensale & pathogène

- *E. coli* commensal de l'intestin :



- *E. coli* responsables de diarrhées :

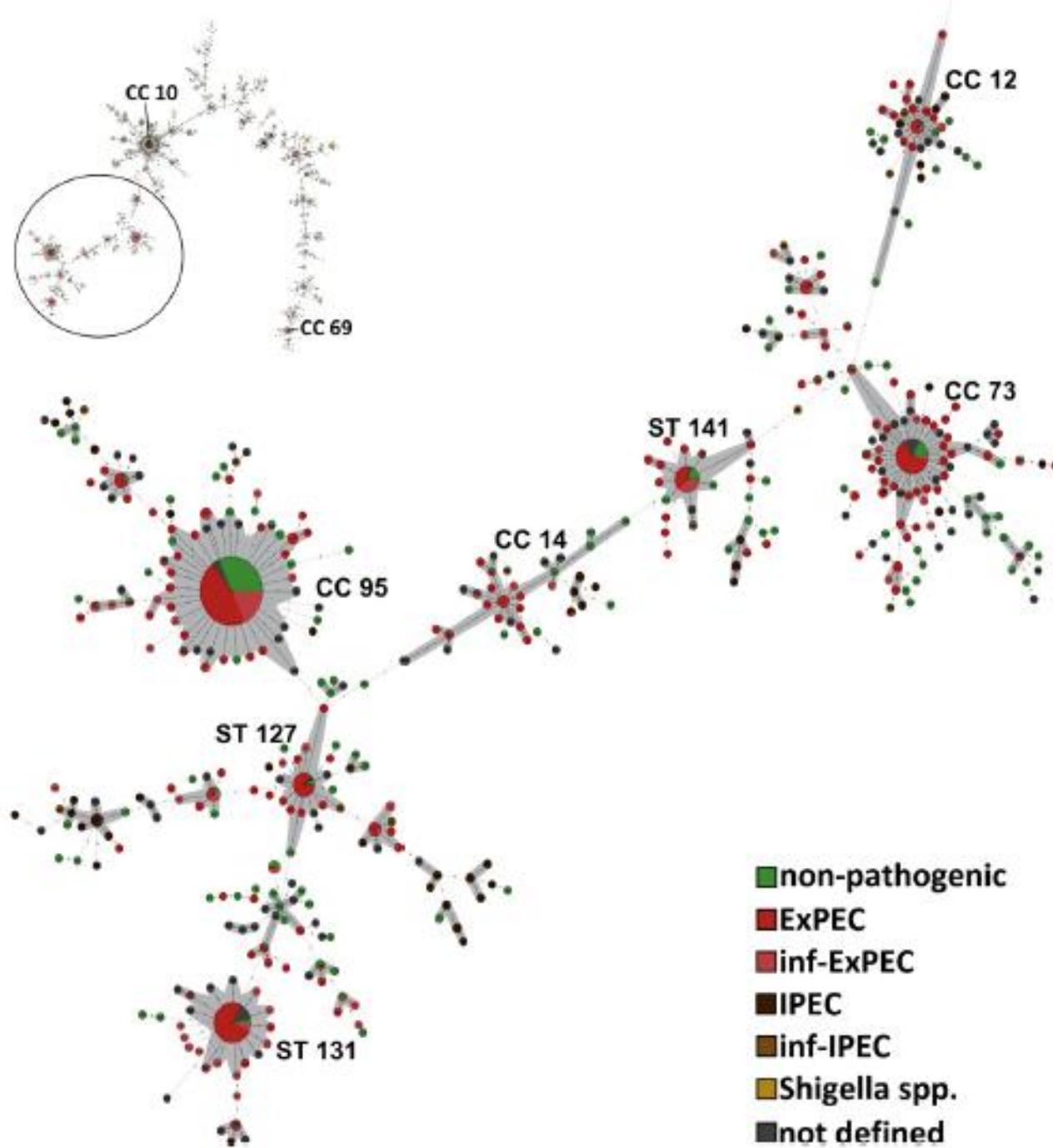


- *E. coli* responsables d'infections extra-intestinales :

Infections urinaires, bactériémies, péritonites, pneumonies nosocomiales, méningites, ...

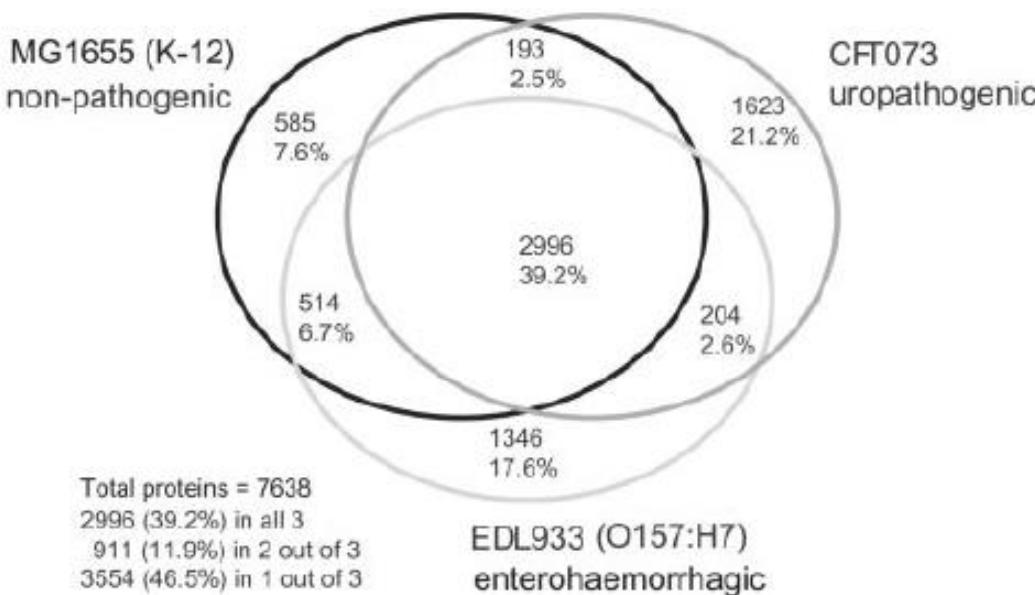
# *E. coli* : génétique de populations

- **Serotyping (Kauffmann, 1940) :**
  - Antigènes O (somatique), H (flagellaire), K (capsulaire)
  - 173 Ag O, 56 Ag H, 80 Ag K
  - Une petite partie des combinaisons O:H sont associées à des pathologies
- **Multilocus enzyme electrophoresis (MLEE) (années 80)**
  - 5 groupes phylogénétiques principaux (A, B1, B2, D, E)
- **Multilocus sequence typing (MLST) (Maiden, 1998)**
  - > 1859 sequence types (ST) et 114 complexes clonaux (CC)
  - CC et ST rassemblent des ExPEC et des souches non pathogènes
- **Phylogroupage par PCR triplex (2000)**
  - 4 groupes phylogénétiques principaux (A, B1, B2, D)
- **Séquençage à haut débit : whole-genome sequencing (2015)**

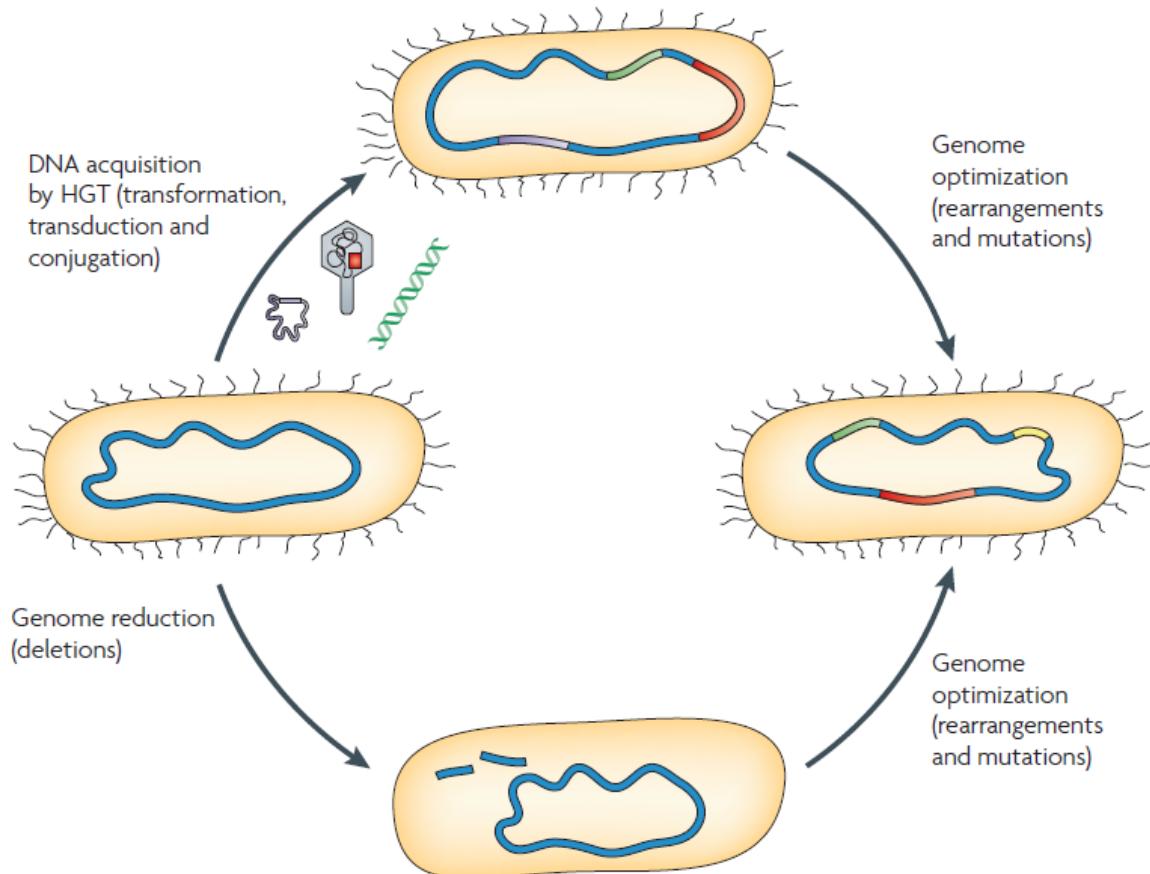


# Plasticité du génome et pathogénicité

- Plasticité = capacité d'un génome à subir des réarrangements
- Une souche pathogène de *E. coli* possède environ 20% d'information génétique supplémentaire par rapport à une souche commensale
- Un répertoire spécifique de gènes de virulence = une pathologie particulière



# Plasticité du génome et pathogénicité



1 million de pb supplémentaires dans le génome d'un *E. coli* pathogène par rapport à une souche de *E. coli* commensale

**Figure 1 | Mechanisms that contribute to bacterial genome evolution.** Genome plasticity results from DNA acquisition by horizontal gene transfer (HGT; for example, through the uptake of plasmids, phages and naked DNA) and genome reduction by DNA deletions, rearrangements and point mutations. The concerted action of DNA acquisition and gene loss results in a genome-optimization process that frequently occurs in response to certain growth conditions, including host infection or colonization.

# Le génome bactérien

- **Core génome :**

- Ensemble de gènes partagés et conservés
- Fonctions essentielles, métabolisme

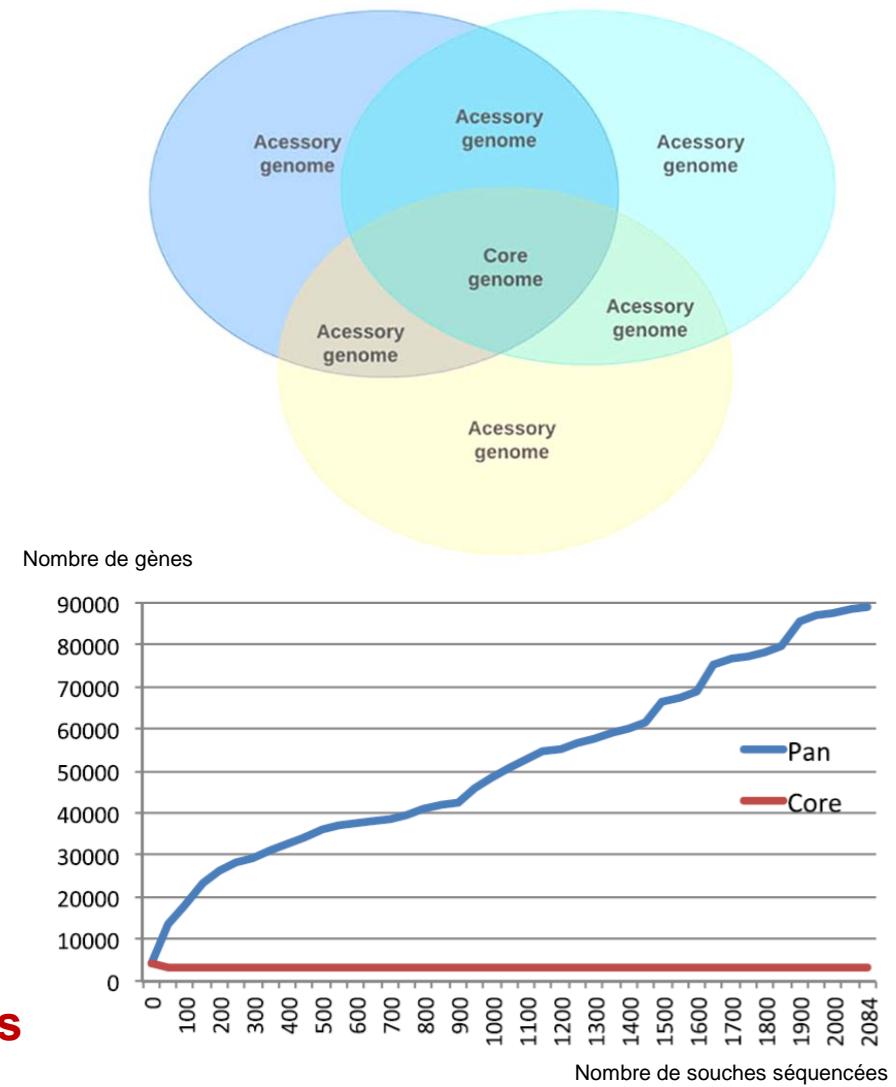
- **Gènes accessoires :**

- Fonctions non indispensables mais favorables à la survie en milieu hostile

- **Pangénomie :**

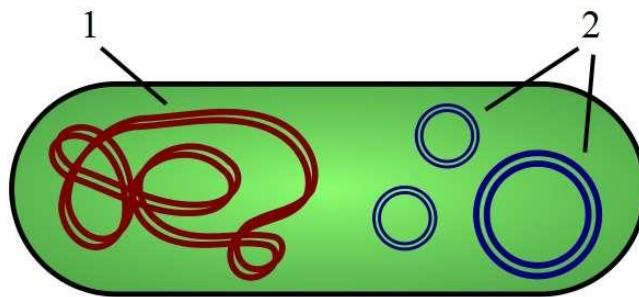
- Tous les gènes de toutes les souches d'une espèce

> Core génome de *E. coli* ≈ 3 100 gènes  
> Pangénomie de *E. coli* ≈ 89 000 gènes



# Supports génétiques de la virulence

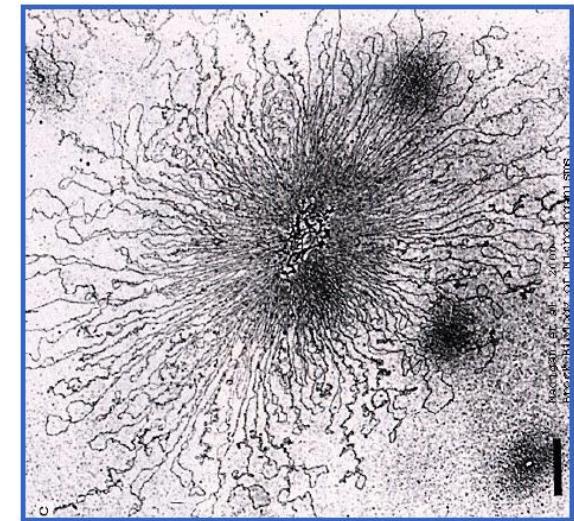
- Les bactéries sont des **organismes procaryotes unicellulaires** :
  - absence d'organites délimités par une membrane (mitochondries, golgi, RE...)
  - nucléoïde bactérien non délimité par une membrane
- Matériel génétique (ADN) :
  - chromosome ou nucléoïde (1)
  - plasmides (2)



# Le nucléoïde : Composition

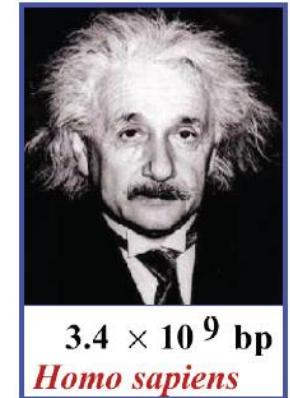
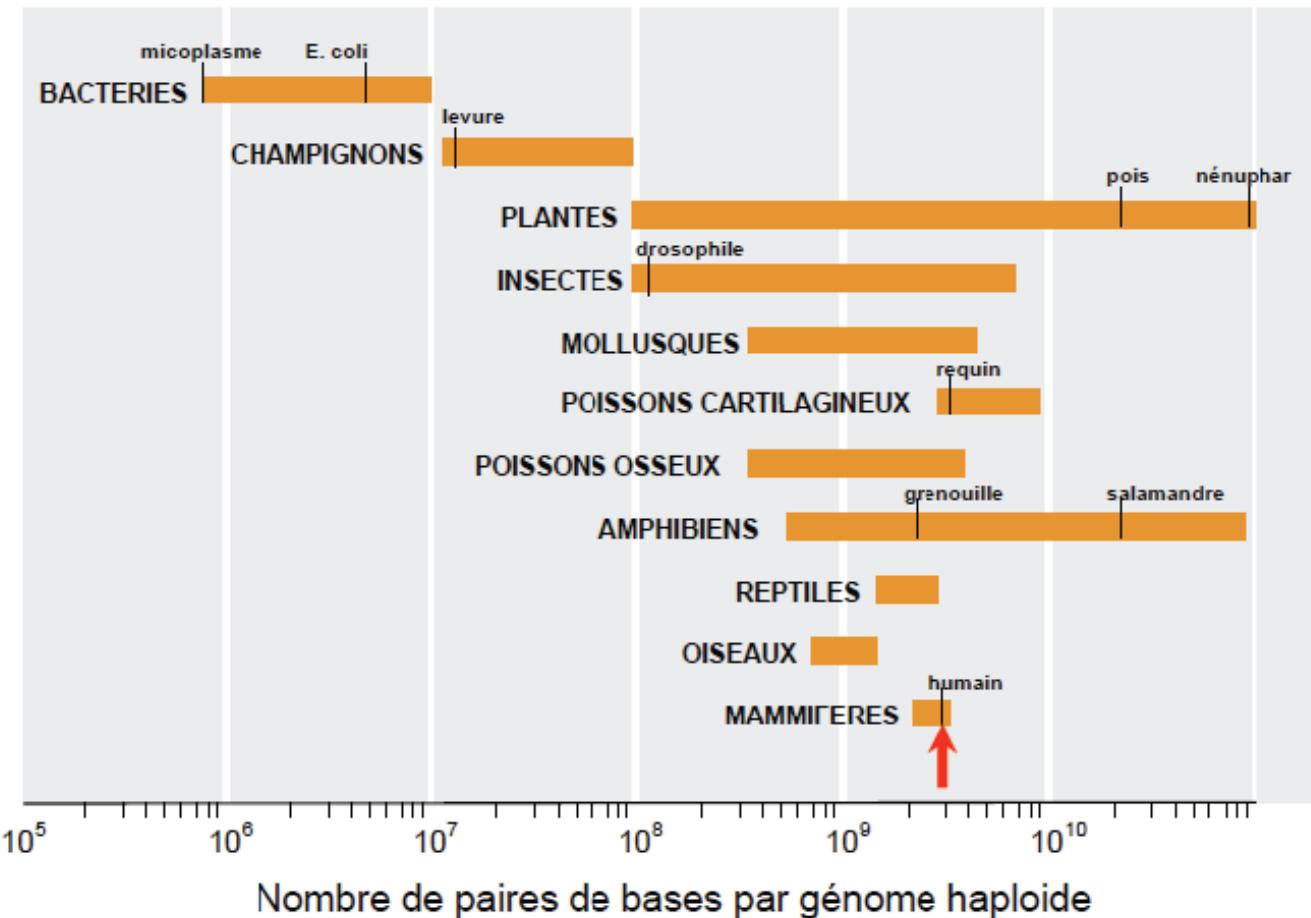
- **Composition de l'appareil nucléaire :**

- **ADN (60%)** : taille variable
- **ARN (30%)** : produits de transcription + ARN polymérase
- **Protéines (10%)** :
  - ✓ Rôle dans la structure, lors de la réplication, dans l'expression génétique
  - ✓ Chez *E. coli*, une douzaine de protéines sont impliquées dans la compaction du génome (Fis, IHF,H-NS,...)



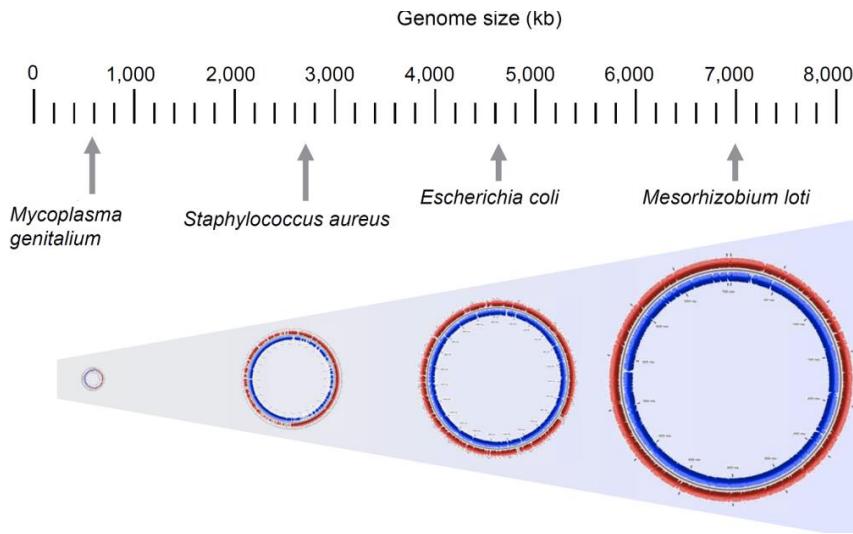
# Le nucléoïde : nombre, structure, taille

- Ordre du million de paires de bases (pb) :



# Le nucléoïde : nombre, structure, taille

- Taille variable selon les espèces
- ADN près de 1000 fois plus long que la bactérie
- Filament continu d'une double chaîne d'ADN pelotonnée



	Nombre de paires de Bases	Nombre de protéines
<i>Mycoplasma genitalium</i>	580073	467
<i>Treponema pallidum</i>	1 138 011	1031
<i>Haemophilus influenzae</i>	1 830 138	1709
<i>Neisseria meningitidis MC58</i>	2 272 325	2025
<i>Vibrio cholerae</i>	4 033 464	3827
<i>Mycobacterium tuberculosis</i>	4 411 529	3918
<i>Pseudomonas aeruginosa</i>	6 264 403	5565

1995 : Publication de la 1ère séquence complète du génome d'un organisme vivant : *Haemophilus influenzae*

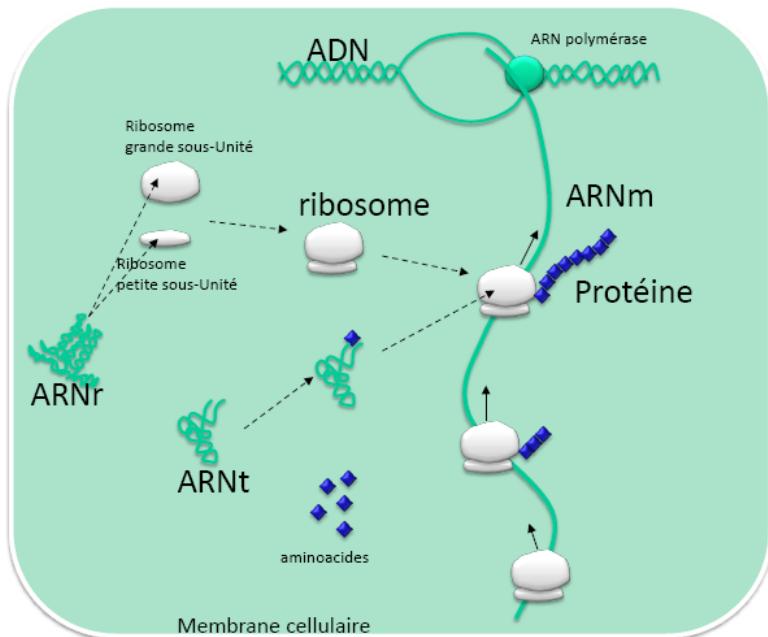
# Le nucléoïde : nombre, structure, taille

- Presque toujours **chromosome unique et circulaire**

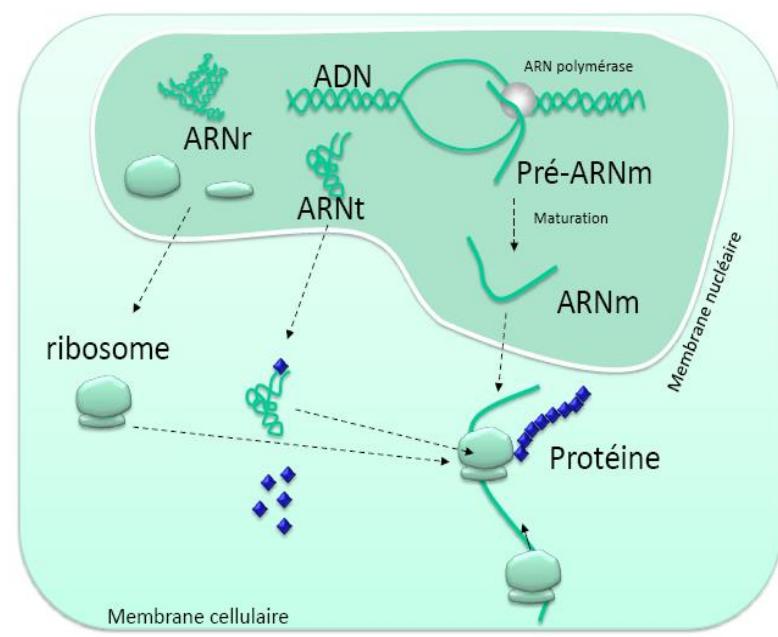
	<b>Chromosome</b> (Nombre, forme, taille)
<i>E. coli</i>	1 circulaire (4,6 Mb)
<i>P. aeruginosa</i>	1 circulaire (6,3 Mb)
<i>Vibrio cholerae</i>	2 circulaires (2,9 & 1,1 Mb)
<i>Brucella suis</i> 1	2 circulaires (2 & 1 Mb)
<i>Brucella suis</i> 3	1 circulaire (3,1 Mb)
<i>A. tumefaciens</i>	1 circulaire (3 Mb) 1 linéaire (2,1 Mb)
<i>Borrelia</i>	1 linéaire (0,9 Mb)

# La transcription chez les bactéries

- Pas de modification post-transcriptionnelle des ARNm
- Transcription et traduction ont lieu simultanément dans le cytoplasme
- Demi-vie très courte des ARNm



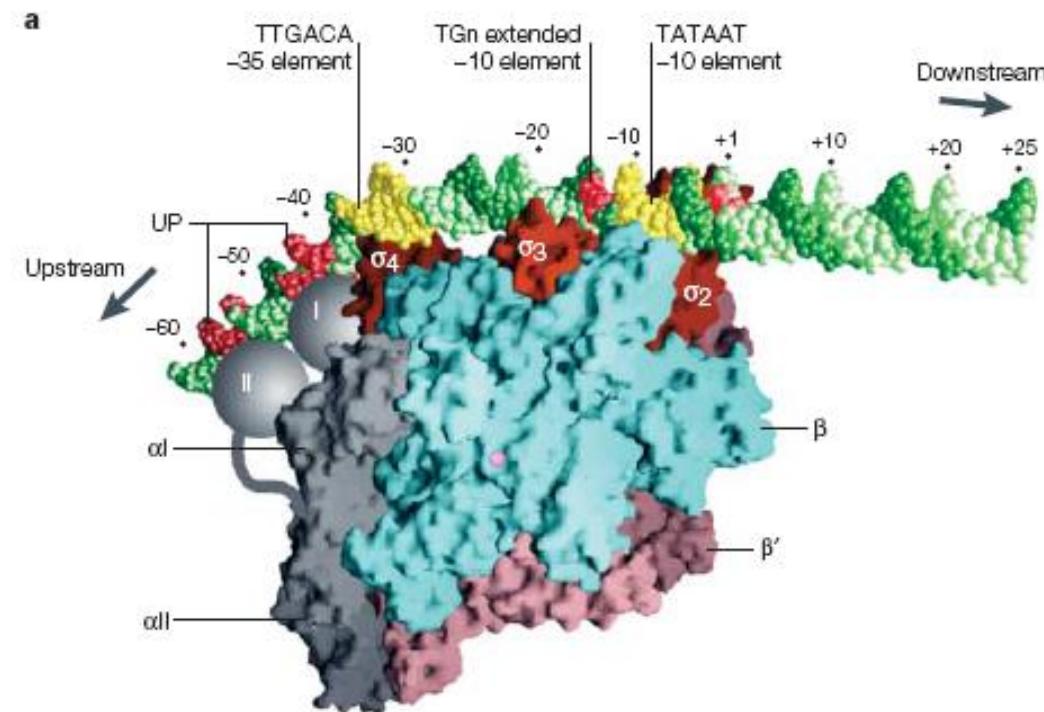
Cellule procaryote



Cellule eucaryote

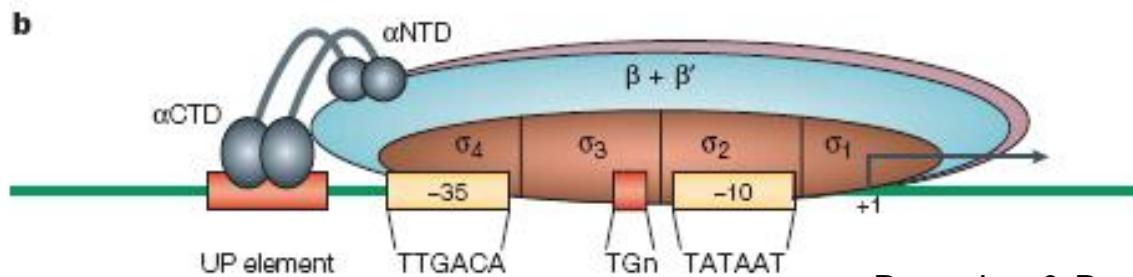
# La transcription chez les bactéries

- Assurée par une **ARN polymérase** constituée de 5 sous-unités :
  - 2 $\alpha$ ,  $\beta$  et  $\beta'$**  : affinité très élevée pour l'ADN, sans spécificité
  - facteur  $\sigma$**  : reconnaît une séquence consensus sur le promoteur des gènes



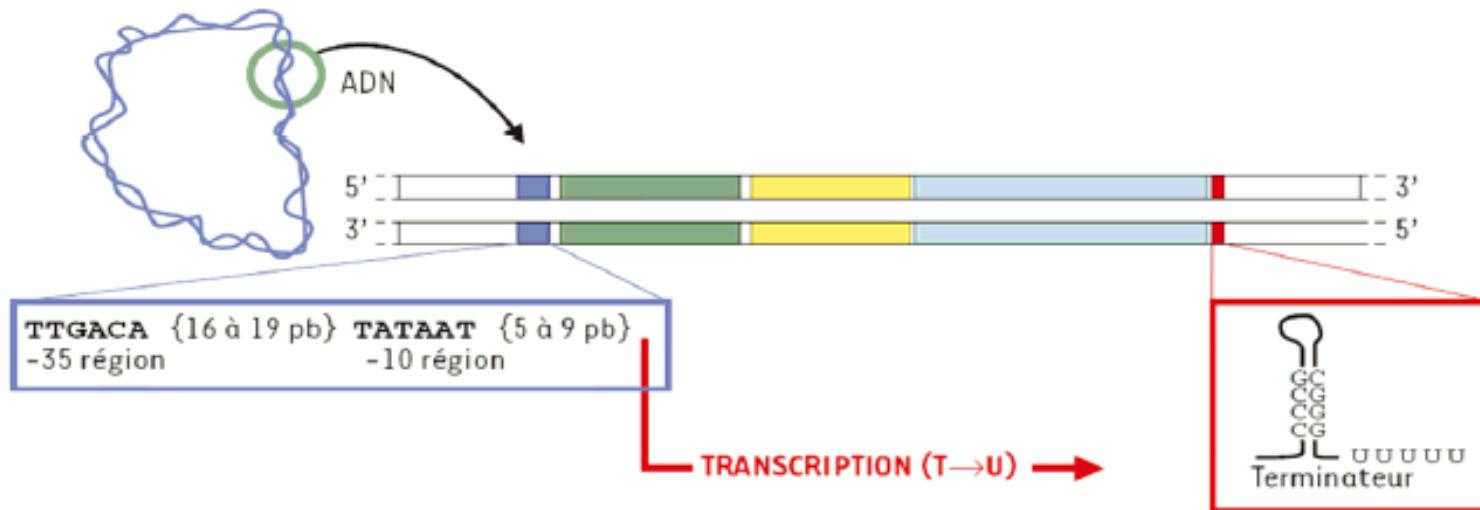
# L'ARN polymérase bactérienne

- Fixation optimale nécessite :
  - présence de séquences consensus en position -10 (boîte TATA) et -35 (boîte Pribnow)
  - espacement de 16 à 19 bases entre ces séquences
- Étroitement régulée :
  - facteurs de transcription
  - protéines régulatrices (répresseur, atténuateur...)
  - type de facteur  $\sigma$  fixé



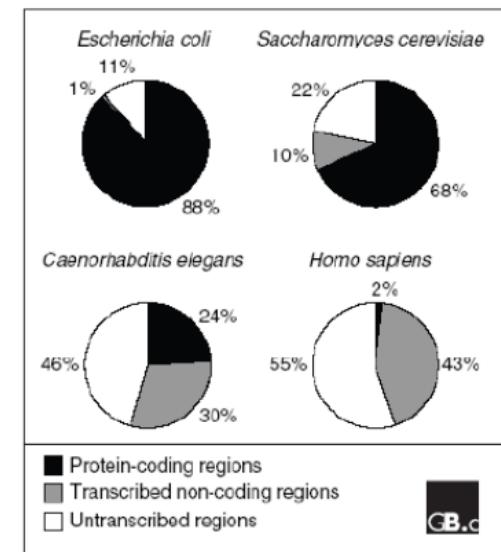
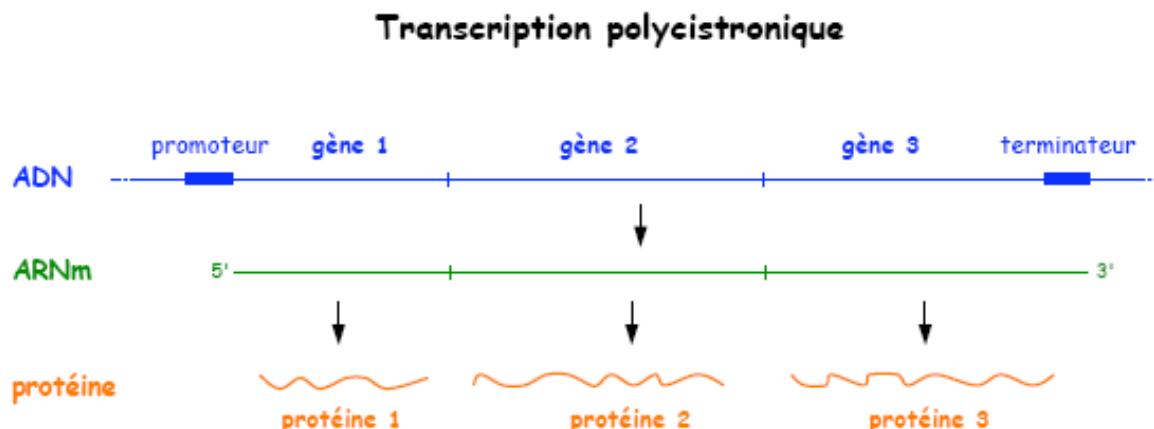
# La transcription chez les bactéries

- Formation d'un complexe ouvert où l'ADN est dénaturé
- Sigma est chassé de l'ADN et la polymérase peut alors se déplacer sur l'ADN
- L'elongation se poursuit jusqu'à ce que l'ARN forme une boucle complémentaire qui déstabilise la polymérase (site de terminaison)
- L'ARNm est relâché et la polymérase est récupérée pour un nouveau cycle transcriptionnel



# La transcription chez les bactéries

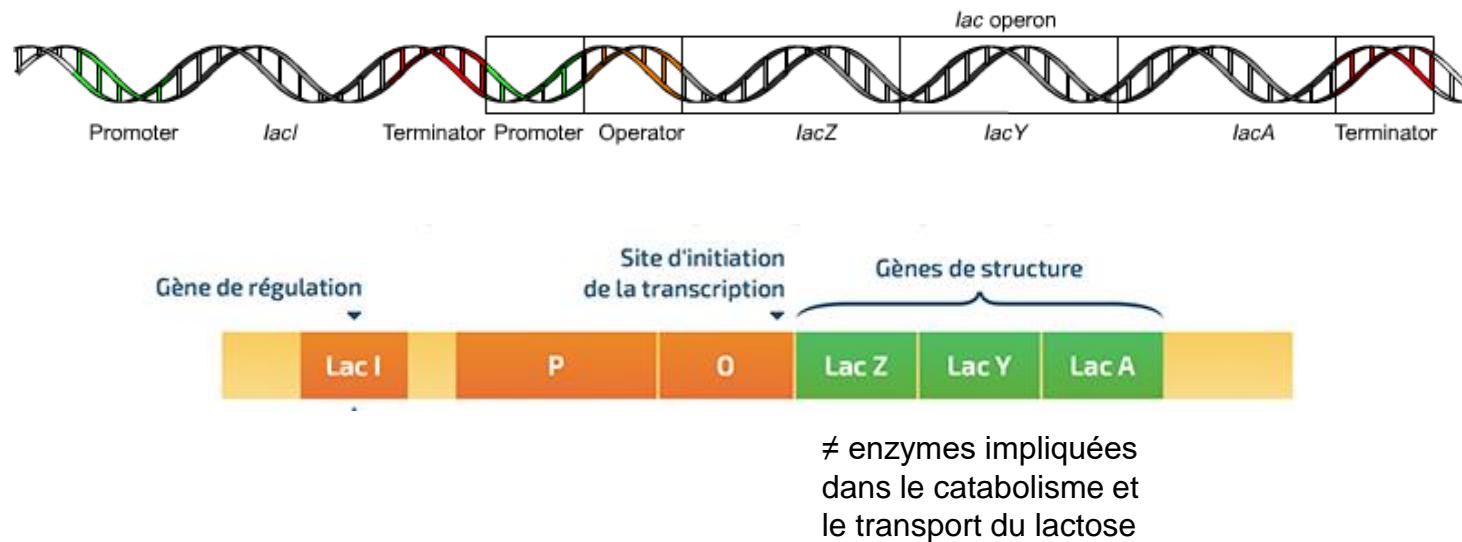
- La plus grande partie du génome des bactéries est composé de séquences codantes ( $\neq$  avec eucaryotes)
- Les gènes bactériens peuvent être organisés en **opérons** : ensemble de gènes sous la dépendance d'un promoteur unique et transcrits sous la forme d'un seul **ARNm polycistronique**



# Gènes organisés en opérons

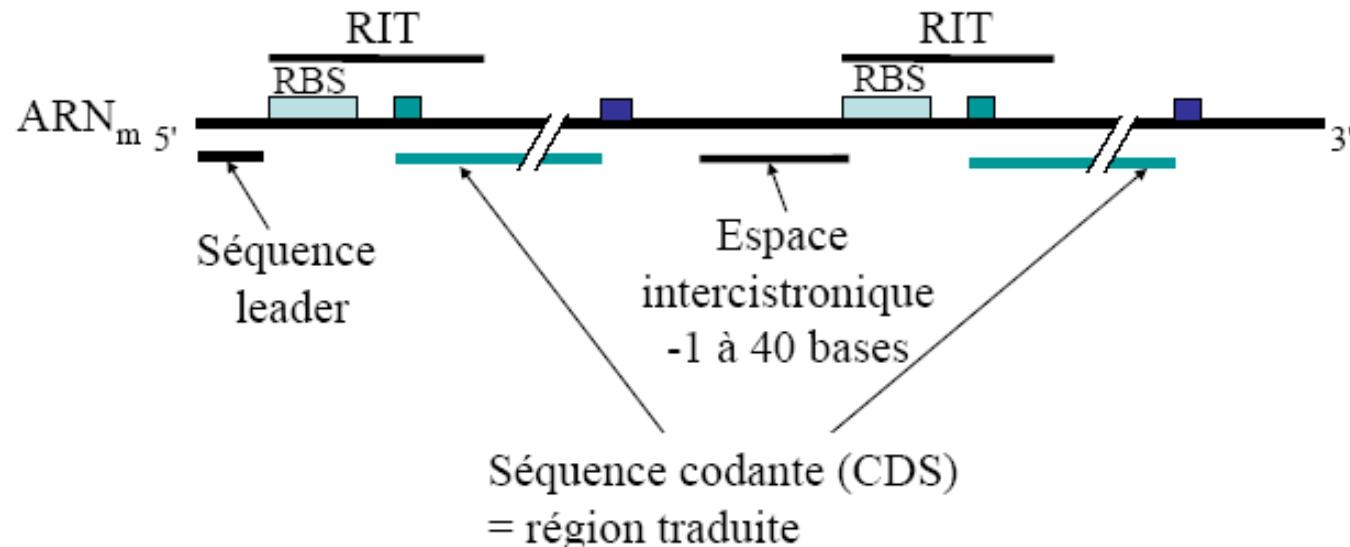
- En général, les gènes d'un même opéron codent pour des protéines d'une même voie métabolique (ex : opéron lactose, opéron tryptophane)

Exemple : opéron lactose d' *E. coli* (gènes *lacZ*, *lacY* et *lacA*)



# Gènes organisés en opérons

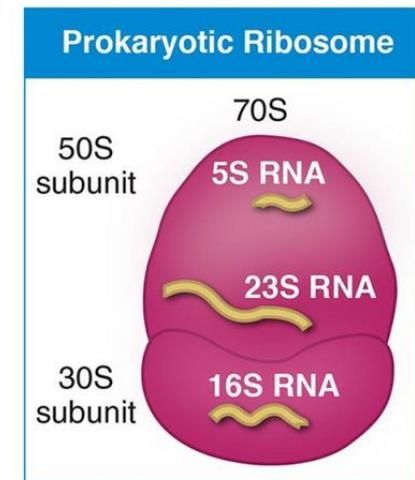
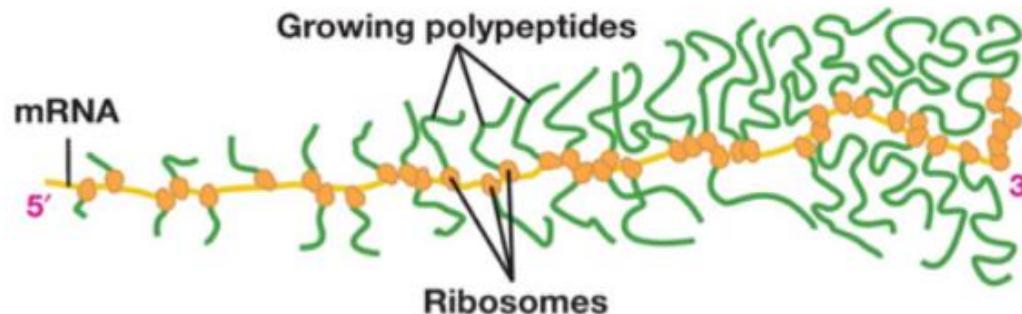
- Un site de fixation pour les ribosomes (= séquence de Shine-Dalgarno) est présent dans l'ARNm près du site de début de la traduction de chaque région codant une protéine



■ Codon d'initiation de la traduction  
■ Codon stop

RIT: région d'initiation de la traduction  
RBS: site de fixation du ribosome

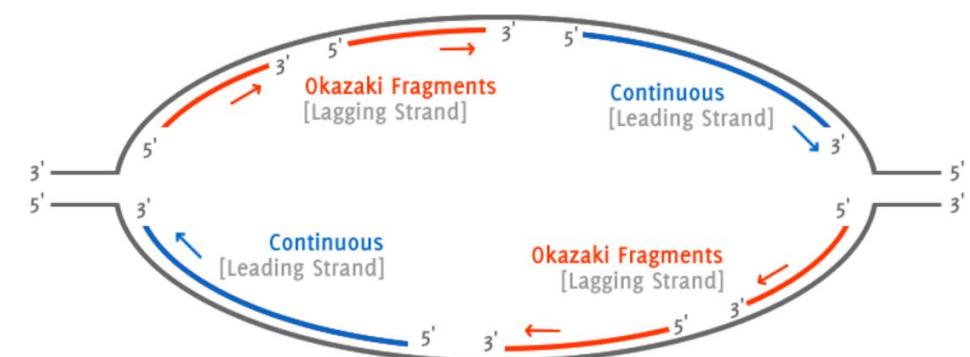
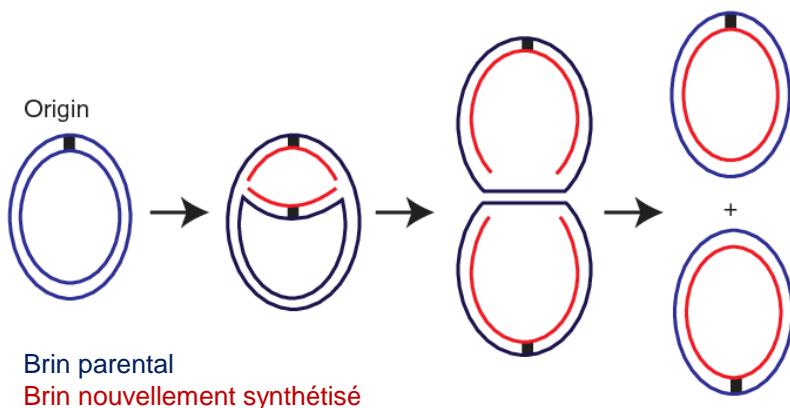
# La traduction chez les bactéries



- **Initiation** : liaison sous-unité 30S - séquence de Shine-Dalgarno (ARNm), appariement d'un fMet-ARNt puis fixation de la sous-unité 50S
- **Elongation** : déplacement du ribosome le long de l'ARNm et accrochage séquentiel des acides aminés
- **Terminaison** : lecture d'un codon STOP, relargage de la protéine synthétisée et séparation des 2 sous-unités.

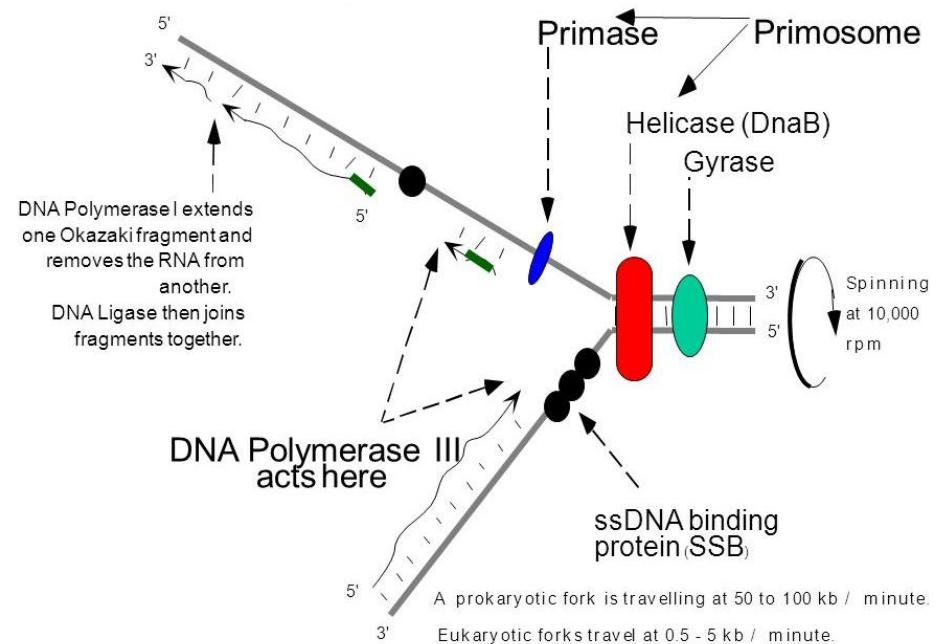
# La réPLICATION chez les bactéries

- Permet la duplication de l'ADN pour être transmis aux générations suivantes
- Rapide chez les bactéries (1000 nucléotides/s chez *E. coli*)
- Débute en un site appelé **origine de réPLICATION (oriC)**
- Séparation du double brin d'ADN et formation de 2 fourches de réPLICATION
- Synthèse des nouveaux brins de façon continue ou non (**fragments d'Okazaki**)



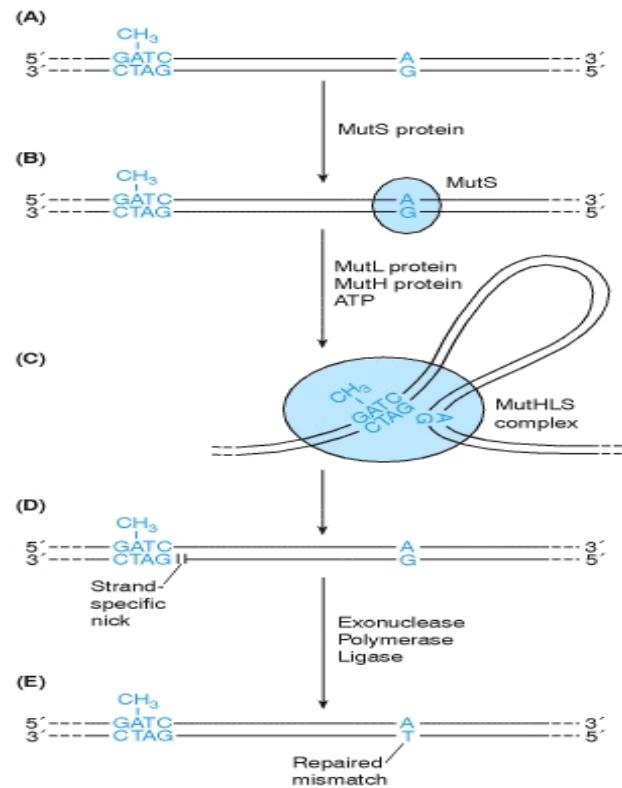
# La réPLICATION chez les bactéries

- Action coordonnée de plusieurs enzymes = **le réplisome**
- déroulement de l'ADN (hélicase, gyrase...)
- synthèse d'amorces d'ARN (primase)
- synthèse du nouvel ADN (ADN polymérase III)
- ligation des nouveaux brins (ligase)
- méthylation des nouveaux brins (méthylase)



# La réPLICATION chez les bactéries

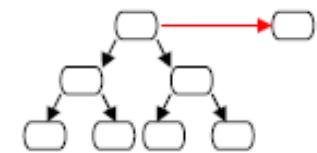
- Taux d'erreurs de réPLICATION extrêmement faible ( $< 1/10^9$  pb)
- Propriétés de relecture et de réparation des ADN polymérases
- Correction des mésappariements par les protéines MutS/MutH/MutL



# Les échanges génétiques chez les bactéries

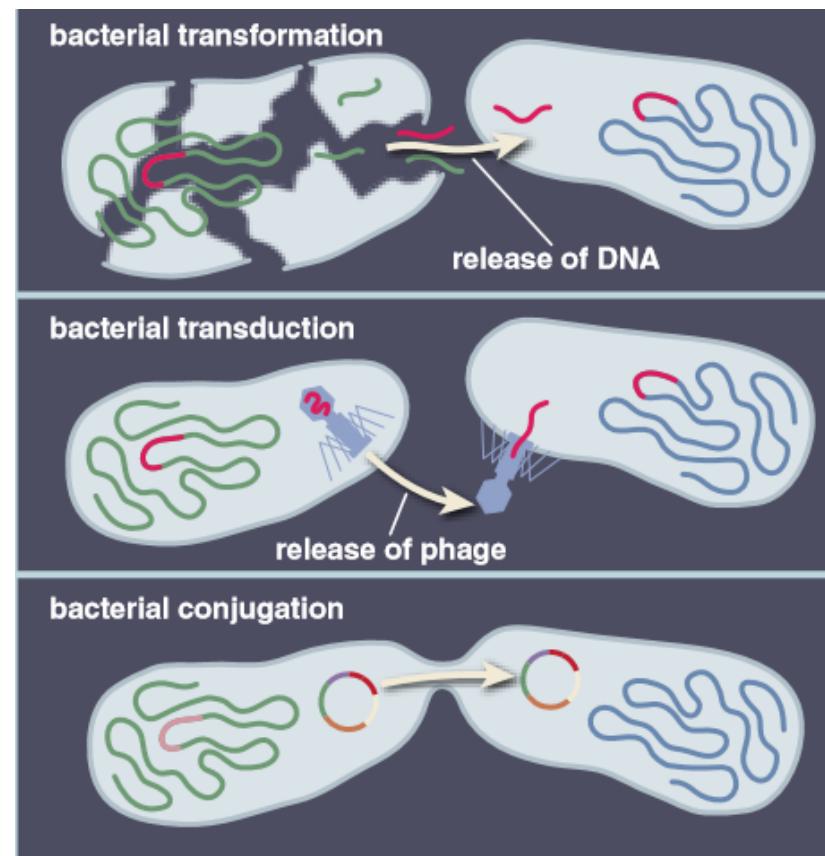
- **Échanges génétiques horizontaux :**

Permettent la dissémination rapide d'informations génétiques entre différentes espèces bactériennes



## 3 mécanismes :

- **Transformation** : intégration d'ADN libre par une bactérie
- **Transduction** : ADN transféré par un phage
- **Conjugaison** : ADN transféré sous forme de plasmide (*plasmides conjuguatifs*)



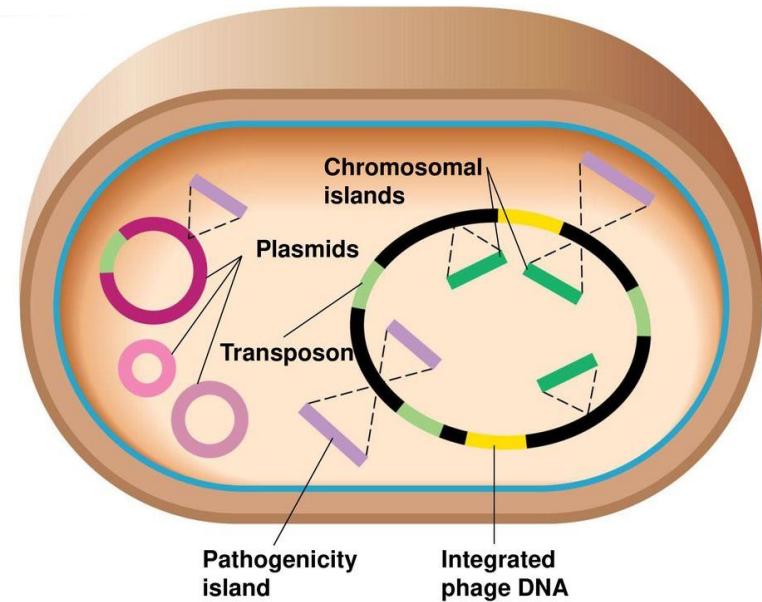
# Les échanges génétiques chez les bactéries

- **Eléments génétiques mobiles :**

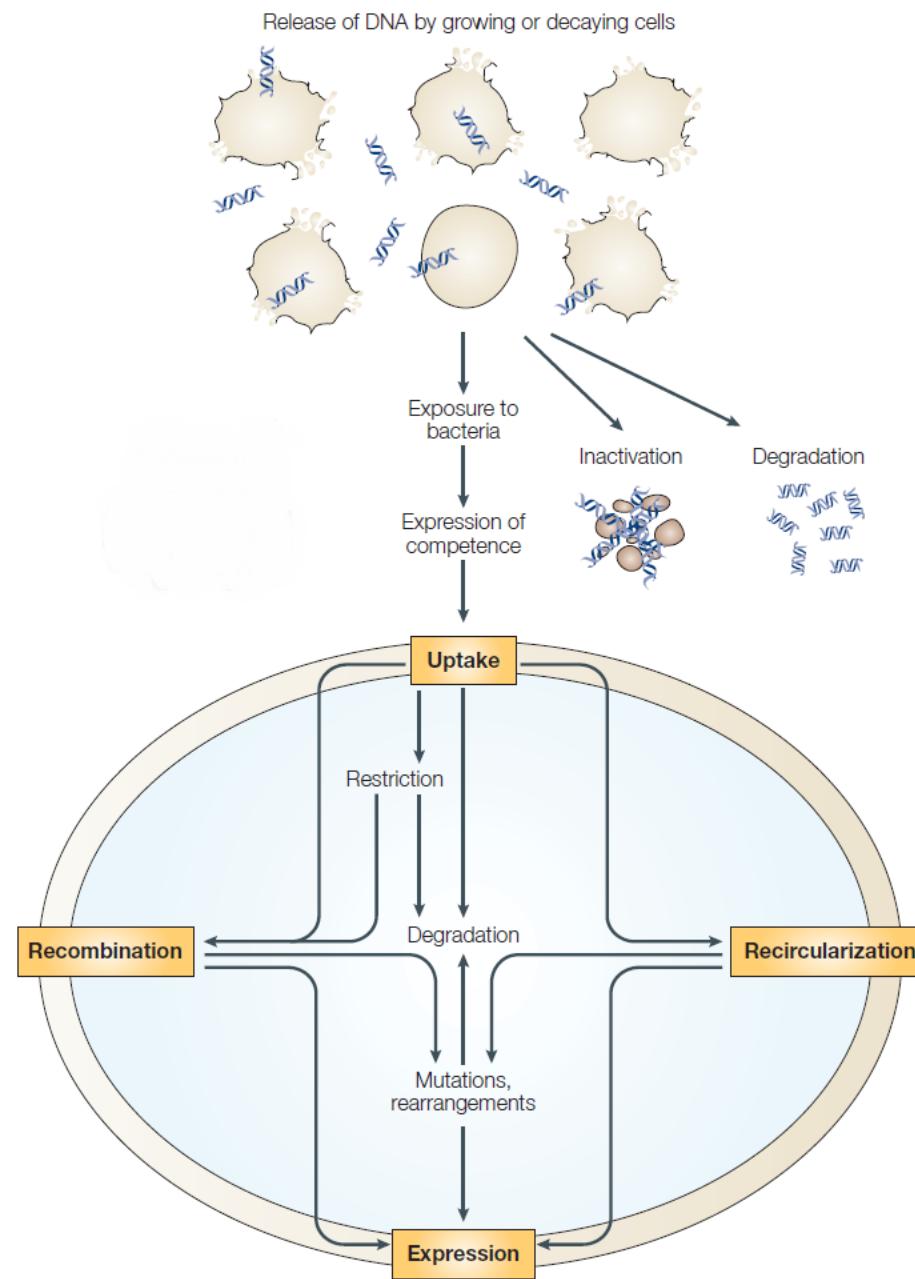
Morceaux d'ADN échangés entre bactéries (génome accessoire)

- **plasmides**
- **éléments transposables** (séquences d'insertion, transposons)
- **îlots de pathogénicité**
- **prophages**

Fonctions non indispensables mais utiles pour occuper une niche écologique donnée

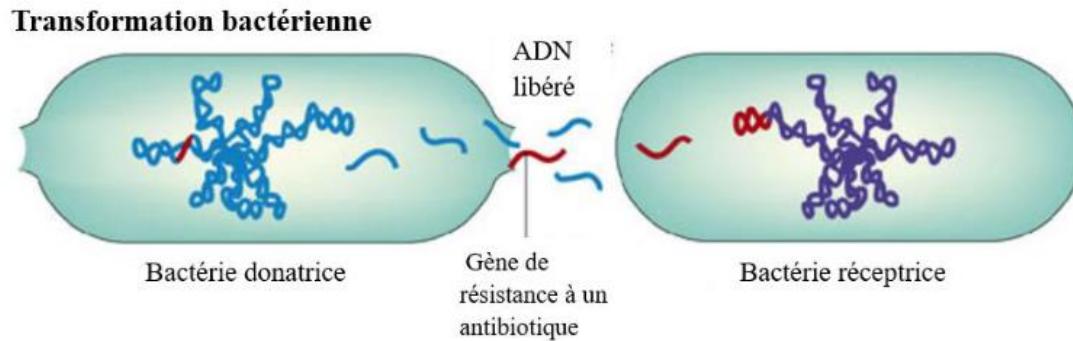


# Transfert d'ADN par transformation



- 1<sup>er</sup> mécanisme découvert de transfert horizontal d'ADN
- Transfert d'**ADN nu** entre bactéries très proches
- Entrée de fragments d'ADN dans des **bactéries** dites **compétentes** puis **recombinaison**

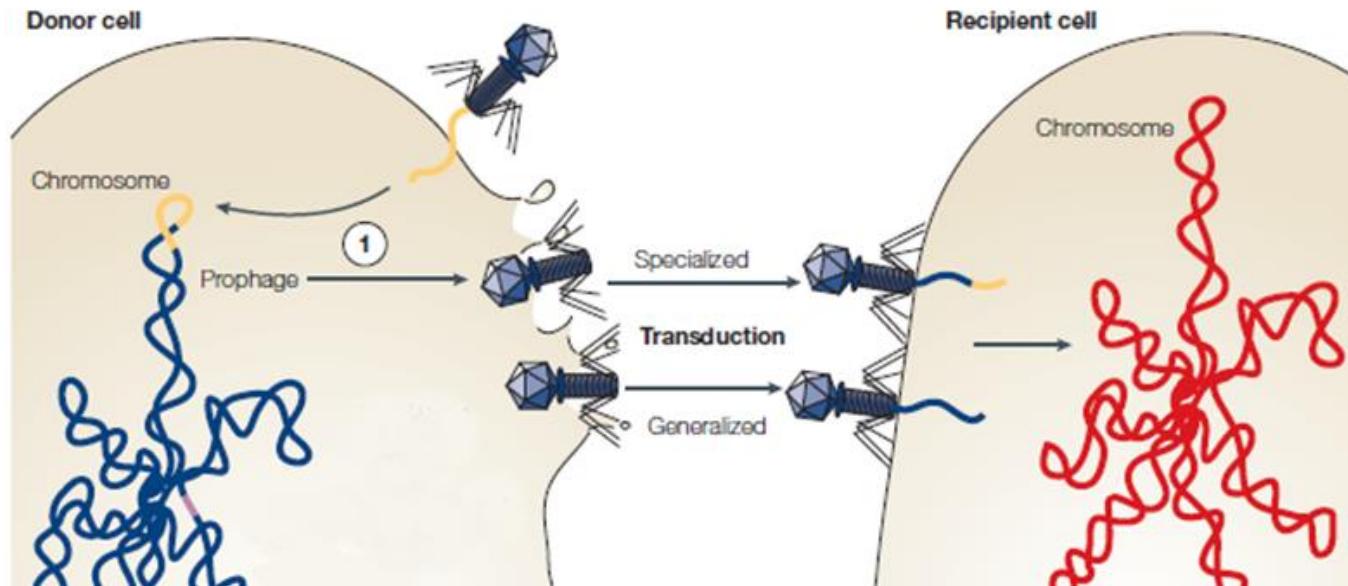
# Transfert d'ADN par transformation



- Bactéries naturellement **transformables**
- Certaines ne le sont qu'à des moments particuliers de leur croissance
- **Compétence naturelle :**
  - *Haemophilus, Neisseria, Streptococcus* (résistance aux  $\beta$ -lactamines)
  - Fin de phase exponentielle de croissance
- **Compétence artificielle :**
  - Traitement au  $\text{CaCl}_2$  et choc thermique
  - Introduire de l'ADN recombinant chez les bactéries

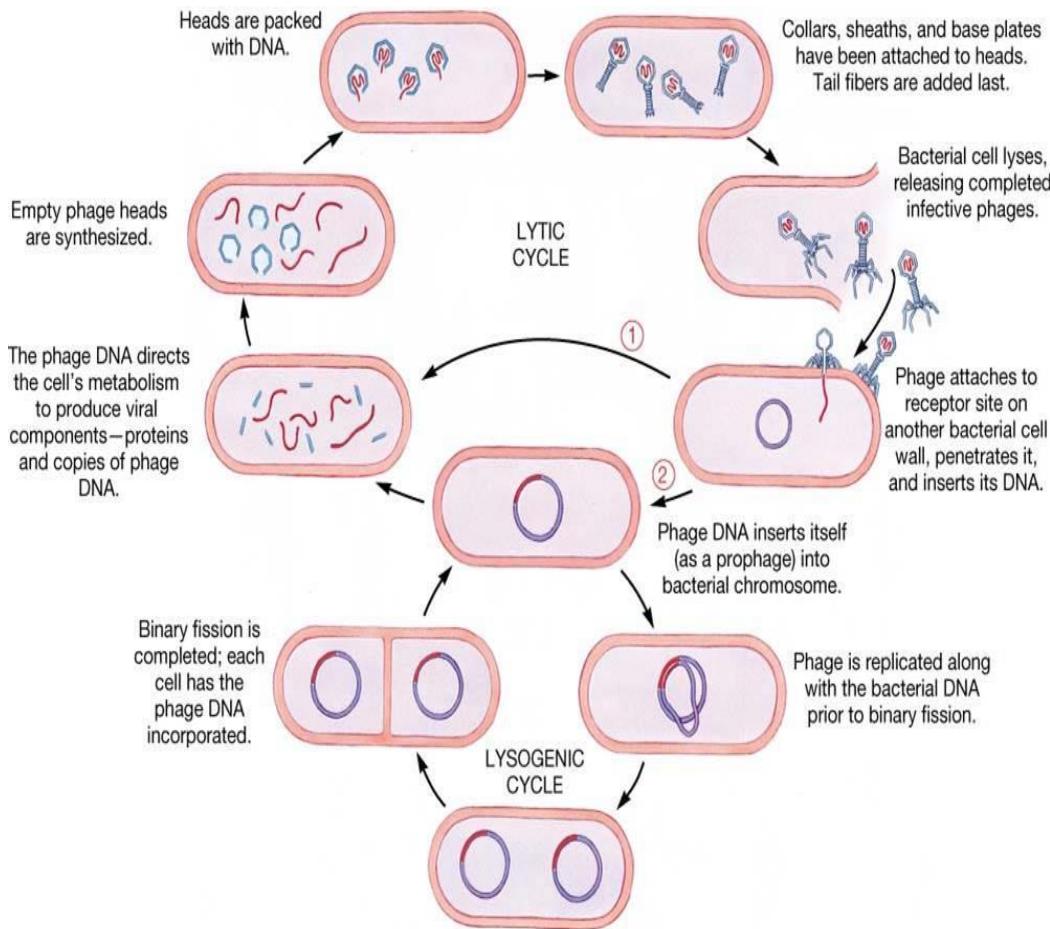
# Transfert d'ADN par transduction

- Transfert de gènes par l'intermédiaire de **bactériophages** (virus de bactéries)



- Phages virulents ou lytiques :**  
Les bactéries explosent en multiples phages
- Phages tempérés :**  
S'intègrent dans chromosome bactérien (**prophages**)

# Transfert d'ADN par transduction



- **Transduction généralisée :**

**Phages virulents** entraînant une lyse bactérienne

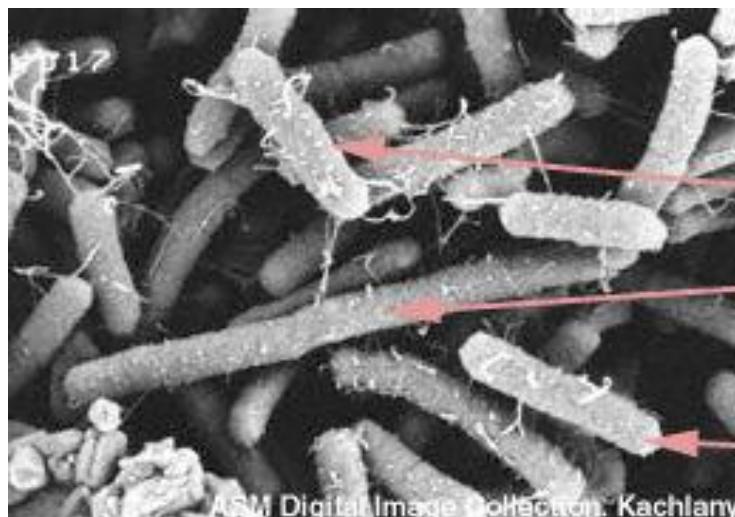
> *cycle lytique*

- **Transduction spécialisée :**

**Phages tempérés** intégrés dans des régions particulière du chromosome (**prophages**)

> *cycle lysogénique*

# Transfert d'ADN par transduction

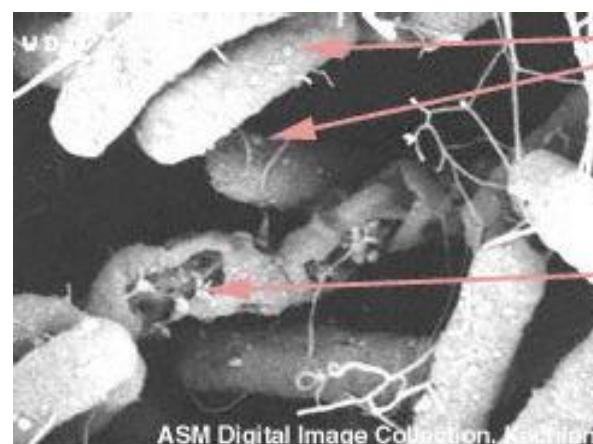
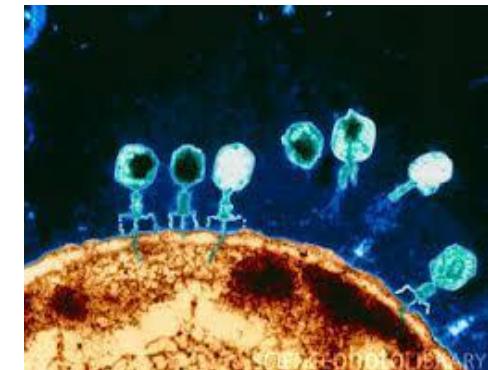


Unidentified Phage particles

Dividing *E.coli* cell

*E.coli* cell

This scanning electron micrograph (SEM) shows *Escherichia coli* cells with phage particles attached to the outside of the cells.



Unidentified phage particles

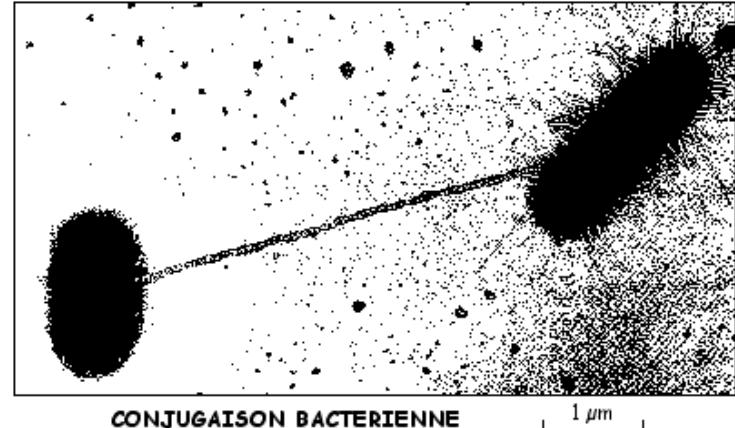
*E. coli* cell with disrupted cell envelope, presumably due to phage release.

Unerupted *E. coli* cell

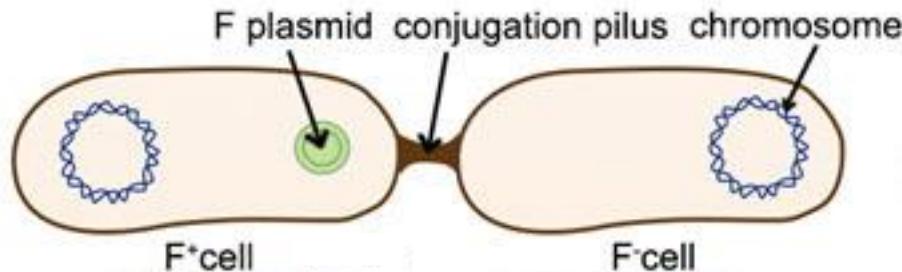
This scanning electron micrograph (SEM) shows *Escherichia coli* cells with disrupted cell envelopes, presumably due to phage release.

# Transfert d'ADN par conjugaison

- Echange d'ADN sous forme de **plasmide**
  - Rôle essentiel des pili sexuels (facteurs de fertilité ou facteurs F)



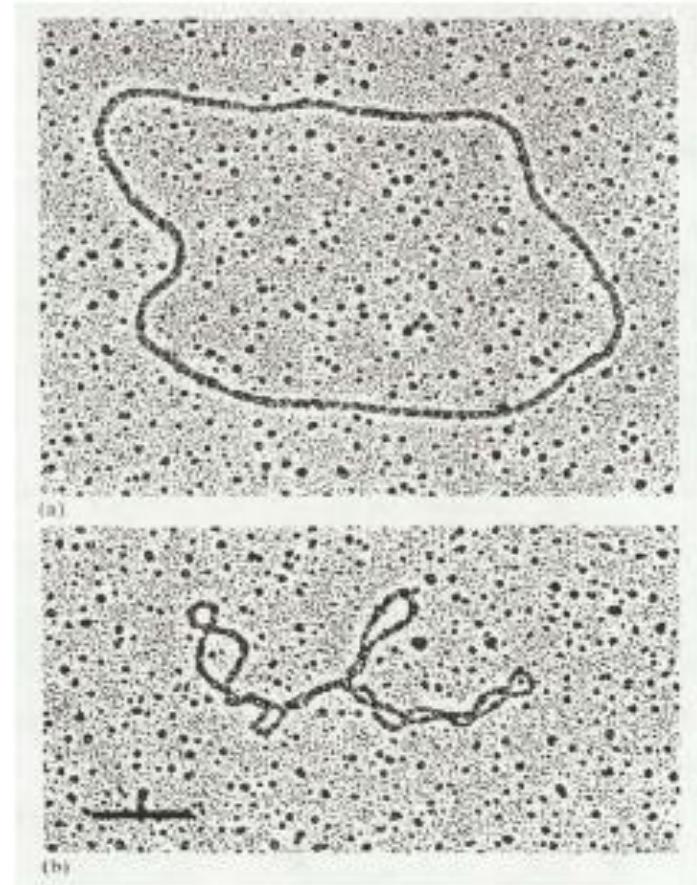
- > Bactérie donneuse : bactérie mâle (F+)
- > Bactérie réceptrice : bactérie femelle (F-)



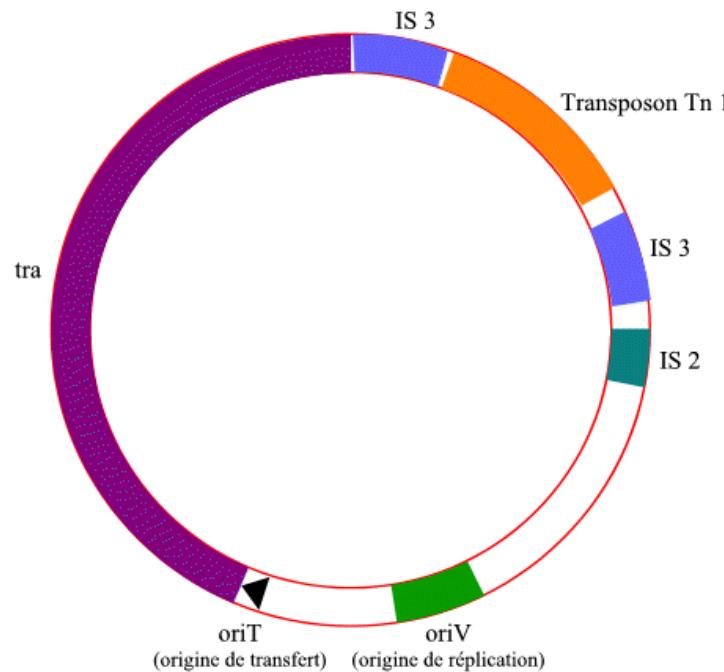
Mécanisme très impliqué dans la diffusion de la résistance aux antibiotiques

# Les plasmides

- Molécules d'ADN bicaténaire, circulaires et cytoplasmiques
- Petite taille (5 à 4000 < chromosome)
- Capables de réPLICATION autonome
- Transmission verticale & horizontale +++
- Non indispensables au métabolisme normal de la bactérie hôte
- Caractérisés par :
  - type et origine de réPLICATION (oriV)
  - nombre de copies
  - compatibilité (InC)
  - spécificité d'hôte
  - mode de transfert



# La conjugaison



- **Plasmides conjuguatifs**  
→ portent l'opéron *tra*
- Plasmides non-conjugatifs

Opéron *tra* :

## Facteur de fertilité ou facteur F

Synthèse des pili sexuels

*traA, traL, traE, traK, traB, traV, traC, traW, traU, traF, traQ, TraH, traG*

Exclusion de surface

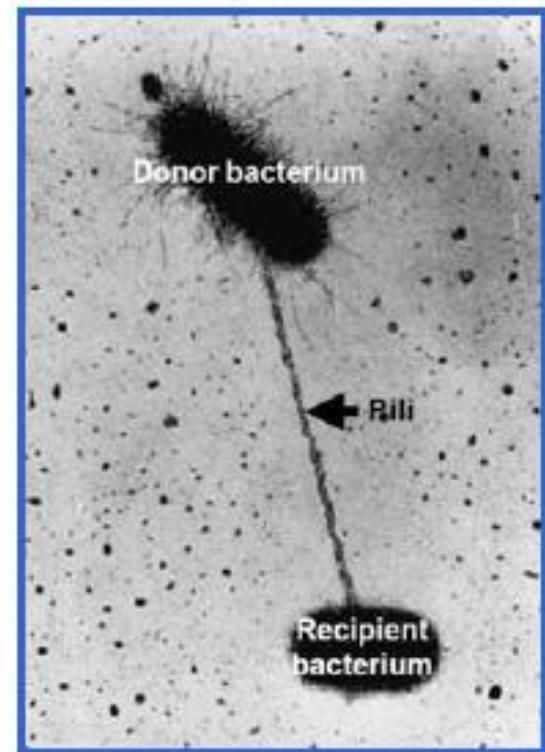
*traS, traT*

Transfert de l'ADN

*traM, traY, traD, traI, traZ*

Régulation

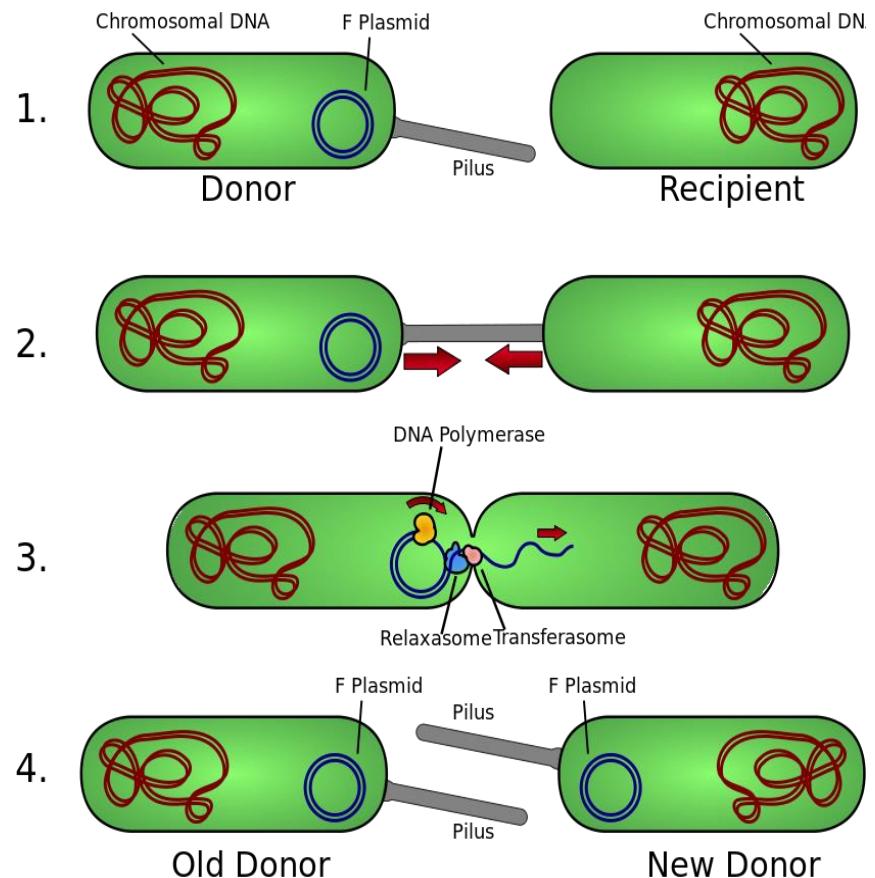
*finP, finO, traJ*



# La conjugaison

## Transfert en 4 étapes :

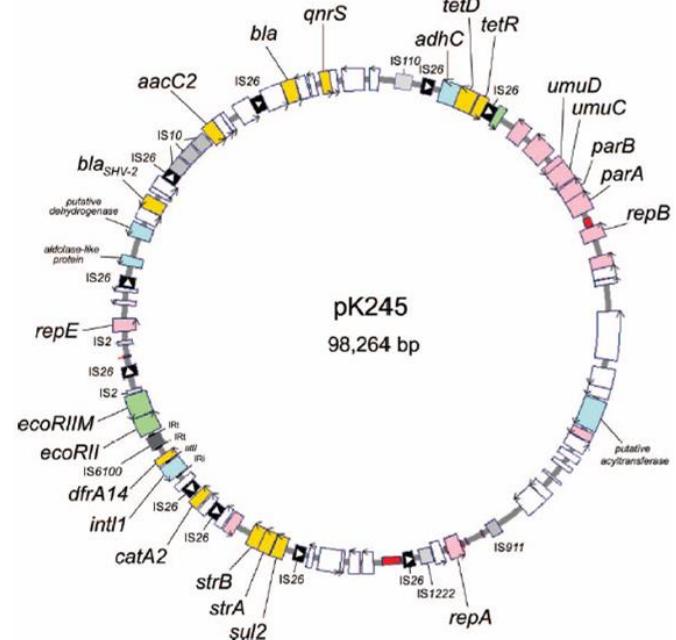
- 1) Reconnaissance entre donneur ( $F^+$ ) et accepteur ( $F^-$ ) grâce à la synthèse du pili
- 2) Transfert d'un des 2 brins du plasmide
- 3) Synthèse du brin complémentaire chez l'accepteur et le donneur
- 4) Recircularisation du plasmide chez l'accepteur et le donneur



Finalement, le facteur F persiste chez la bactérie donatrice ( $F^+$ ) et une copie du plasmide est acquise par la bactérie réceptrice qui devient  $F^+$

# Les plasmides

- **Supports de gènes impliqués dans une meilleure adaptation à l'environnement :**
  - résistance aux antibiotiques
  - résistance aux antiseptiques
  - résistance aux métaux lourds
  - résistance aux bactériophages
  - synthèse de bactériocines
  - pouvoir pathogène des colibacilles entéropathogènes
  - pouvoir pathogène des staphylocoques dans l'impétigo (exfoliatine)
  - ...



# Eléments transposables

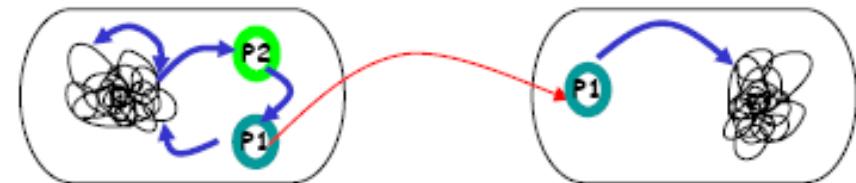
Séquences d'insertion (IS)

Transposons composites

Transposons non composites

Transposons conjugatifs

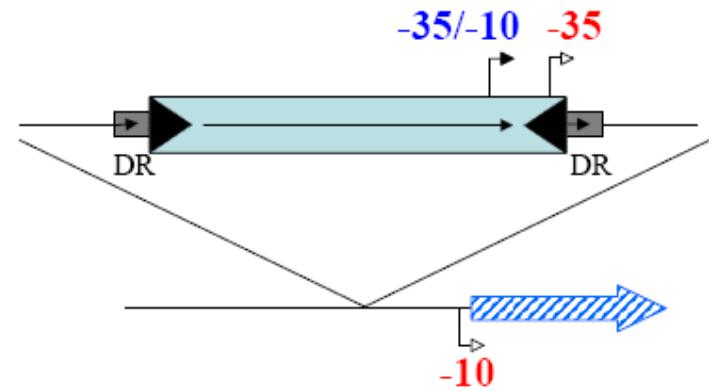
Ilots de pathogénicité



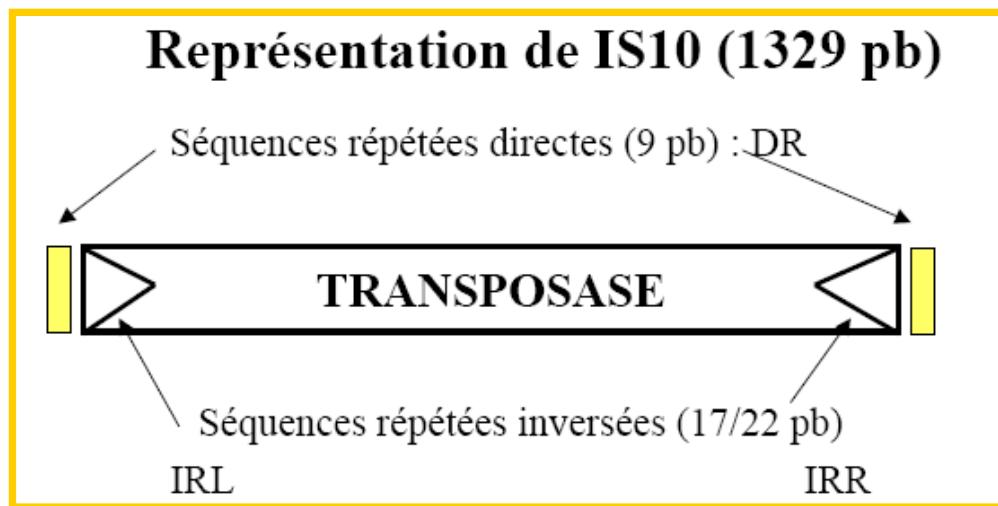
- Capables d'être transférés d'un site donneur vers un site cible, localisé ou non sur une même molécule d'ADN, plasmidique ou chromosomique
- Séquences d'ADN linéaire, mobiles, n'apparaissant jamais à l'état libre
- À l'origine de variations génétiques et de remaniement des gènes sur le chromosome bactérien
- Transposition fait intervenir une recombinase :
  - transposases
  - Intégrases
  - résolvases / invertases

# Séquences d'insertion (IS)

- $\approx 1000$  IS recensées
- Petite taille (800 à 2500 pb)
- Eléments mobiles les plus simples
- Ne codent que pour leur mobilité (pas d'autre fonction)
- Localisées dans le chromosome et/ou les plasmides (1 ou plusieurs copies)
- Interviennent dans l'expression de certains gènes en les inactivant par interruption ou en les activant

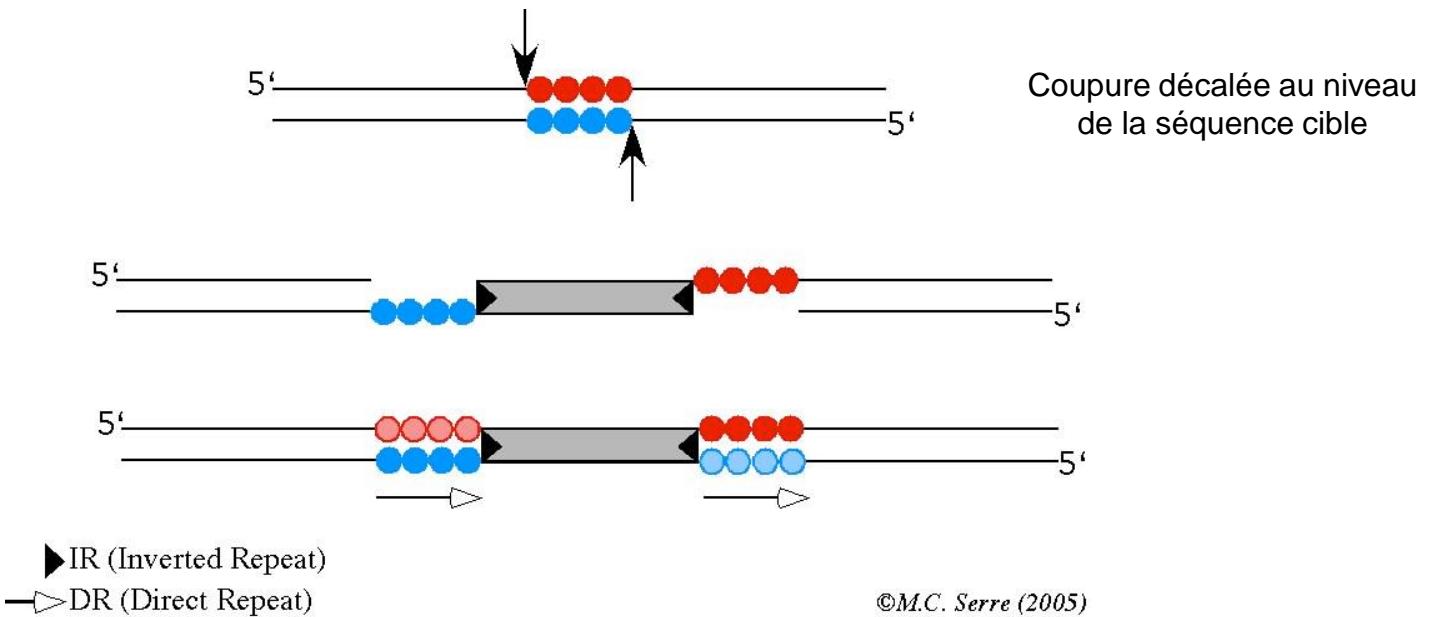


# Séquences d'insertion (IS)



- Délimitées par de courtes séquences répétées (jusqu'à 40 pb) en orientation inverse (IRS)
- Cadre de lecture codant pour une transposase
- Transposase excise le transposon et coupe la cible

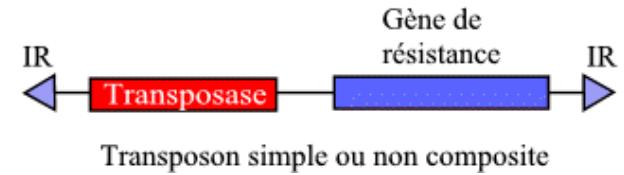
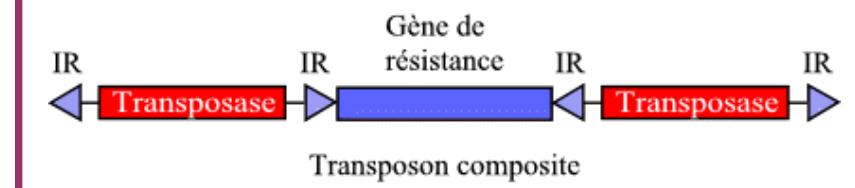
# Séquences d'insertion (IS)



- L'insertion de ces éléments génétiques conduit toujours à la duplication d'une courte séquence d'ADN (5 à 9 paires de bases)
- Il existe des sites d'insertion privilégiés, particulièrement riches en adénine et thymine

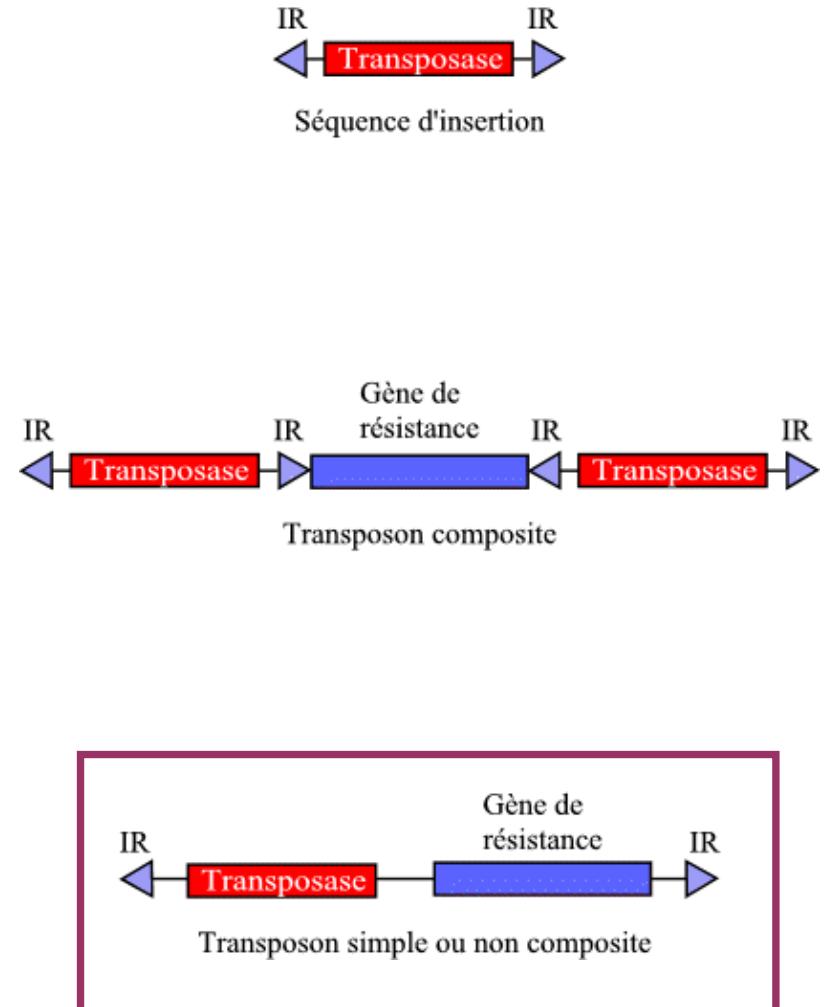
# Transposons composites

- Taille moyenne (3 à 10 kb)
- Encadrés de 2 copies d'IS en orientation directe ou inverse
- Gènes centraux non liés aux fonctions de transposition
- IS fonctionnent de paire et emportent le segment d'ADN qu'elles encadrent
- Généralement, seule 1 des 2 IS code pour une transposase fonctionnelle



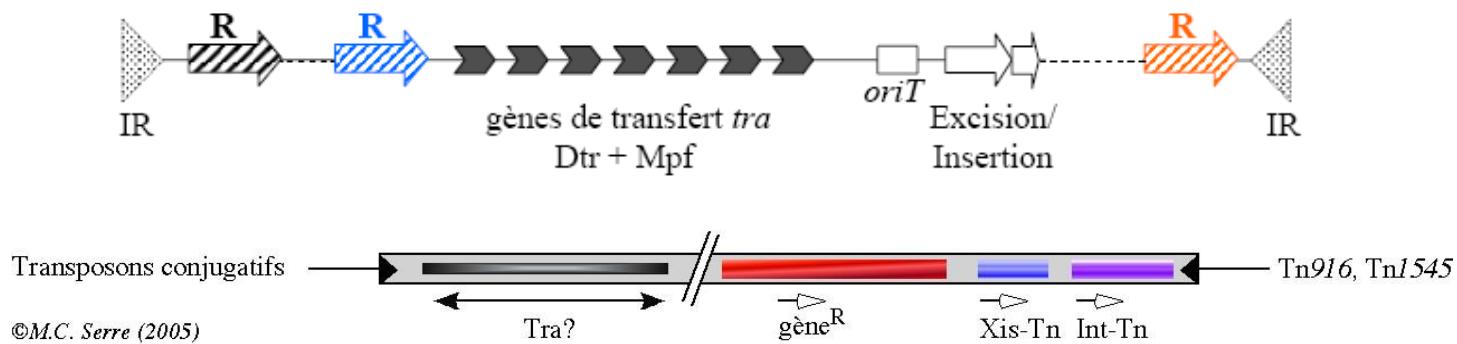
# Transposons non composites

- Taille allant jusqu'à 70 kb
- Absence d'IS à leurs extrémités, mais présence de séquences terminales répétées en orientation inverse (IRS)
- Codent pour une transposase, une résolvase
- 2 séquences répétées directes (DR) aux extrémités



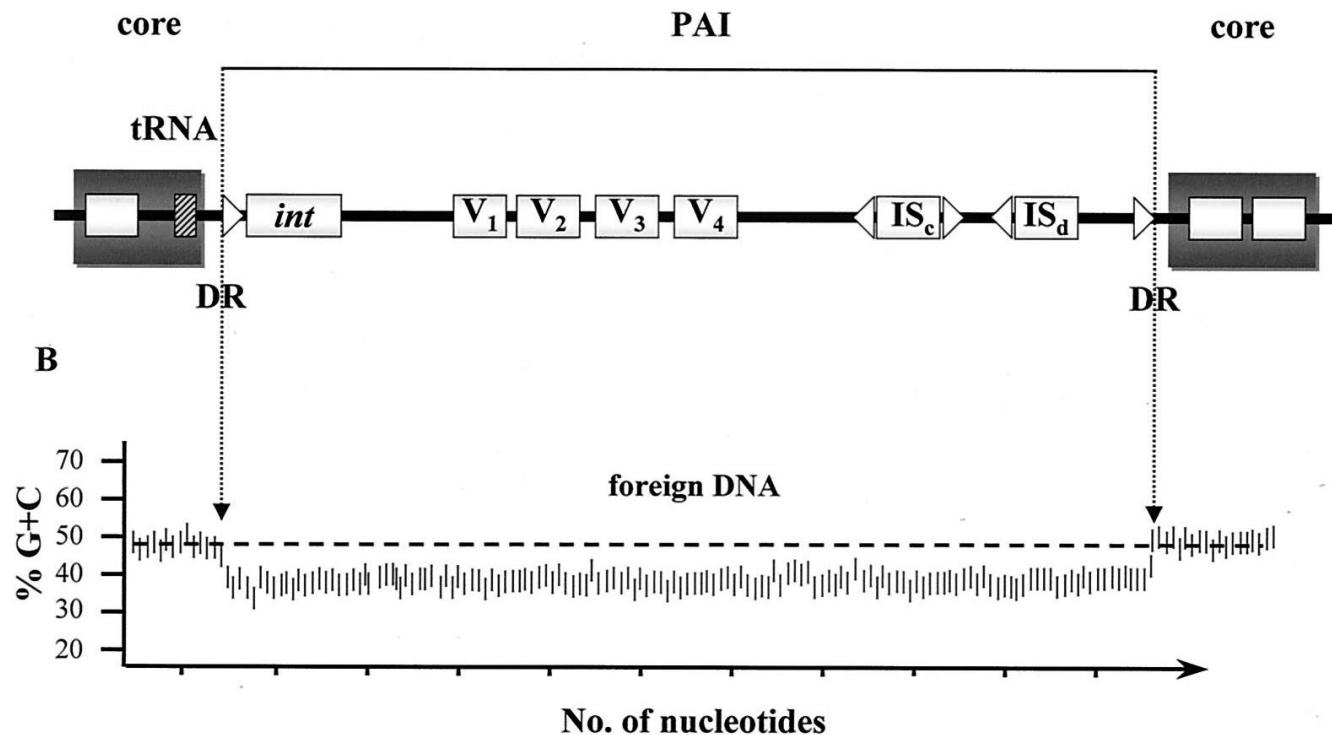
# Transposons conjuguatifs

- Eléments de grande taille
- Seuls capables de transfert intercellulaire
- Propriétés d'intégration des bactériophages (gène d'intégrase) et propriétés de transfert des plasmides conjuguatifs (gènes *tra*)
- Ne possèdent pas les fonctions de réPLICATION



# Ilots de pathogénicité

- Fragments d'ADN de grande taille (20 à 200 kb)
- % G-C différent du chromosome
- Bornes : courtes séquences répétées inversées
- Association avec des gènes de transfert



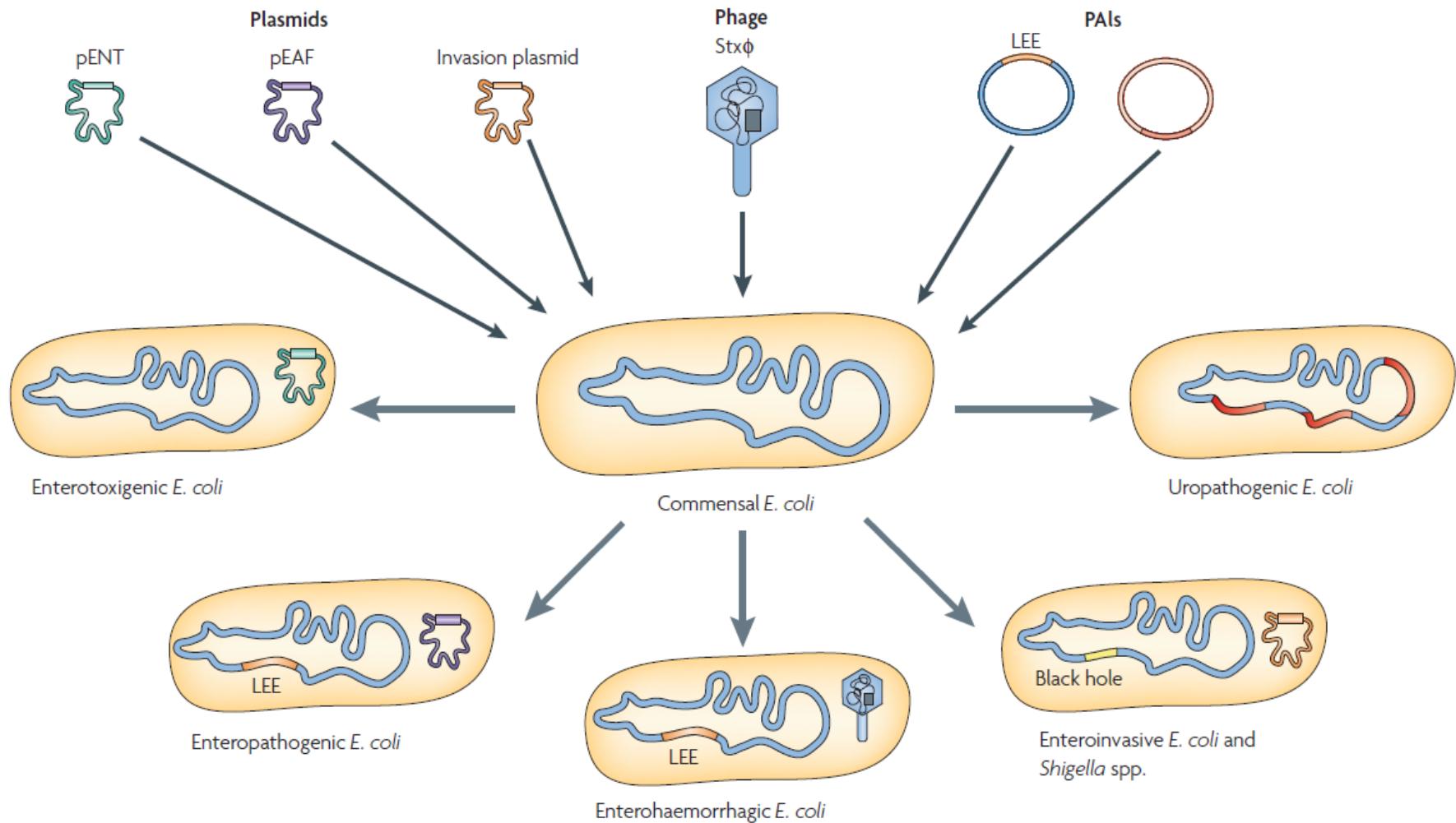
# Ilots de pathogénicité

- Caractérisent les souches pathogènes
- Portent des gènes de virulence et de régulation
- Issus d'éléments transposables ou de bactériophages
- Proviennent probablement de transferts horizontaux

TABLE 1 Major virulence features encoded by pathogenicity islands

Virulence feature	Examples
Adherence factors	Diarrheagenic <i>Escherichia coli</i> Uropathogenic <i>Escherichia coli</i> <i>Vibrio cholerae</i> <i>Listeria</i> spp.
Toxins	Uropathogenic <i>Escherichia coli</i> <i>Staphylococcus aureus</i>
Iron uptake systems	Uropathogenic <i>Escherichia coli</i> <i>Shigella flexneri</i> <i>Yersinia</i> spp.
Invasions, modulins, effectors	Diarrheagenic <i>Escherichia coli</i> <i>Salmonella</i> spp. <i>Shigella</i> spp. <i>Listeria</i> spp.
Type III secretion systems	Diarrheagenic <i>Escherichia coli</i> <i>Pseudomonas syringae</i> <i>Erwinia</i> spp. <i>Yersinia</i> spp. <i>Salmonella</i> spp. <i>Shigella</i> spp.
Type IV secretion system	<i>Helicobacter pylori</i> <i>Agrobacterium tumefaciens</i>

# Supports génétiques de la virulence



Chaque sous groupe de *E. coli* pathogène à une signature génétique complexe

# *Escherichia coli* uropathogène (UPEC)

- Infections urinaires, physiopathologie :



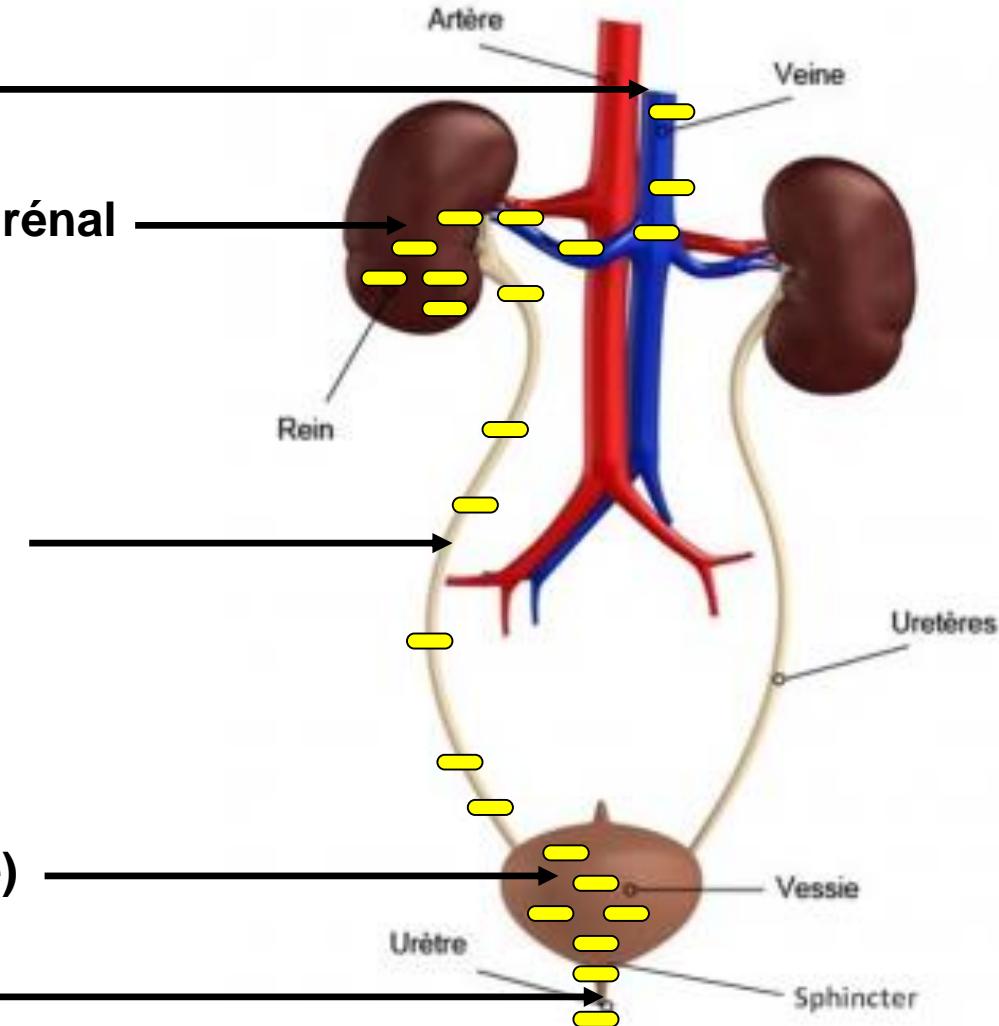
5 Bactériémie

4 Infection du parenchyme rénal  
(pyélonéphrite)

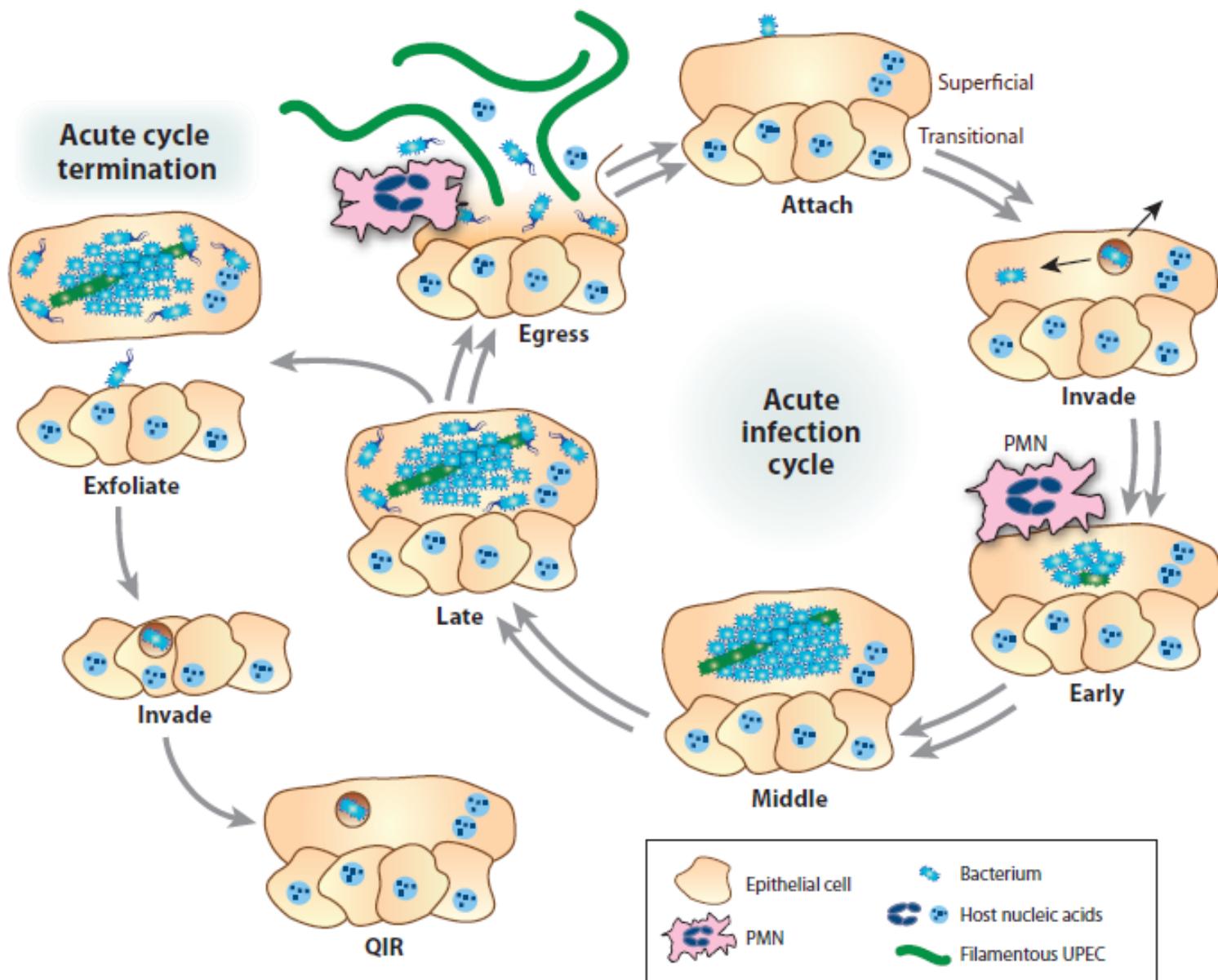
3 Colonisation ascendante  
des uretères

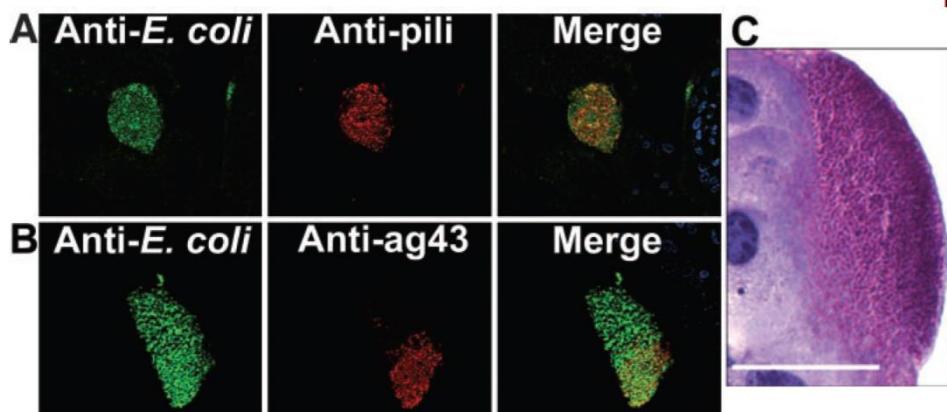
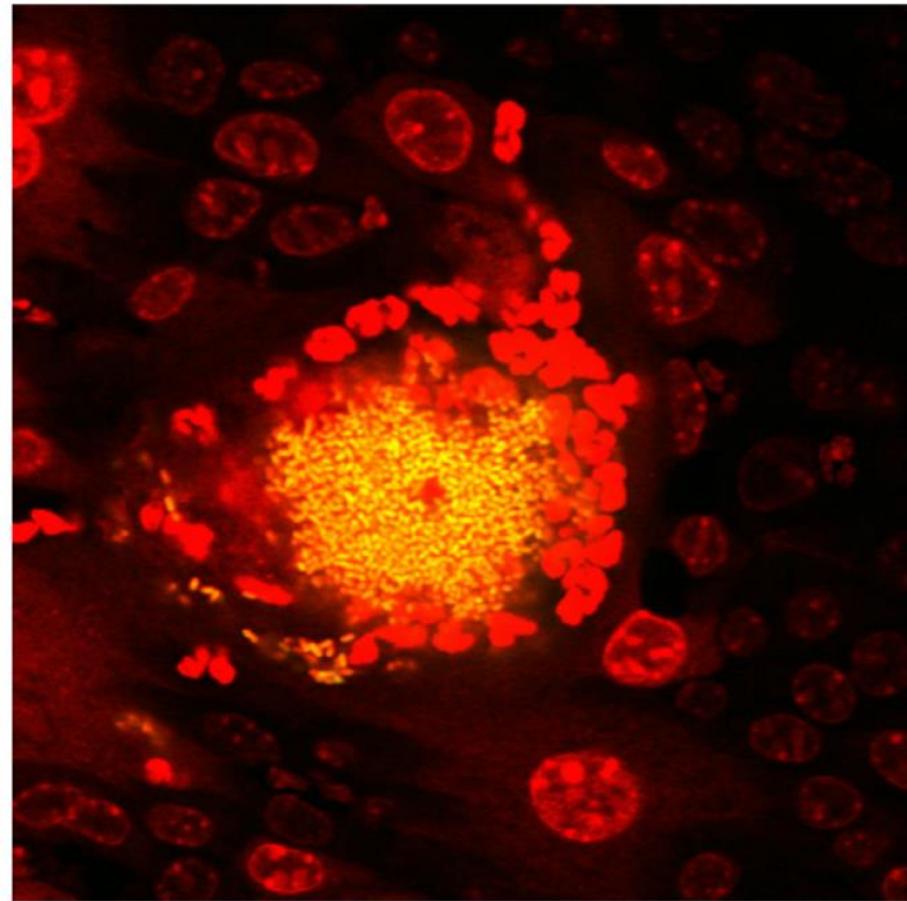
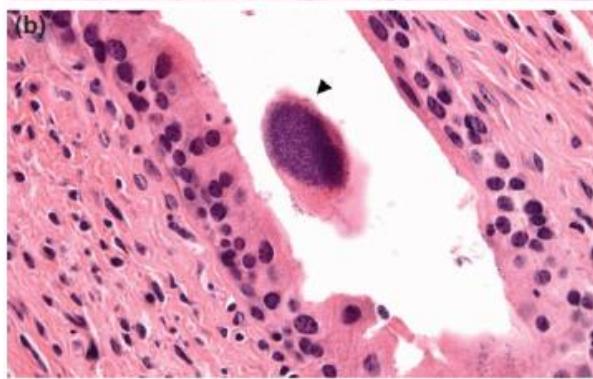
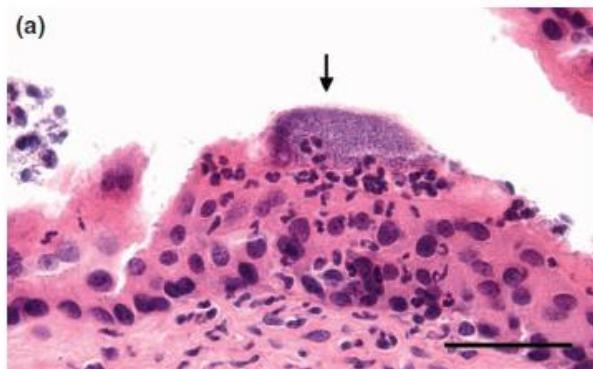
2 Infection vésicale (cystite)

1 Colonisation urétrale



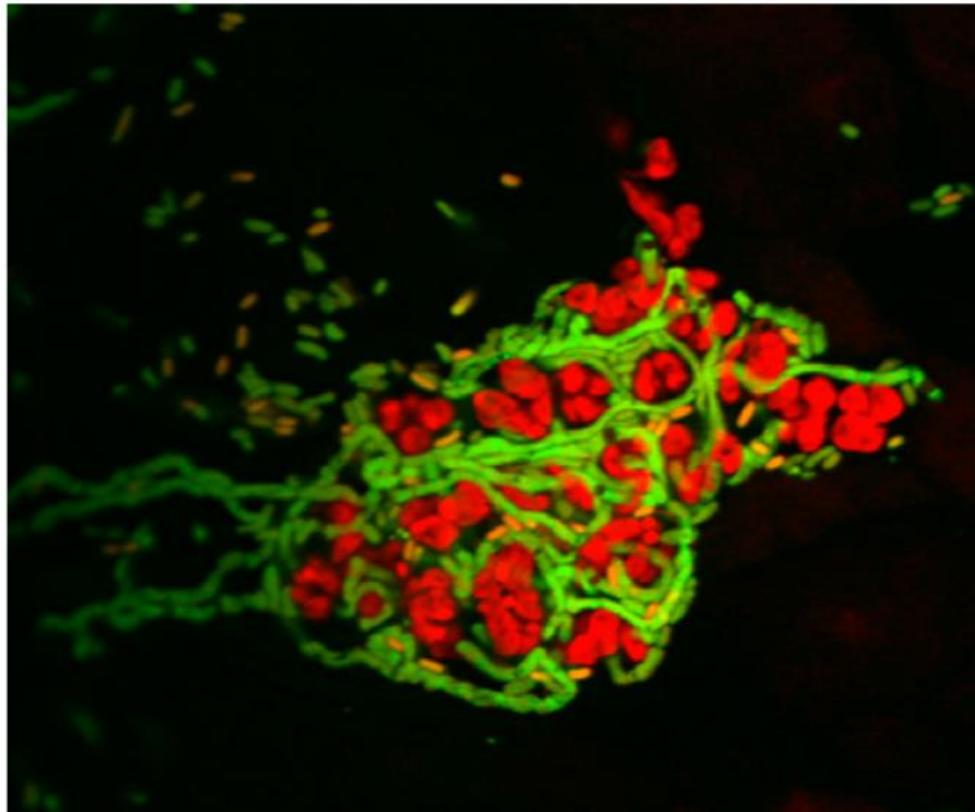
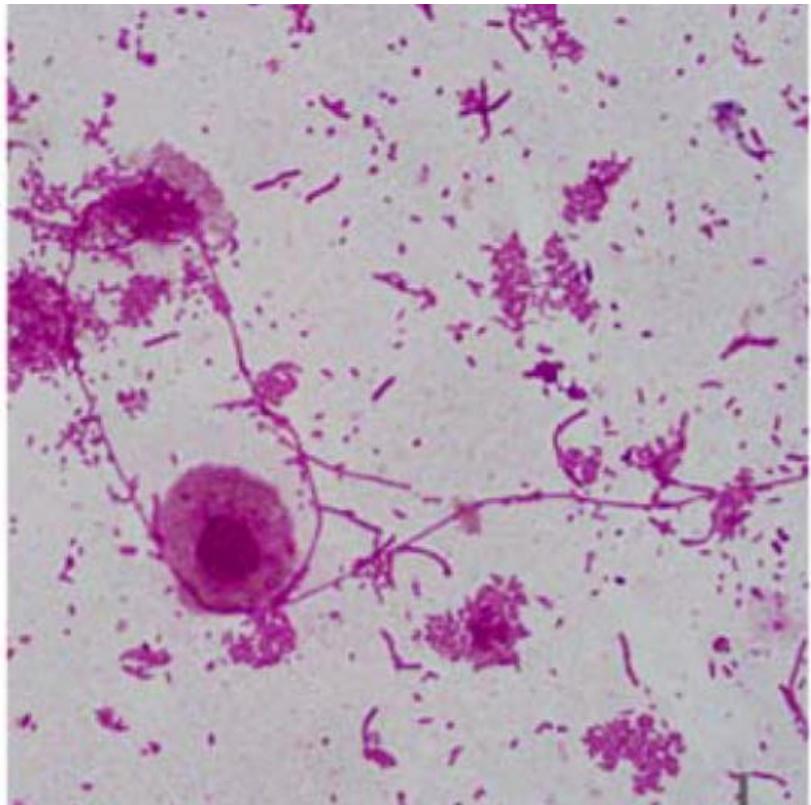
# *Escherichia coli* uropathogène (UPEC)





## Communautés bactériennes Intracellulaires (IBC)

Fig. 3. Bacteria in pods express type 1 pili, antigen 43, and polysaccharides. Pods were immunostained

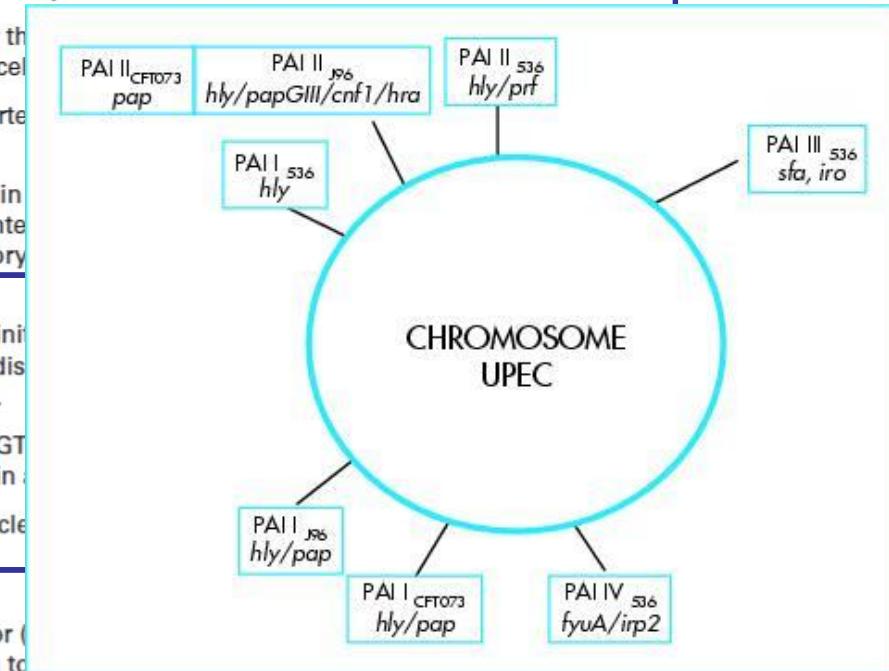


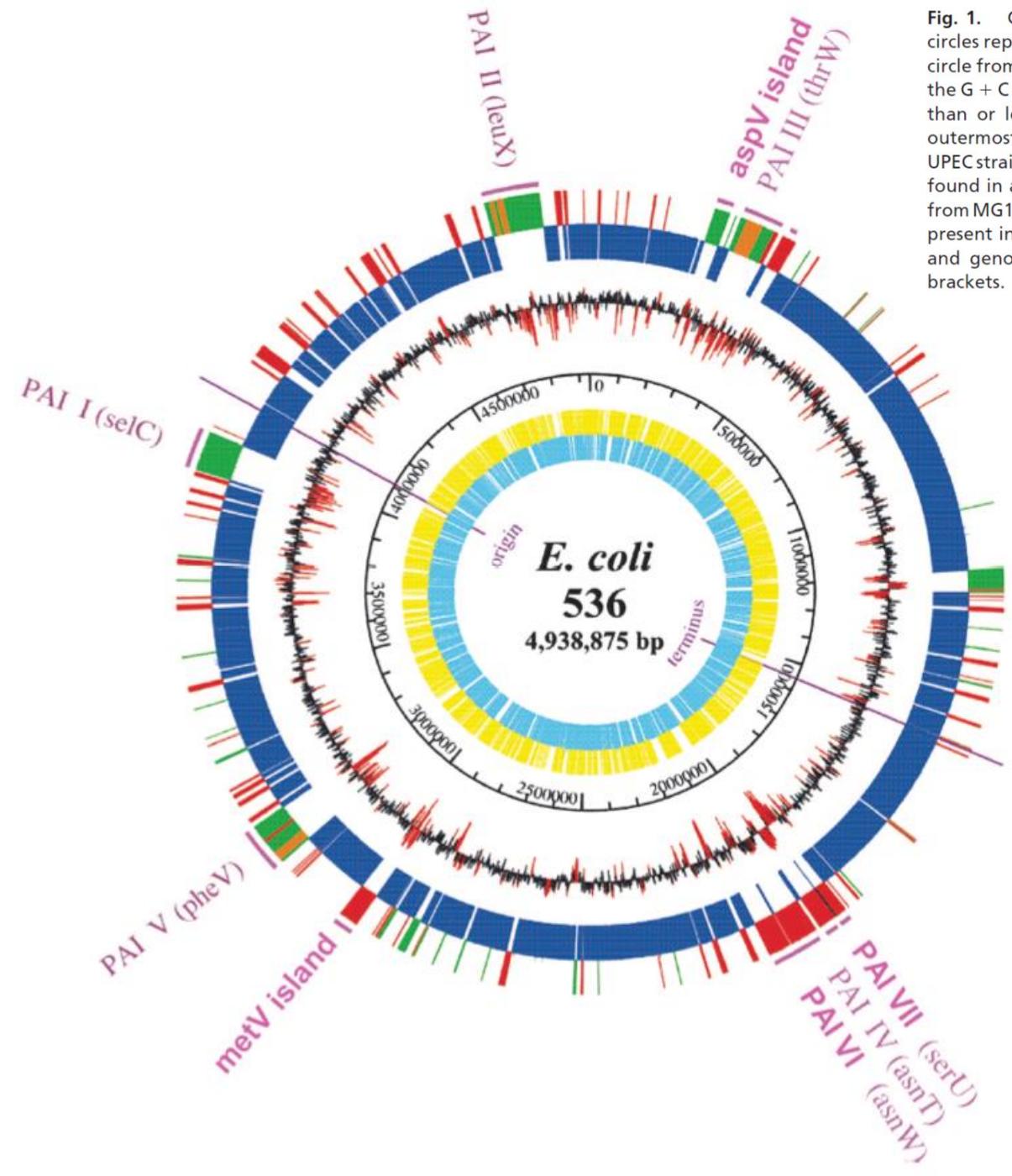
## Filmentation :

- Résistance à la phagocytose par les polynucléaires neutrophiles
- Requiert l'expression d'un inhibiteur de la division cellulaire : SulA

Virulence Factor	Function
<b>Adhesins</b>	
CUP pili (e.g. type 1 pili, P pili, S pili, F9 pili)	Adherence to a broad range of biotic and abiotic surfaces. For example type 1 pili Mediate adherence to bladder epithelial cells and catheter material. Type 1 and P pili have been shown to induce TLR4 signaling independent of LPS.
Antigen 43	AIDA-I-type autotransporter protein involved in adhesion, cell aggregation, biofilm formation, and long-term colonization of the urinary tract.
UpaB	Autotransporter protein that binds to extracellular matrix components, such as fibrinogen, fibronectin, and laminin.
UpaC	Autotransporter protein that promotes biofilm formation on abiotic surfaces.
UpaG	Trimeric autotransporter that promotes cell aggregation, biofilm formation, and adherence to bladder cells through binding fibronectin and laminin.
UpaH	AIDA-I-type autotransporter protein that promotes biofilm formation and bladder colonization.
Curli fibers	Amyloid fibers involved in adhesion, cell aggregation, and biofilm formation. Curli have been shown to interact with the human LL-37 anti-microbial peptide, and elicit a pro-inflammatory immune response through induction of IL-8.
<b>Toxins</b>	
$\alpha$ -Hemolysin	Pore-forming toxin that initiates degradation of host cell protein paxillin, along with other targets, to disrupt cell-to-cell interactions and signaling cascades leading to lytic events.
Cytotoxic necrotizing factor-1 (CNF-1)	Constitutively activates GTPases Rho, Rac, and Cdc42 through deamidation leading to alterations in actin cytoskeleton arrangement.
Cytotoxic necrotizing factor-1 (CNF-1)	Blocks eukaryotic cell cycle in G <sub>2</sub> phase, leading to the development of giant mononucleated cells.
<b>Immune suppression</b>	
TcpC	Toll/interleukin-1 receptor (TIR) UPEC protein homolog, which inhibits MyD88 and downstream signaling to suppress the host innate immune response.
YbcL	Suppresses acute neutrophil migration.
<b>Metal acquisition and chelation</b>	
Siderophores (e.g. hydroxamateaerobactin, enterobactin, salmochelin, versiniabactin)	Bacterial-secreted molecules for scavenging essential metals from the host environment. UPEC yersiniabactin has also been shown to protect UPEC from the harmful effects of copper ions through chelation.
<b>Motility</b>	
Flagella	Allows for motility within the urinary tract and can serve in adhesion to surfaces.

Virulence Factor	Function
<b>Adhesins</b>	
CUP pili (e.g. type 1 pili, P pili, S pili, F9 pili)	Adherence to a broad range of biotic and abiotic surfaces. For example type 1 pili Mediate adherence to bladder epithelial cells and catheter material. Type 1 and P pili have been shown to induce TLR4 signaling independent of LPS.
Antigen 43	AIDA-I-type autotransporter protein involved in adhesion, cell aggregation, biofilm formation, and long-term colonization of the urinary tract.
UpaB	Autotransporter protein that binds to extracellular matrix components, such as fibrinogen, fibronectin, and laminin.
UpaC	Autotransporter protein that promotes biofilm formation on abiotic surfaces.
UpaG	Trimeric autotransporter that adherence to bladder cells.
UpaH	AIDA-I-type autotransporter colonization.
Curli fibers	Amyloid fibers involved in have been shown to interact with other targets, to diselicit a pro-inflammatory response.
<b>Toxins</b>	
$\alpha$ -Hemolysin	Pore-forming toxin that interacts with other targets, to dislead to lytic events.
Cytotoxic necrotizing factor-1 (CNF-1)	Constitutively activates GTPases leading to alterations in cell morphology.
Cytotoxic necrotizing factor-1 (CNF-1)	Blocks eukaryotic cell cycle progression in mononucleated cells.
<b>Immune suppression</b>	
TcpC	Toll/interleukin-1 receptor (TIR) domain containing downstream signaling toxin.
YbcL	Suppresses acute neutrophil migration.
<b>Metal acquisition and chelation</b>	
Siderophores (e.g. hydroxamateaerobactin, enterobactin, salmochelin, versiniabactin)	Bacterial-secreted molecules for scavenging essential metals from the host environment. UPEC yersiniabactin has also been shown to protect UPEC from the harmful effects of copper ions through chelation.
<b>Motility</b>	
Flagella	Allows for motility within the urinary tract and can serve in adhesion to surfaces.





**Fig. 1.** Genetic map of the UPEC strain 536 chromosome. The two inner circles represent all putative genes, depending on ORF orientation. The third circle from the center gives the scale. The fourth circle from the center shows the G + C distribution. Regions with a highly aberrant G + C content (greater than or less than 2-fold standard deviation) are highlighted in red. The outermost circles show the result of a three-way genome comparison with the UPEC strain CFT073 and *E. coli* MG1655 (K-12) genomes: blue, backbone genes found in all three strains; red, genes present in 536 and CFT073 but absent from MG1655; green, genes found in 536 only; orange, genes of 536, which are present in CFT073 but located in a different genomic region. Pathogenicity and genomic islands (GEIs) are highlighted; flanking tRNAs are given in brackets.

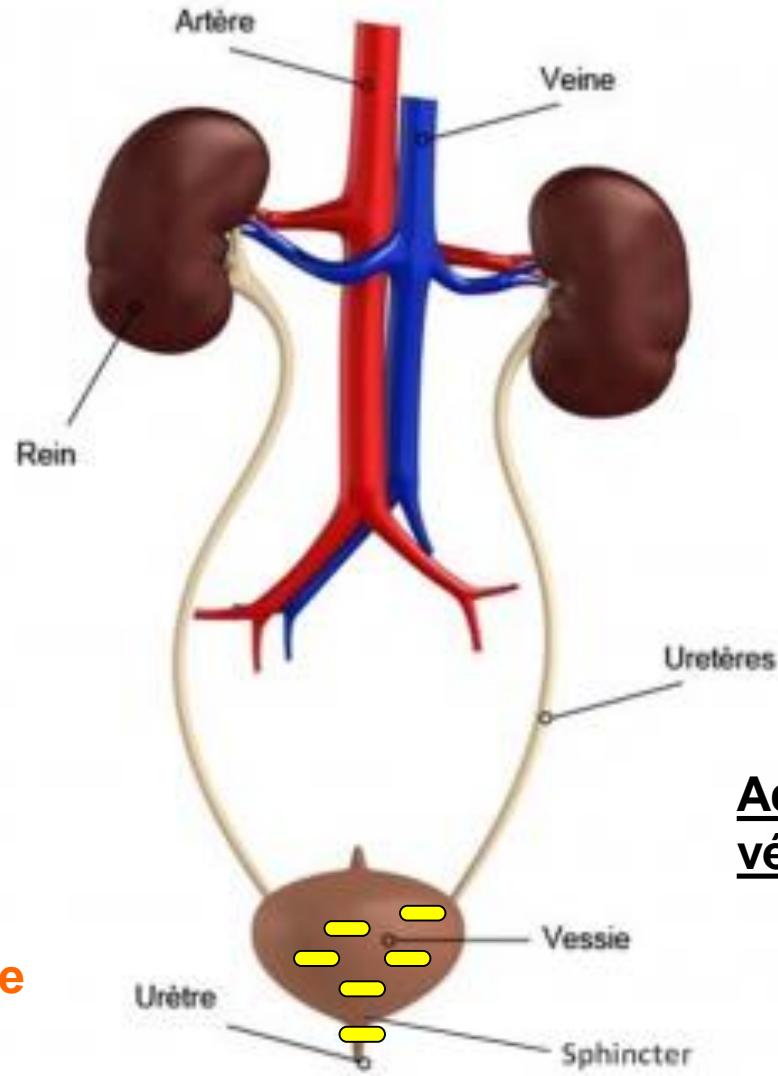
Comparaison *E. coli* 536  
(UPEC) vs CFT073 (UPEC)  
vs K-12 (commensale)

Bleu: core génome  
Rouge: gènes absents chez K-12  
Vert: gènes absents chez K-12 & CFT073  
Orange: gènes autre endroit chez CFT073

432 gènes pourraient contribuer à l'urovirulence

# ***Escherichia coli* uropathogène (UPEC)**

# FACTEURS DE L'HÔTE



# FACTEURS BACTERIENS

# Récepteurs Mannose

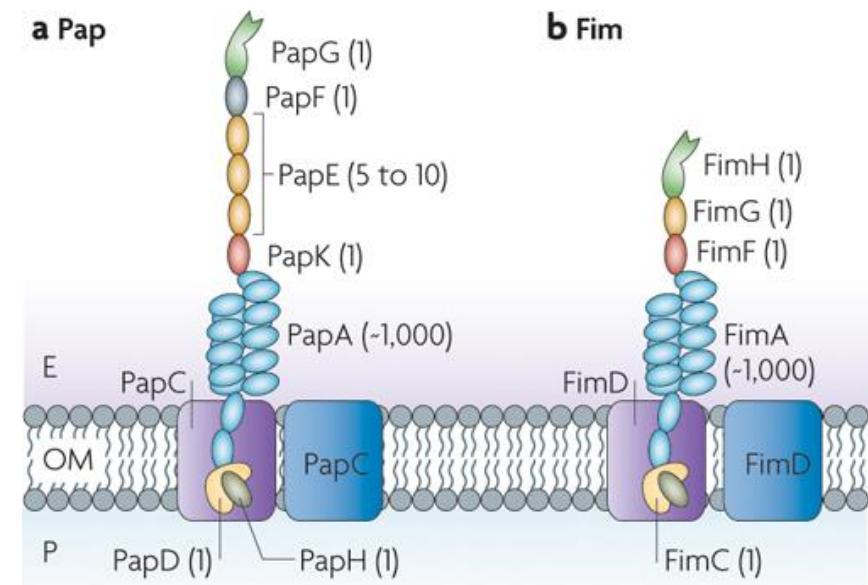
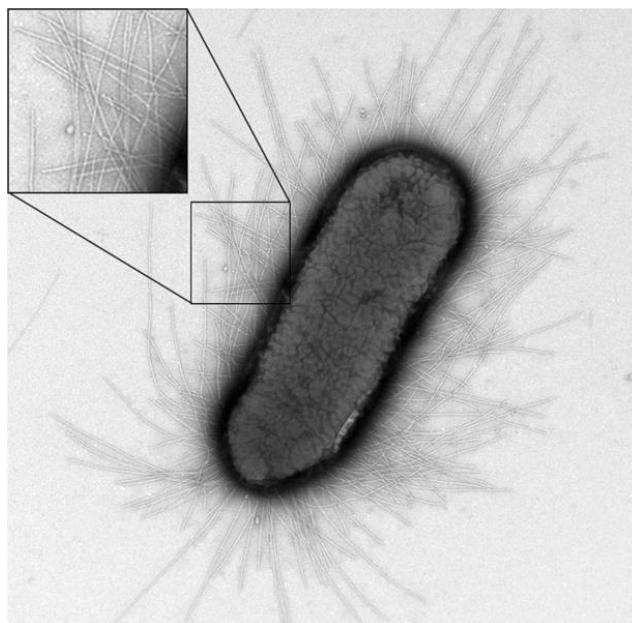
## Adhésion à l'épithélium vésical :

# Pili type 1 « on »

## PapGIII

# *Escherichia coli* uropathogène (UPEC)

- Rôle des adhésines (fimbriae ou pili) :
  - Pili de type 1 (opéron fim – 9 gènes) :  
Adhésion aux récepteurs mannose des cellules de l'épithélium vésical
  - Pili de type P (opéron pap – 11 gènes) :  
Adhésion aux récepteurs  $\alpha$ Gal [1-4]  $\beta$ Gal des cellules eucaryotes



# *Escherichia coli* uropathogène (UPEC)

## FACTEURS DE L'HÔTE

Récepteurs Gal-Gal

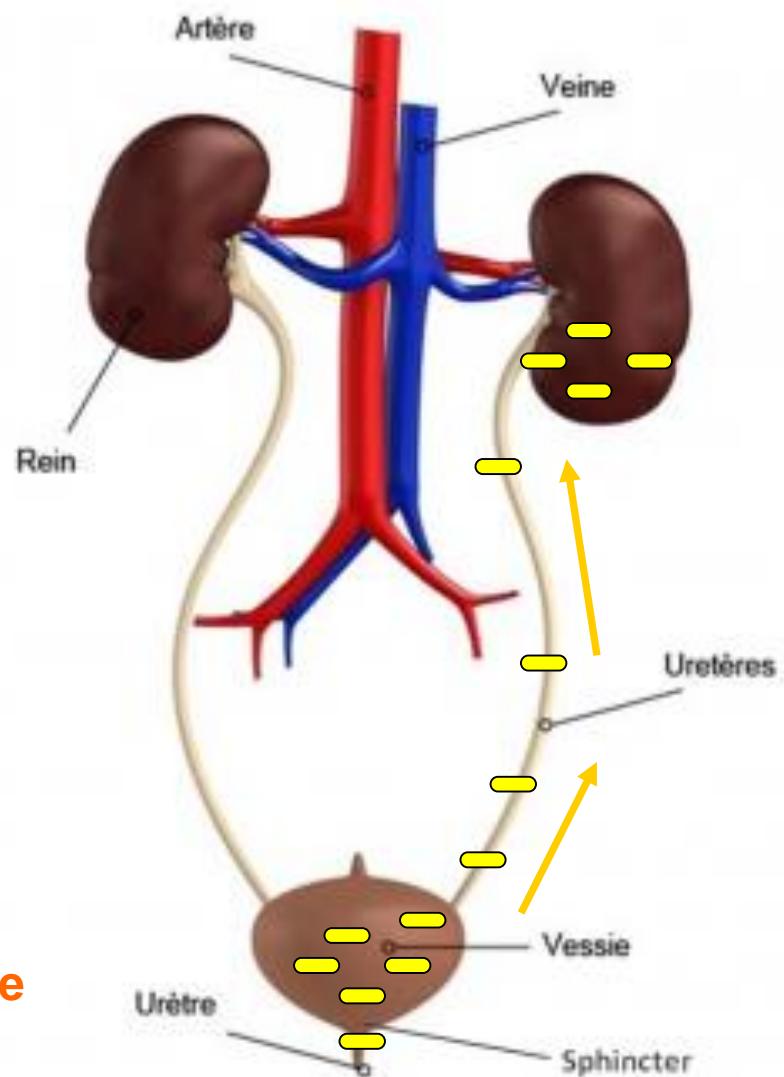
Récepteurs Mannose

## FACTEURS BACTERIENS

PapGII

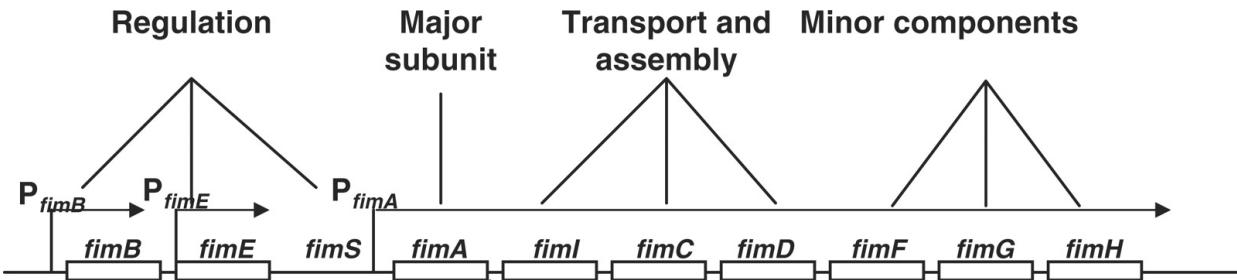
Pili type 1 « off »

infection ascendante  
jusqu'au néphron

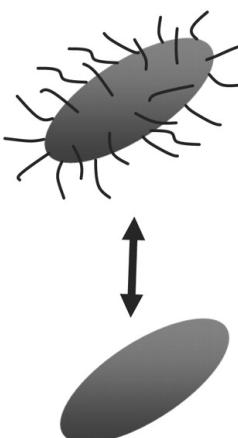
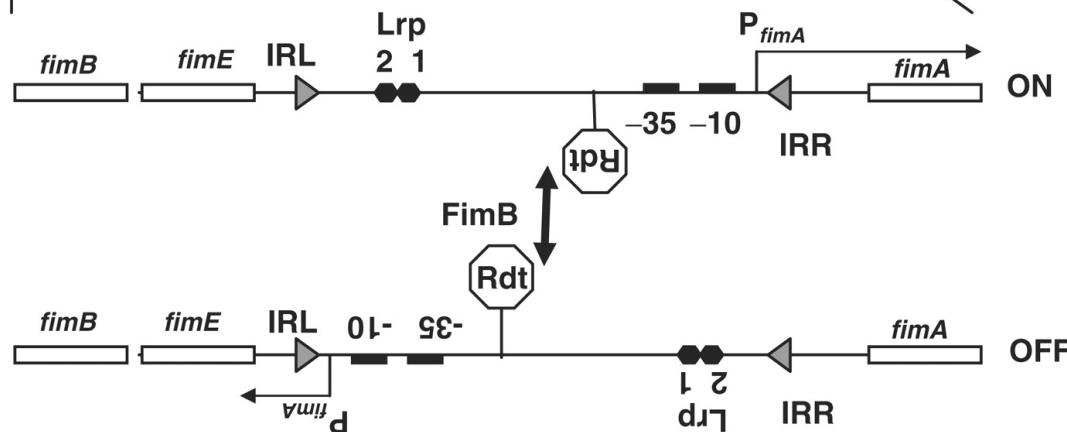


# *Escherichia coli* uropathogène (UPEC)

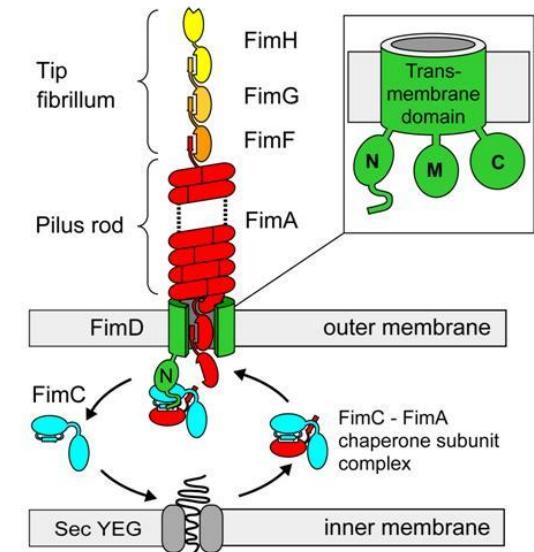
- Pili type 1 switch « ON/OFF » :



invertible *fim* switch (*fimS*)



Dorman et al. Nucl. Acids Res. 2009;37:672-678

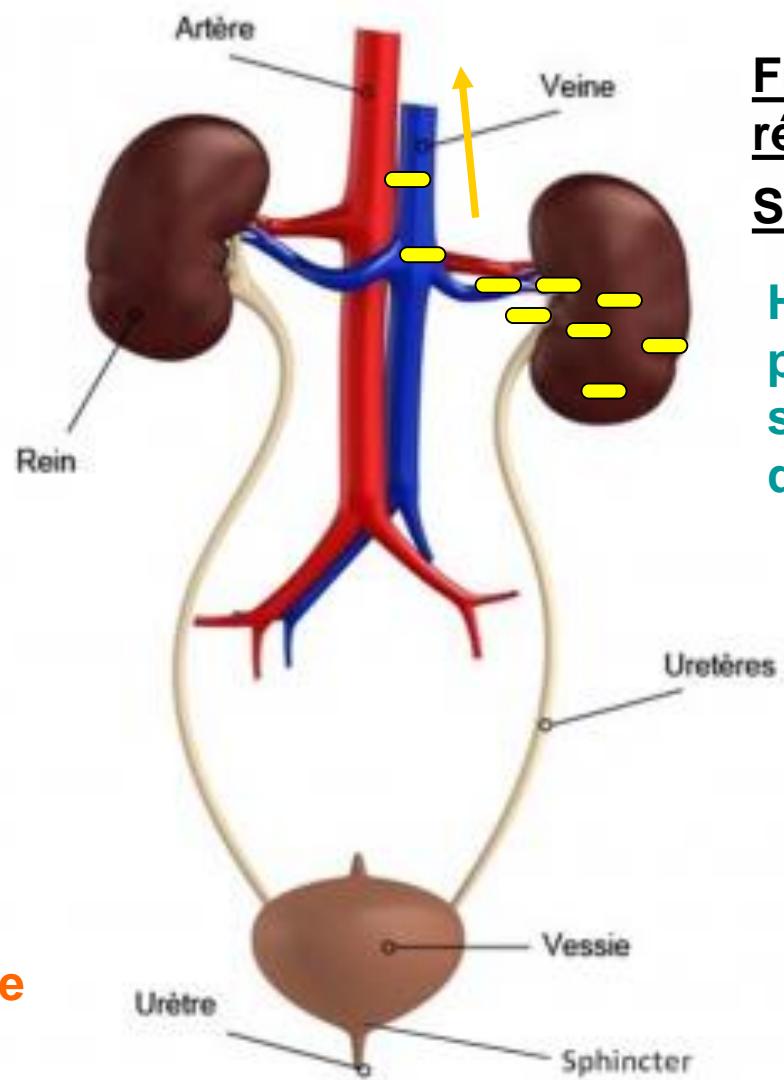


# *Escherichia coli* uropathogène (UPEC)

## FACTEURS DE L'HÔTE

Récepteurs Gal-Gal

Récepteurs Mannose



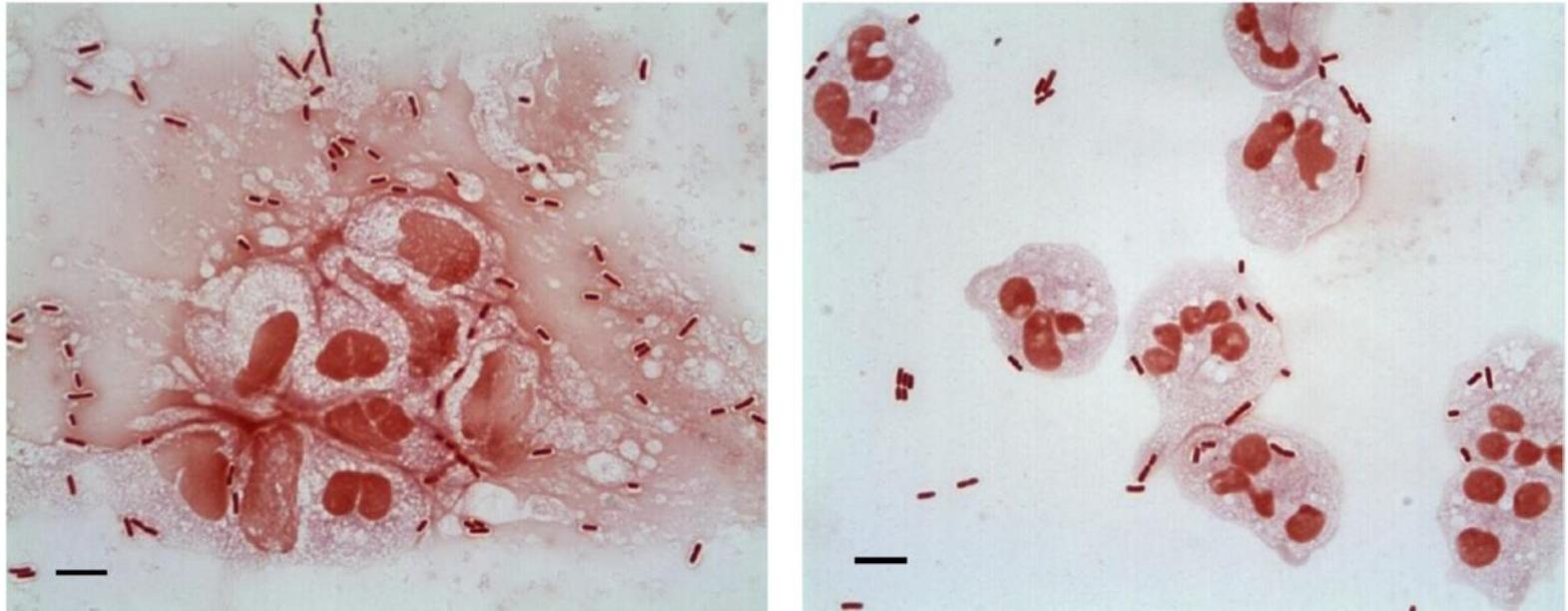
## FACTEURS BACTERIENS

Fragilisation des tissus rénaux

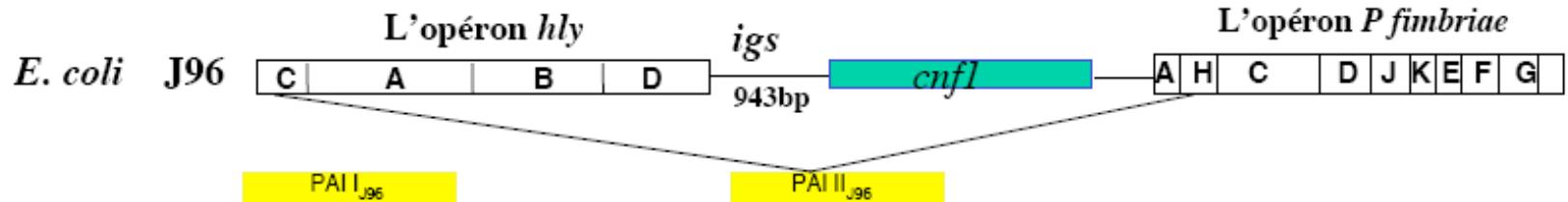
Survie dans le sang

Hémolysine, toxines, protectines, capsule, systèmes de capture du fer...

$\alpha$ -hémolysine (opéron *hly*) : rôle cytotoxique direct sur les cellules rénales, les globules rouges et les polynucléaires neutrophiles

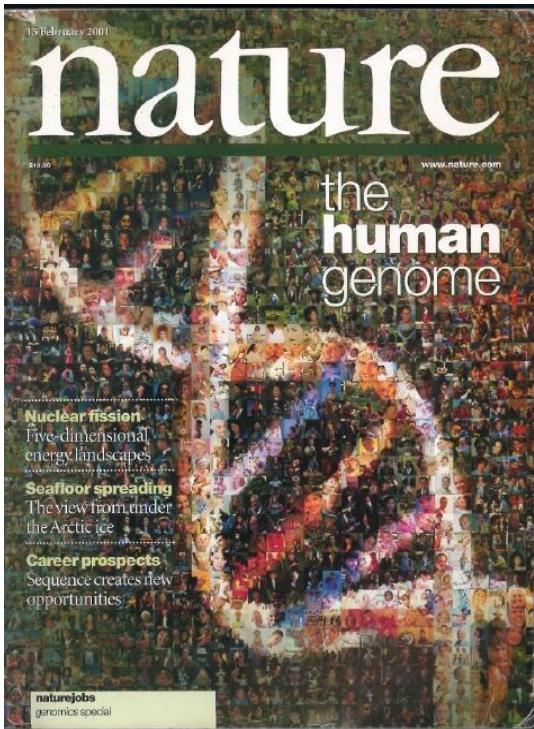


PAI II J96 (110 kb) contenant les gènes *hly*, *papGIII*, *cnf1* & *hra* :



# **Séquençage nouvelle génération (NGS)**

# The human genome project



Février 2001

Octobre 2015  
The 1000 genomes project

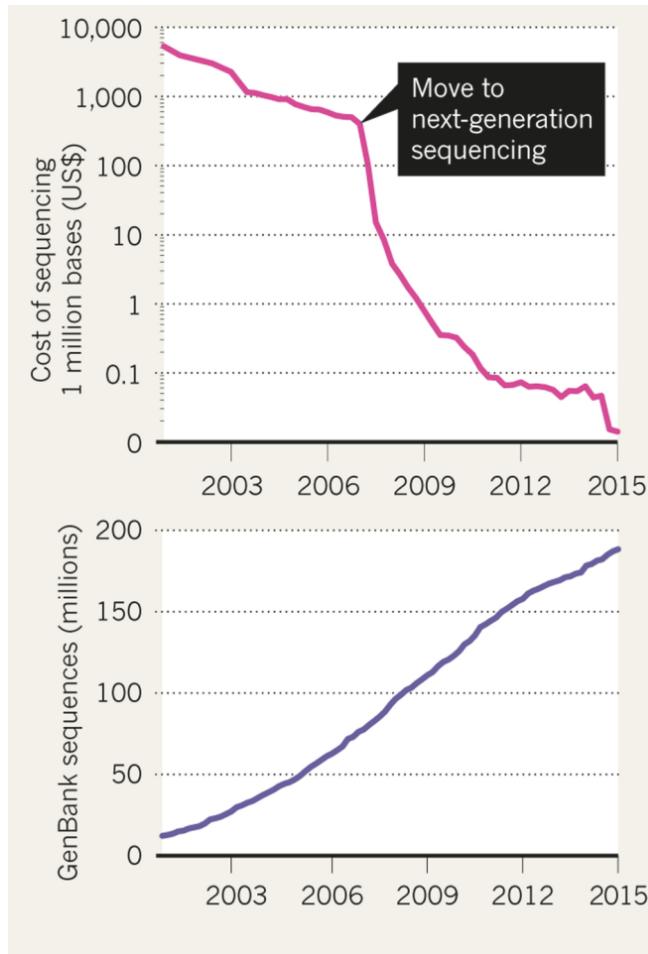


Populations: ● - African; ● - American; ● - East Asian; ● - European; ● - South Asian;

2012-2018 : The 100 000 genomes project

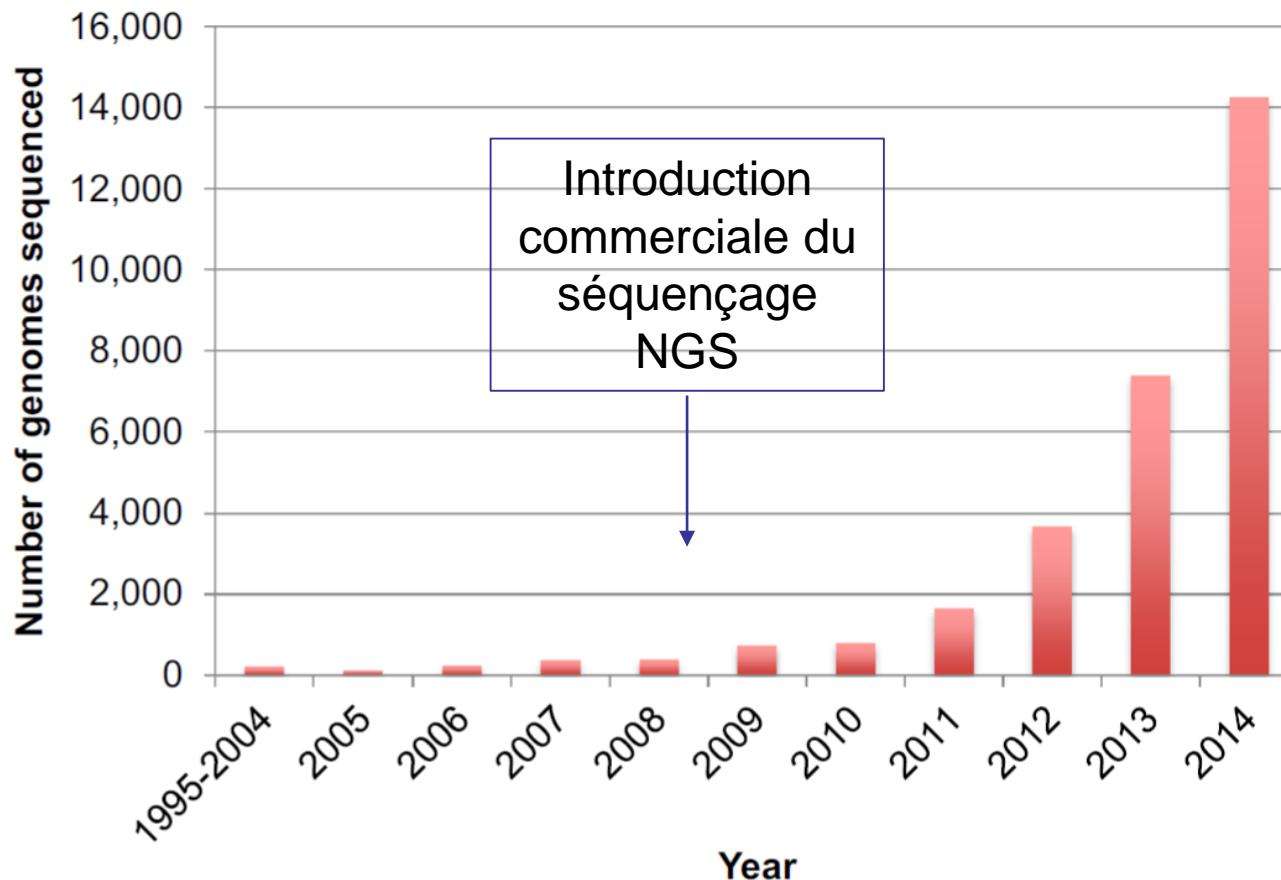
# Les nouvelles technologies de séquençage

## Essor du séquençage nouvelle génération (NGS) et de la bioinformatique



The MinION portable device was used to sequence viral genomes in West Africa's 2014 Ebola outbreak.

# Génomique bactérienne



Nombre de génomes bactériens séquencés chaque année, depuis 20 ans

# Les nouvelles méthodes de séquençage à haut débit

**Objectif** : Séquencer plus vite et de façon plus efficace

**Principe** : Obtention de séquences en très grand nombre, puis assemblage de ces séquences

- **3 étapes :**
  - Préparation de librairies d'ADN
  - Séquençage
  - Analyse bioinformatique

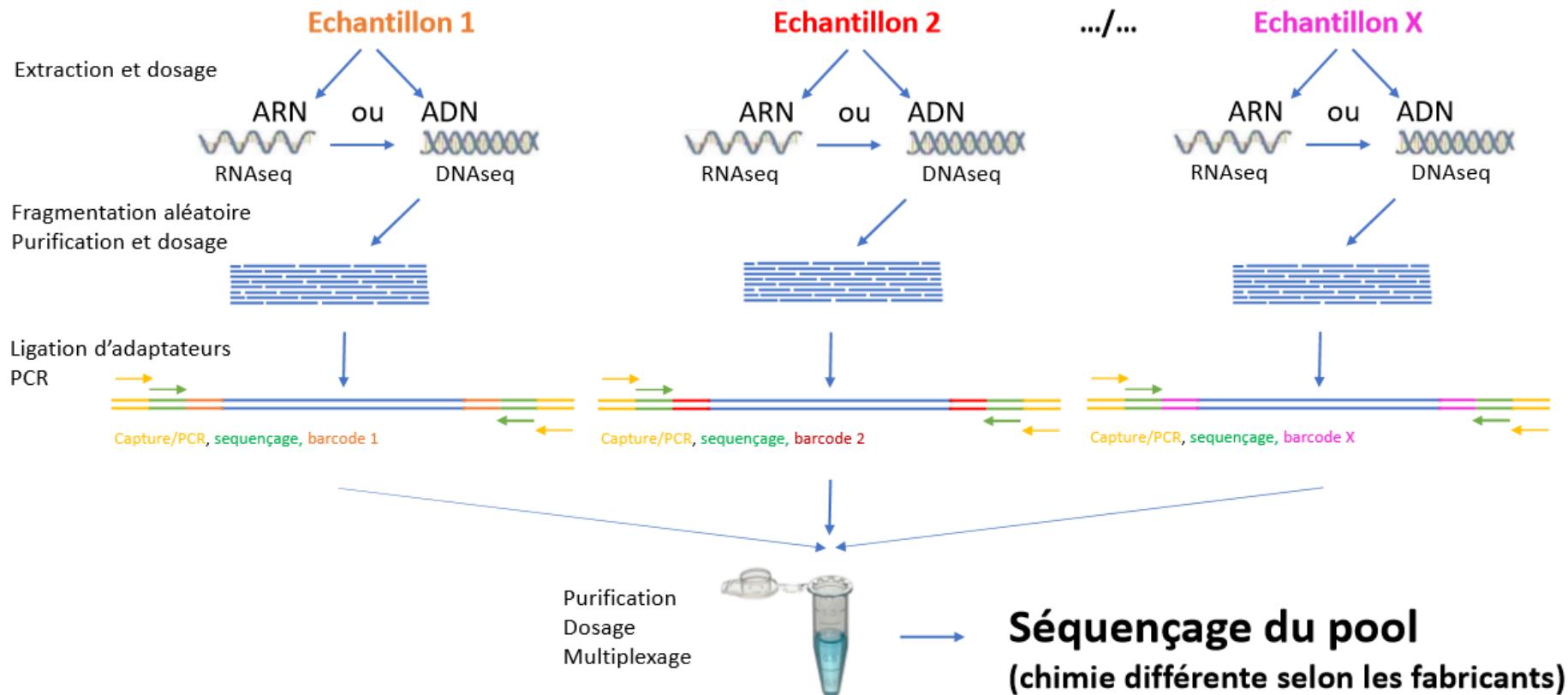
Field revolutionized through the advent of Next Generation DNA Sequencing (NGS) & bioinformatics



# Les systèmes et leurs performances

	Société	Séquenceurs	Nombre de bases/run	Tailles des reads	Taux d'erreur	Durée du séquençage	Coût / Gb
2G	Roche	454	0,3 à 0,7 Gb	400 à 1000 pb	1% (indel)	10 à 23h	9500-40000 \$
	Thermo Fisher Sc.	Ion PGM Ion Proton	0,03 à 15 Gb	200 à 400 pb	1% (indel)	4-7h	80-2400 \$
	Applied Biosystems	SOLiD systems	10-15 Gb	35-75 pb	≤1% (biais AT )	6 à 10j	70-130 \$
	Illumina	MiniSeq	1,6 à 7,5 Gb	1 x 75pb à 2 x 150pb	≤1% (mutation)	7 à 14h	200-300 \$
		MiSeq	0,5 à 15 Gb	1 x 36 pb à 2 x 300 pb	0,1% (mutation)	4 à 56h	110-1000 \$
3G		NextSeq	16 à 120 Gb	1 x 75 pb à 2 x 150 pb	≤1% (mutation)	15 à 29h	33-43 \$
		HiSeq	47-900 Gb	1 x36 pb à 2 x 150 pb	0,1% (mutation)	7h à 6j	7-130 \$
	Pacific Biosciences	PacBio RSII	0,5-1 Gb	~20 kb	10-15% (indel), <1% CCS	30min-4h	1000 \$
	Oxford Nanopore	MinION	10-20 Gb / 48h	jusqu'à 200 kb	5% (2D) 13% (1D)	3min à 48h	50-100 \$

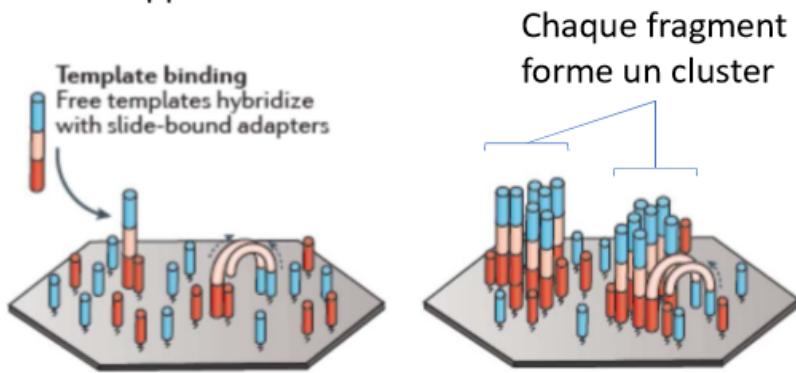
# NGS : Préparation des librairies



# Séquençage 2<sup>ème</sup> Génération

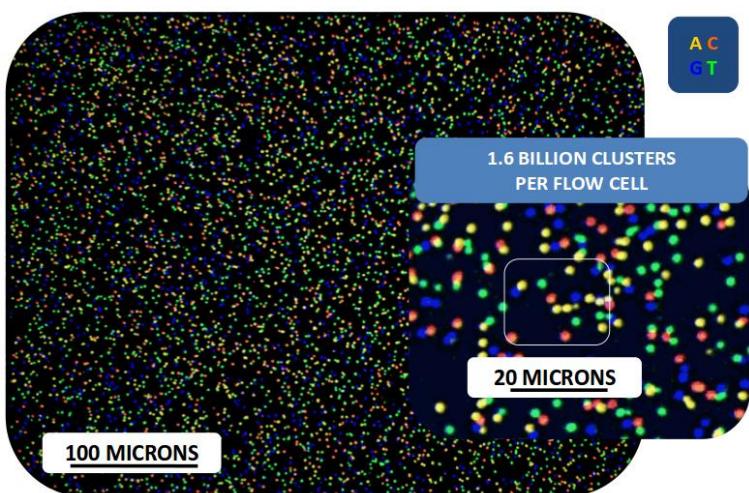
## Ex. Illumina

### 1- Clusterisation par amplification en pont sur support solide

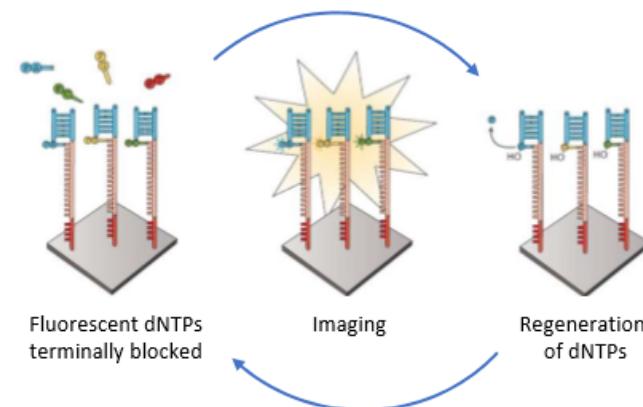


Bridge amplification  
Distal ends of hybridized templates interact with nearby primers where amplification can take place

Cluster generation  
After several rounds of amplification, 100–200 million clonal clusters are formed

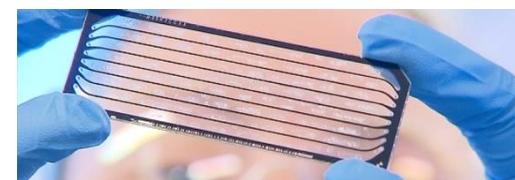


### 2- Lecture cyclique de la séquence pour chaque cluster



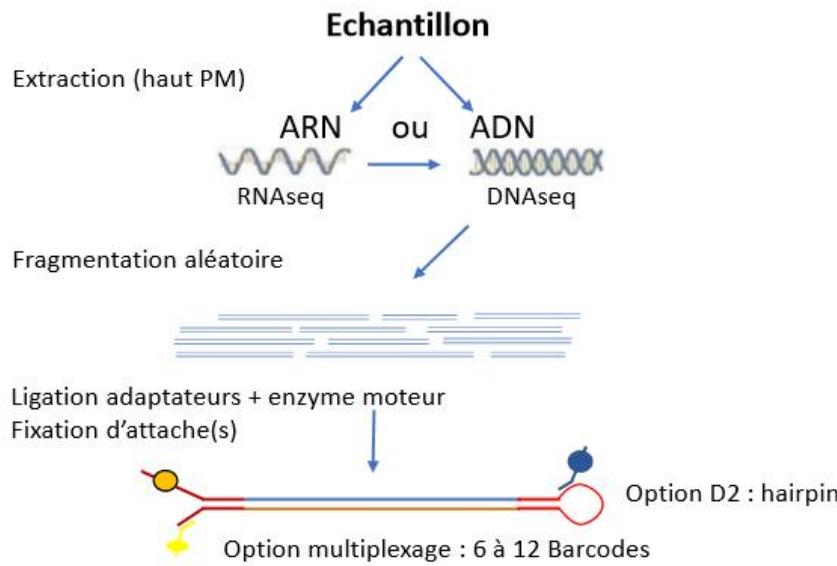
- Cycle 1      Cluster1: A Cluster2: T ...
- Cycle 2      Cluster1: AC Cluster2: TA ...
- Cycle 3      Cluster1: ACG Cluster2: TAT ...

75 à 300 cycles de séquençage = séquence de 75 à 300 bases (+ 1)

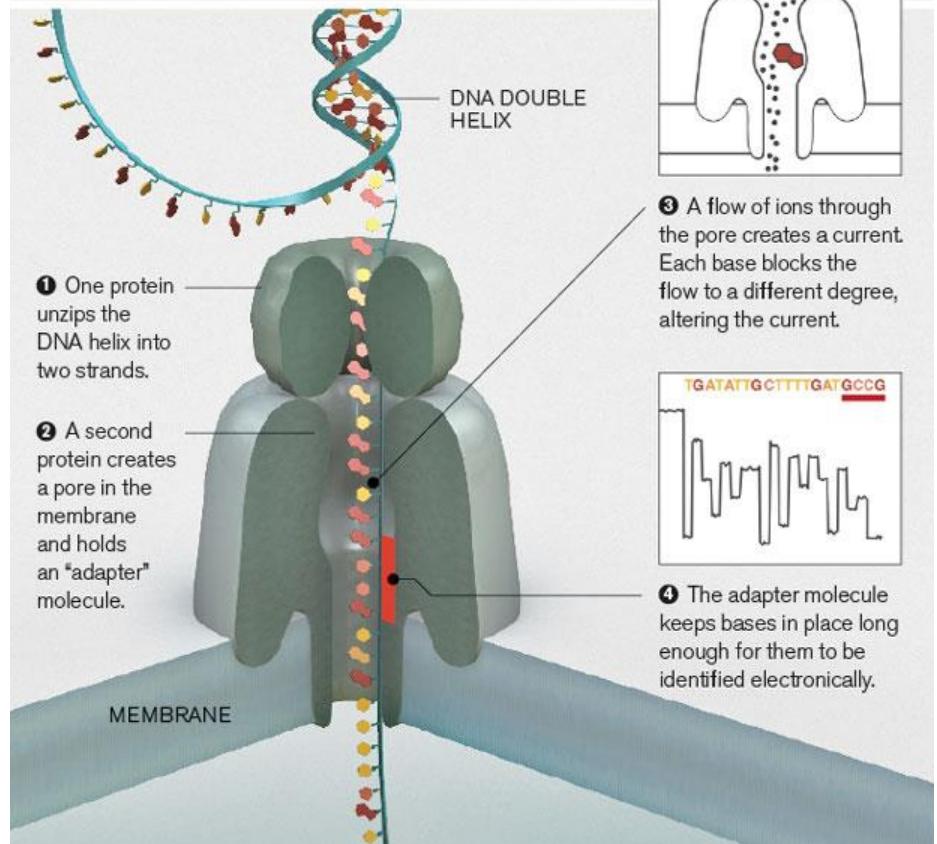


# Séquençage 3ème génération

## Ex. Oxford Nanopore



DNA can be sequenced by threading it through a microscopic pore in a membrane. Bases are identified by the way they affect ions flowing through the pore from one side of the membrane to the other.

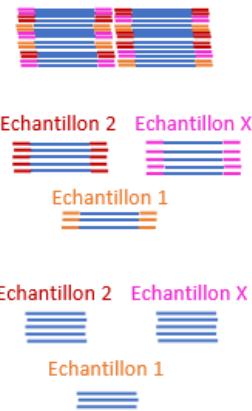


# Analyse informatique des données de séquençage



# Principe de l'analyse des données

## 1- Préparation des reads



## 2a- Assemblage de novo



## 2b- Assemblage par mapping (Reséquençage)



fragments génomiques réassemblés

Analyse des séquences conservées:  
Identification  
Phylogénie et typage (SNP, cgMLST)

Analyse des séquences variables:  
Resistome Virulome  
Mobilome



Identify SNPs in the “core” genome by mapping to a single reference  
-> phylogenetic analysis

*de novo* assembly of accessory genes, followed by functional assignment by BLAST  
-> phenotypic predictions

Development of bioinformatics and standardised database infrastructure for resource sharing

# **SEQUENCAGE GENOMIQUE :**

## **Applications**

- Identifier des facteurs de virulence : pathogénomique
- Etudier l'émergence et la diffusion de pathogènes importants
- Etudier les processus patho-adaptatifs intra-hôte des bactéries

# **Identifier et caractériser les déterminants de la virulence à partir d'une séquence génomique complète**

## **« Pathogénomique »**

### **> 2 Approches :**

- Utilisation d'outils bioinformatiques permettant le décryptage des gènes ou des PAI : recherche de gènes de virulence connus, ou d'homologies avec des facteurs de virulence connus...
- **Approche comparative :** comparaison de souches pathogènes avec d'autres non pathogènes

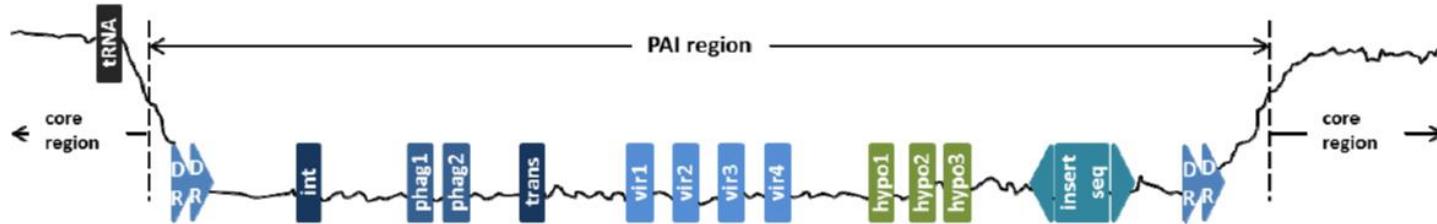
# Décryptage des gènes

## Utilisation de bases de données bioinformatiques :

### Virulence Factor Databases/Servers

MvirDB	A database that contains a collection of publicly available and organized sequences representing known toxins, virulence factors, and antibiotic resistance genes	<a href="http://mvirdb.llnl.gov/">http://mvirdb.llnl.gov/</a>
VFDB	A database that contains all known virulence factors, as well as homologous genes through similarity search	<a href="http://www.mgc.ac.cn/VFs/">http://www.mgc.ac.cn/VFs/</a>
VirulentPred	A web server that predicts virulence factors based on input protein sequences	<a href="http://bioinfo.icgeb.res.in/virulent/">http://bioinfo.icgeb.res.in/virulent/</a>

# Décryptage des PAI



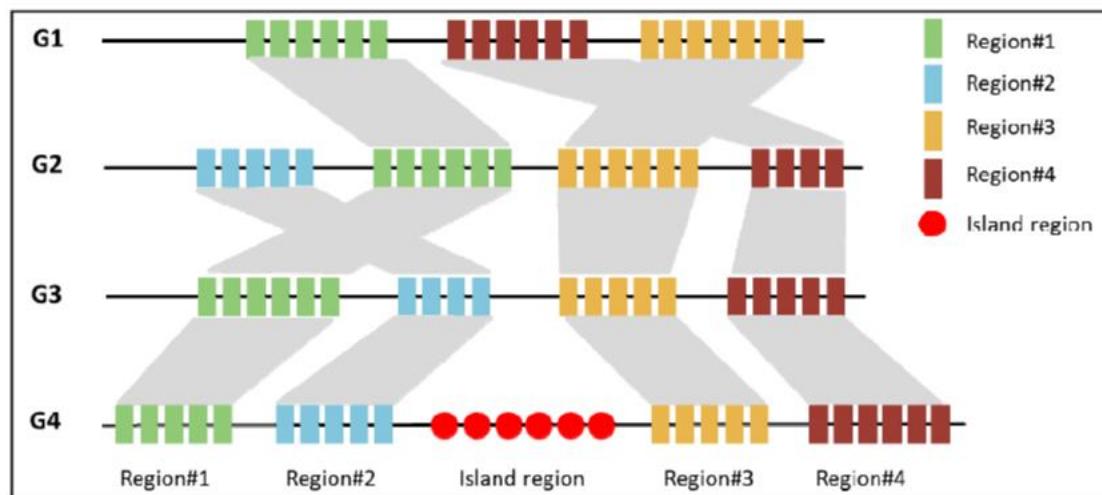
**Table 1.** A list of Pathogenicity Islands (PAI)-associated features and measurement methods.

PAI-Associated Features	Feature Measurement Methods
Different genomic sequence signature	Compute G+C content, GC-skew, codon usage, or other sequence signature tools
Presence of virulence factors	Search through virulence factor database such as VFDB
Presence of mobility genes (integrases, transposes)	Search through NCBI-nr/nt, UniprotKB, Pfam or COG database
High percentage of phage-related genes	Search through NCBI-nr/nt, UniprotKB, Pfam or COG database
Presence of tRNA genes	Use tRNA gene search tool of tRNAscan-SE
High percentage of hypothetical protein genes	Search through NCBI-nr/nt, UniprotKB, Pfam or COG database
Presence of direct repeats	Use repeat finder software REPuter
Presence of insertion sequences	Search through ISfinder database

# Approche comparative

Comparaison de souches proches génétiquement : pathogènes vs non pathogènes

**Figure 2.** A schematic view of genomic region alignment in the comparative genomic based approach for island prediction. Three phylogenetically closely-related reference genomes (G1, G2, and G3) are shown here for the detection of island region in the query genome (G4).



# « Pathogénomique »

## Portrait robot de l'agent pathogène > conséquences cliniques



- La virulence peut être liée à un seul gène (ex. PVL)
- La virulence peut être liée à une mutation dans un gène

- La virulence peut être due à plusieurs gènes agissant à des étapes différentes de l'infection : *quel panel minimum de gènes pour virulence chez un patient ?*
- Présence gène ≠ expression > Transcriptomique
- Gènes de virulence n'ont pas tous été identifiés
- Certaines bactéries virulentes avec peu de gènes de virulence



# Meilleur suivi des épidémies

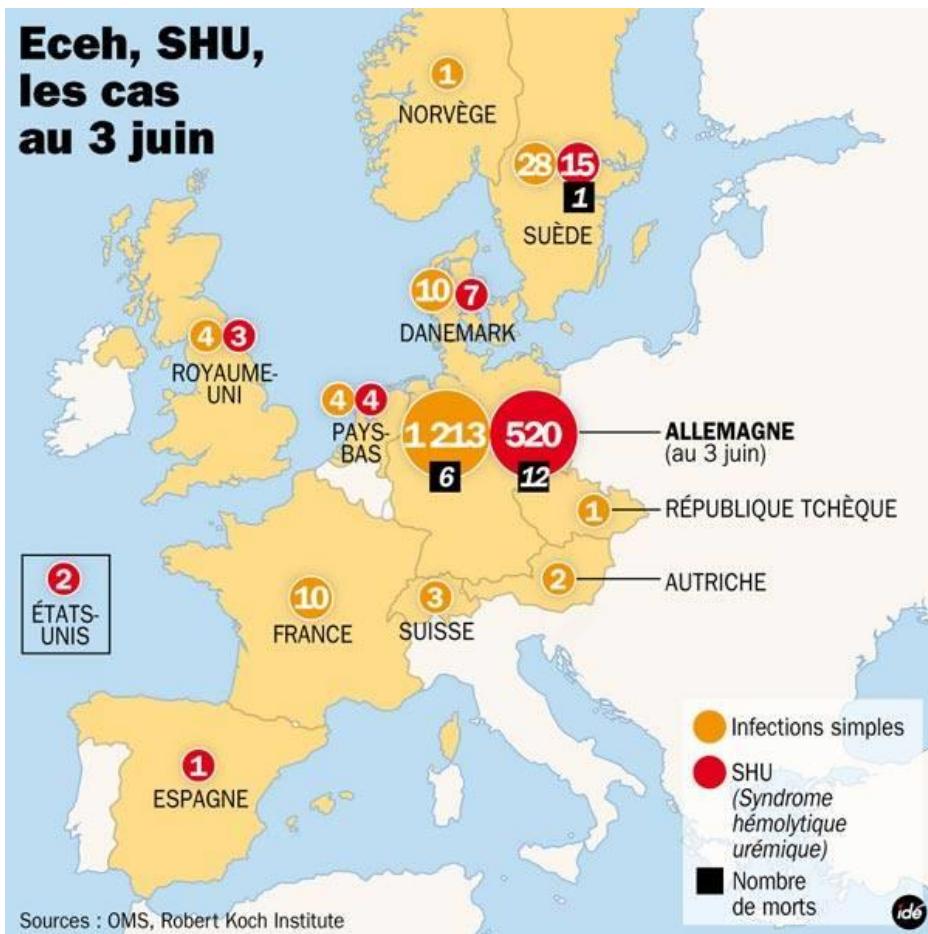
- Suivre la diffusion mondiale et locale des pathogènes importants
- Comprendre les mécanismes d'émergence de bactéries hautement pathogènes et/ou résistantes

## > Analyse phylogénétique détaillée :

- Empreinte génétique précise, basée sur des 100aines de gènes
- Classement des souches en groupes clonaux (cluster) pour identifier les cas groupés – *étude du core génome*
- Recherche de gènes spécifiques de niche – *étude du génome accessoire*

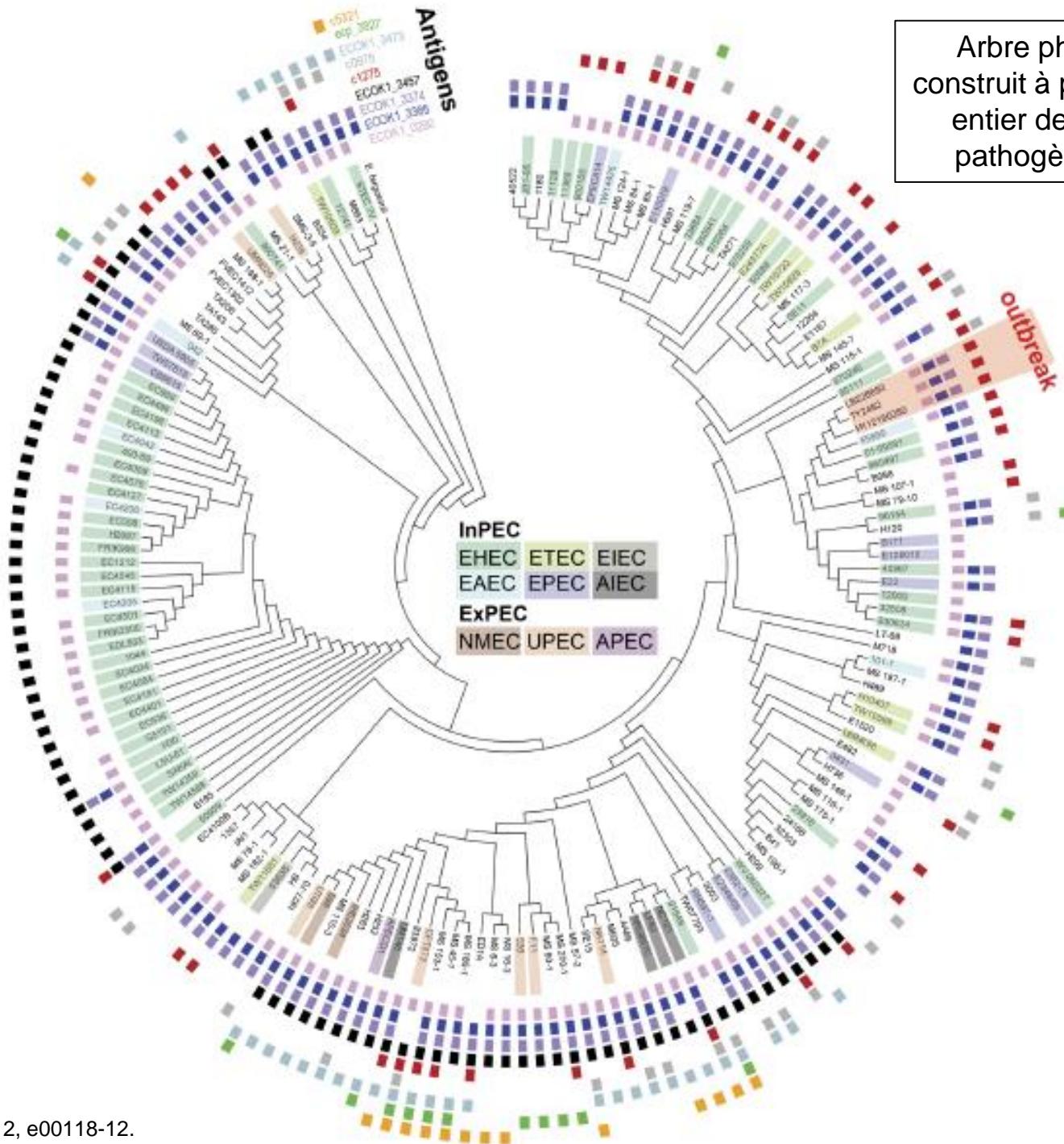
# *E. coli* O104:H4 un pathotype hybride

En Mai-Juin 2011, une épidémie lié à une souche de *E. coli* entéro-hémorragique a eu lieu en Allemagne, puis en Europe



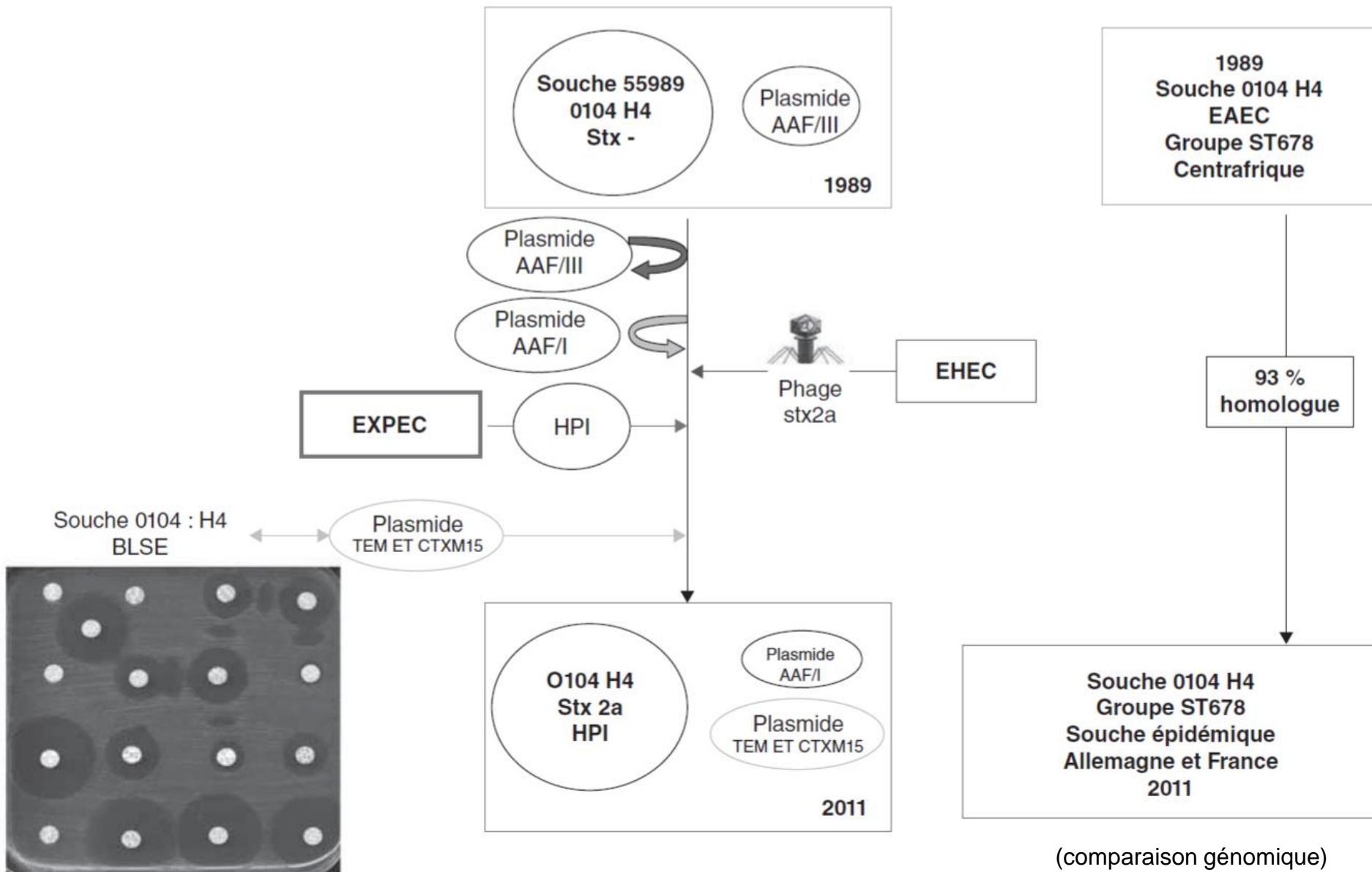
- **Bilan :** 4321 cas (852 SHU, 50 décès) – 12 cas recensés en France
- **Source :** graines germées de Fenugrec importées d'Egypte
- **Souche :** *E. coli* O104:H4 virulence inhabituelle

Jamais impliquée dans épidémie, seulement des cas sporadiques en Allemagne (2001) et Corée (2005)

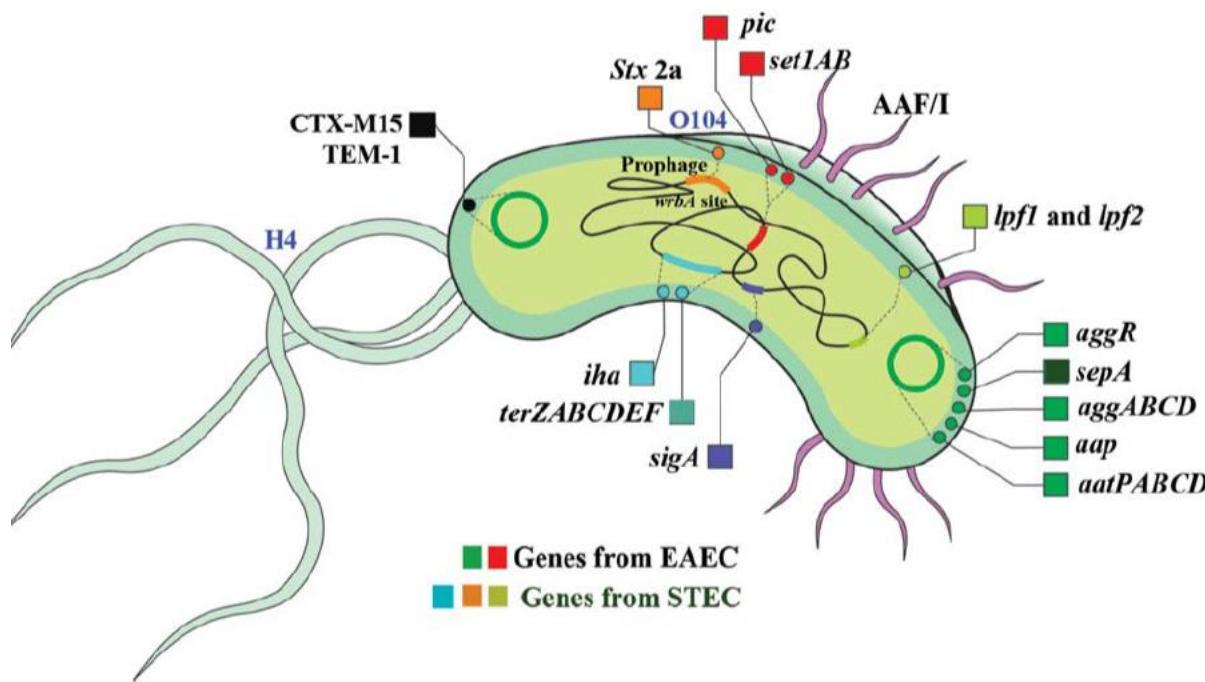


Arbre phylogénétique construit à partir du génome entier de 170 souches pathogènes de *E. coli*

# *E. coli* O104:H4 : un pathotype triple hybride



# *E. coli* O104:H4 : cocktail explosif de facteurs de virulence



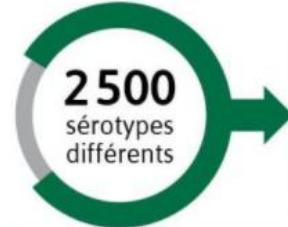
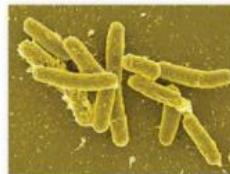
- Adhérence accrue +++ aux cellules épithéliales intestinales
- Action systémique de la shigatoxine > SHU

EAEC : Fimbriae d'adhésion aggrégative de type I (ancrage/inflammation > AAF/III), SPATE proteins (*pic*, *sigA*, *sepA*) : auto-transporteurs impliqués dans colonisation muqueuse et dommages tissulaires, pourraient favoriser la translocation de Stx2 au travers épithélium intestinal, puis dans le sang large panel de FV régulés par le facteur de transcription *aggR*.

EHEC : Shigatoxine Stx2 plus souvent associée à des pathologies sévères que Stx1,  
Iha : adhésion-sidérophore  
Long polar fimbriae (Lpf) impliqué dans adhérence

**■ Les bactéries**

Initialement présentes dans les intestins des animaux sauvages ou domestiques.



Au moins **2 000** d'entre-eux sont pathogènes pour l'homme.

dont la *Salmonella Agona*, en cause dans la contamination chez Lactalis.

**■ La salmonellose...**

Peut être contractée en cas...



... d'ingestion d'aliments contaminés.



... de contact direct avec un animal malade.

\*intensité variant selon le type de salmonelle.

Photo : NHS. Sources : OMS. Santé publique France. Institut Pasteur.

**■ ...et ses symptômes\***

Fièvre



Vomissements

Diarrhée

VISACTU

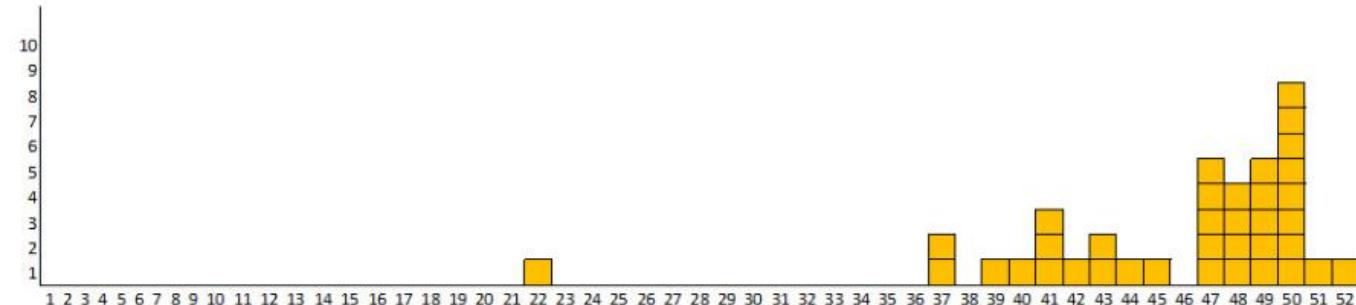
# Salmonellose à *Salmonella Agona*

## • 37 nourrissons atteints fin 2017

- Consommation de laits élaborés sur un même site de production
- Retrait et rappel des produits concernés : aucun nouveau cas depuis



**Figure 1 - Distribution des cas d'infection par *S. Agona* (clone épidémique) par semaine de diagnostic au CNR, France, avril-décembre 2017 (n=37)**





## SCANDALE SANITAIRE

# Lactalis: la salmonelle n'avait pas été éradiquée en 2005

"On peut scientifiquement penser qu'il y a eu contamination à bas bruit de cette usine durant toute la période 2005-2017"

Cette révélation a été rendue possible par la technique dite du "séquençage complet du génome", que les chercheurs ont appliquée à la collection de souches de salmonelle conservées au fil du temps à l'Institut Pasteur.

[...] les travaux effectués dans l'usine en 2017 ont pu réactiver cette salmonelle",



EMBL-EBI

 ENA  
European Nucleotide Archive

Examples

Home Search & Browse Submit & Update Software About ENA Support

Study: PRJEB23906

Genomic analysis of Salmonella Agona outbreaks in France

View: Project XML Study XML

Name  
S. Agona clinical case in infant

Submitting Centre  
Institut Pasteur

Secondary accession(s)  
ERP105685

Description  
Genomic epidemiology of infections of salmonellosis due to Salmonella Agona in France

**Usine de Craon:** une tour de séchage de l'usine est à l'origine de la contamination

## INFO FRANCE INFO.

### L'ADN PARLE

L'Institut Pasteur a recensé, entre 2006 et 2016, 25 cas de nourrissons contaminés à la même souche de salmonelle que celle découverte fin 2017 à l'usine de Craon de Lactalis. L'épidémie de 2005 avait touché 146 nourrissons. Le bilan dépasse donc 200 enfants malades à cause d'une souche unique.

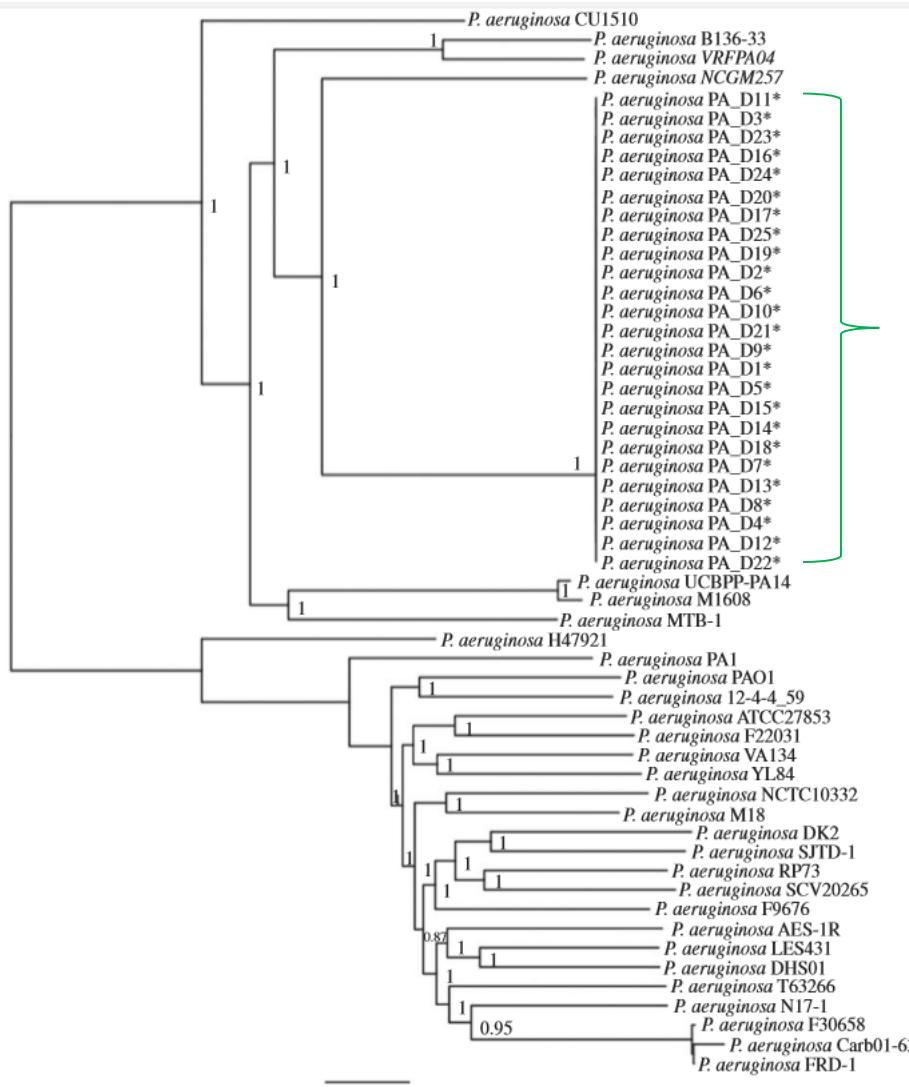
# Meilleurs compréhension de l'évolution intra-hôte des bactéries

## Processus patho-adaptatif

- Etude de la micro-évolution bactérienne en réponse aux pressions de sélection spécifiques de l'hôte, rencontrées au cours d'une infection (autres micro-organismes, réponse immunitaire, stress oxydatif, antibiothérapie...)
  - Apporte des informations sur le répertoire de facteurs bactériens impliqués dans l'interaction avec l'hôte
  - Recherche de mutations patho-adaptatives, fonctionnellement importantes dans la pathogénèse
  - Trouver nouvelles cibles thérapeutiques

# The rapid in vivo evolution of *Pseudomonas aeruginosa* in ventilator-associated pneumonia patients leads to attenuated virulence.

Wang K<sup>1,2</sup>, Chen YQ<sup>3</sup>, Salido MM<sup>4</sup>, Kohli GS<sup>4</sup>, Kong JL<sup>1</sup>, Liang HJ<sup>5</sup>, Yao ZT<sup>2</sup>, Xie YT<sup>6</sup>, Wu HY<sup>7</sup>, Cai SQ<sup>1</sup>, Drautz-Moses DL<sup>4</sup>, Darling AE<sup>8</sup>, Schuster SC<sup>4,9</sup>, Yang L<sup>4,9</sup>, Ding Y<sup>10,9,11</sup>.



- 25 isolats de *P. aeruginosa* collectés chez 4 patients au cours de PAVM dans un hôpital Chinois
- Analyse phylogénétique des 25 génomes : un seul clone épidémique

## Patient 1 – 4 isolats

D1 D20 D38D39

## Patient 2 – 7 isolats

D1 D14 D25 D29 D50 D77D78

## Patient 3 – 8 isolats

D1 D8 D27 D36 D43 D50D57 D68

## Patient 4 – 6 isolats

D1D2D8 D21D24D25

## Evolution adaptative de *P. aeruginosa* au niveau pulmonaire au cours de PAVM

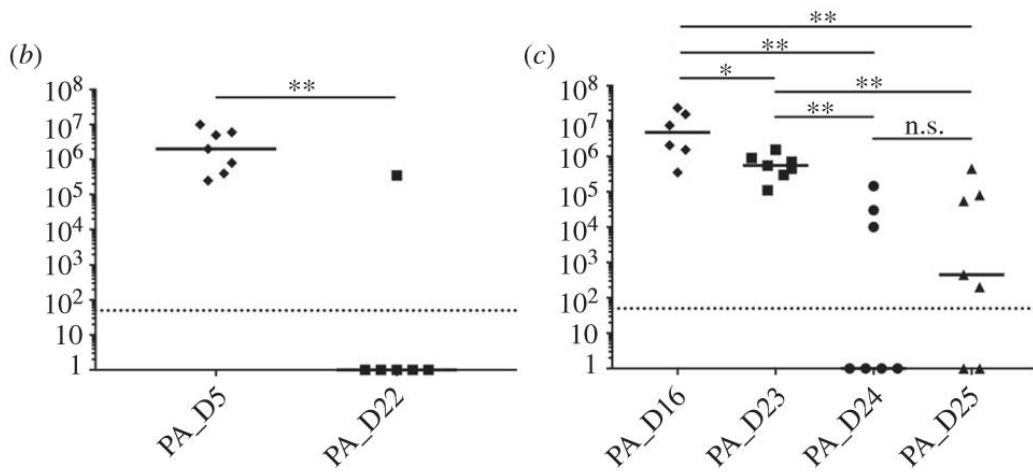
isolate	time	coding region change	amino acid change	type	non-synonymous	remarks on <i>lasR</i> quorum sensing
<i>patient 1</i>						
PA_D1	Day 1	—	—	—	—	functional <i>lasR</i> quorum sensing
PA_D3	Day 20	<i>mpl</i> , 782T > G	Val261Gly	SNV	yes	functional <i>lasR</i> quorum sensing
PA_D8	Day 38	<b>deletion of <i>lasR</i></b> <b>deletion of <i>mpl</i></b>	— —	— —	— —	<i>lasR</i> mutant
PA_D9	Day 39	<b>deletion of <i>lasR</i></b> <b>deletion of <i>mpl</i></b>	— —	— —	— —	<i>lasR</i> mutant

- Apparition de mutations *lasR* aussi chez patients 2, 3 et 4
- Apparition de mutations *mpl* aussi chez patients 2 et 3
- Par ailleurs, apparition de mutations *pvdS* chez patients 2, 3 et 4

### ➤ Convergence d'évolution du génome de *P. aeruginosa* au cours d'infections *in vivo*

- LasR : système du *quorum sensing* impliqué dans la régulation de la virulence
- Mpl : rôle dans la résistance aux β-lactamines ?
- PvdS : synthèse de pyoverdine

- Moindre virulence des isolats collectés tardivement
- Moindre virulence des mutants *lasR* et *pvdS* dans un modèle murin d'infection pulmonaire aigue



**Figure 5.** The abilities of early and late isolates to cause acute lung infections in a murine pulmonary infection model. Mice were infected with each strain by intranasal inoculation at 10<sup>6</sup> CFU per mouse. The total bacterial count recovered from the lungs after 24 h of inoculation are shown in the figure. Comparison between CFU recovered from the lungs of mice infected with (a) PA\_D1 (ancestor stain from patient 1) and PA\_D21 (*pvdS* mutant from patient 2), (b) PA\_D5 (ancestor stain from patient 3) and PA\_D22 (*lasR/pvdS* mutant from patient 3), (c) PA\_D16 (ancestor stain from patient 4) and PA\_D23 (*lasR* mutant from patient 4), PA\_D24 (*lasR/pvdS* mutant from patient 4), PA\_D25 (*lasR/NRPS* mutant from patient 4). Solid lines: the median for each group. Dotted lines: the detection limit of CFU counting ( $n = 6$  for PA\_D21 and PA\_D16, and  $n = 7$  for the other isolates; \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , n.s.  $p > 0.05$ ; Mann–Whitney test).

#### Patient 3

PA\_D5 : souche ancestrale  
PA\_D22 : mutant *lasR/pvdS*

#### Patient 4

PA\_D16 : souche ancestrale  
PA\_D23 : mutant *lasR*  
PA\_D24 : mutant *lasR/pvdS*  
PA\_D25 : mutant *lasR/NRPS*

Evolution rapide de *P. aeruginosa* au cours de PAVM vers une virulence atténuée par l'intermédiaire de mutations *lasR* et *pvdS*

# Conclusion

- Facteurs de virulence bactériens impliqués dans le développement des infections
- Pouvoir pathogène dicté par le panel de gènes de virulence exprimés
- Echanges génétiques entre bactéries et modifications du génome contribuent à la diversité bactérienne et ont des conséquences sur la Santé humaine (résistance aux antibiotiques, gain de virulence, persistance).
- Mécanismes moléculaires à l'origine de l'évolution des bactéries de mieux en mieux compris, grâce aux nouvelles techniques de séquençage à haut-débit (NGS)
- WGS = Analyse de génomes tant sur le plan de la virulence, que de la résistance, ou de leur éventuel lien de clonalité