GenEpi-BioTrain Virtual training 23

Virulence Profiling of *E. coli* Pathotypes

Dr. Jette Kjeldgaard (jetk@food.dtu.dk)
Technical University of Denmark (DTU), National Food Institute

<March 2026>

Intended Learning Objectives

Specific objectives of this session:

1. Understand the challenges in typing of *E. coli* and *Shigella* in the laboratory
2. Understand the biological diversity within the *E. coli* pathotypes (incl. *Shigella*) and the significance for their virulence.
3. Learn about the challenges in quality control of whole-genome sequencing (WGS) *E. coli* data.
4. Analyse and interpret the different sub-typing methods, including MLST, cgMLST, and serotyping of *E. coli*.
5. Apply virulence profiling and typing using simple bioinformatics tools and suggest the resulting pathotype
6. Learn about real-life challenges in tracing emerging clones of pathogenic *E. coli*.

Outline

This session consists of the following elements

Day 1 – March 3

1. Introduction to *E. coli* pathotypes incl. *Shigella*
2. Introduction to typing tools and virulence factor databases
3. Intro to an exercise on *E. coli* subtyping and virulence profiling

Day 2 – March 10

4. Discuss results of typing and virulence profiling
5. Learn about outbreaks with specific *E. coli* pathotypes

Agenda

Agenda	Speakers	Hours
ECDC introduction and rules of conduct		09:00 – 09:10
***1. E. coli* virulence and pathotypes** ○ *E. coli* diagnostic in the laboratory – incl. notifiable results ○ *E. coli* pathotypes; virulence and genetic markers	Dr Cecilia Jernberg (ECDC) Dr Jette Kjeldgaard (DTU)	09:10 – 09:30 09:30 - 10:20
Break		10:20 – 10:30
***1. E. coli* genomics and quality control (QC)** ○ What to be aware of, when working with a diverse organism as *E. coli* ○ Presentation of QualiBact – species-specific QC parameters	João Cardoso, bioinformatician (DTU)	10:30 – 11:00
3. Taxonomic analysis and typing: ○ Concept, application, WGS-based tools	Dr Jette Kjeldgaard	11:00 – 11:20
Break		11:20 - 11:30
3. Virulence gene databases ○ Available databases; specificity, inputs, and outputs ○ Applications of databases	João Cardoso, bioinformatician	11:30 – 11:50
3. Interpretation of virulence gene results ○ Virulence profiling and pathotypes	Dr Jette Kjeldgaard	11:50 – 12:10
3. Introduction to the exercise.	Dr Jette Kjeldgaard	12:10 – 12:20
Q&A and conclusion	Dr Jette Kjeldgaard and João Cardoso	12:20 – 12:30

Intended Learning Objectives of this session

Introduction to *E. coli* pathotypes incl. *Shigella*

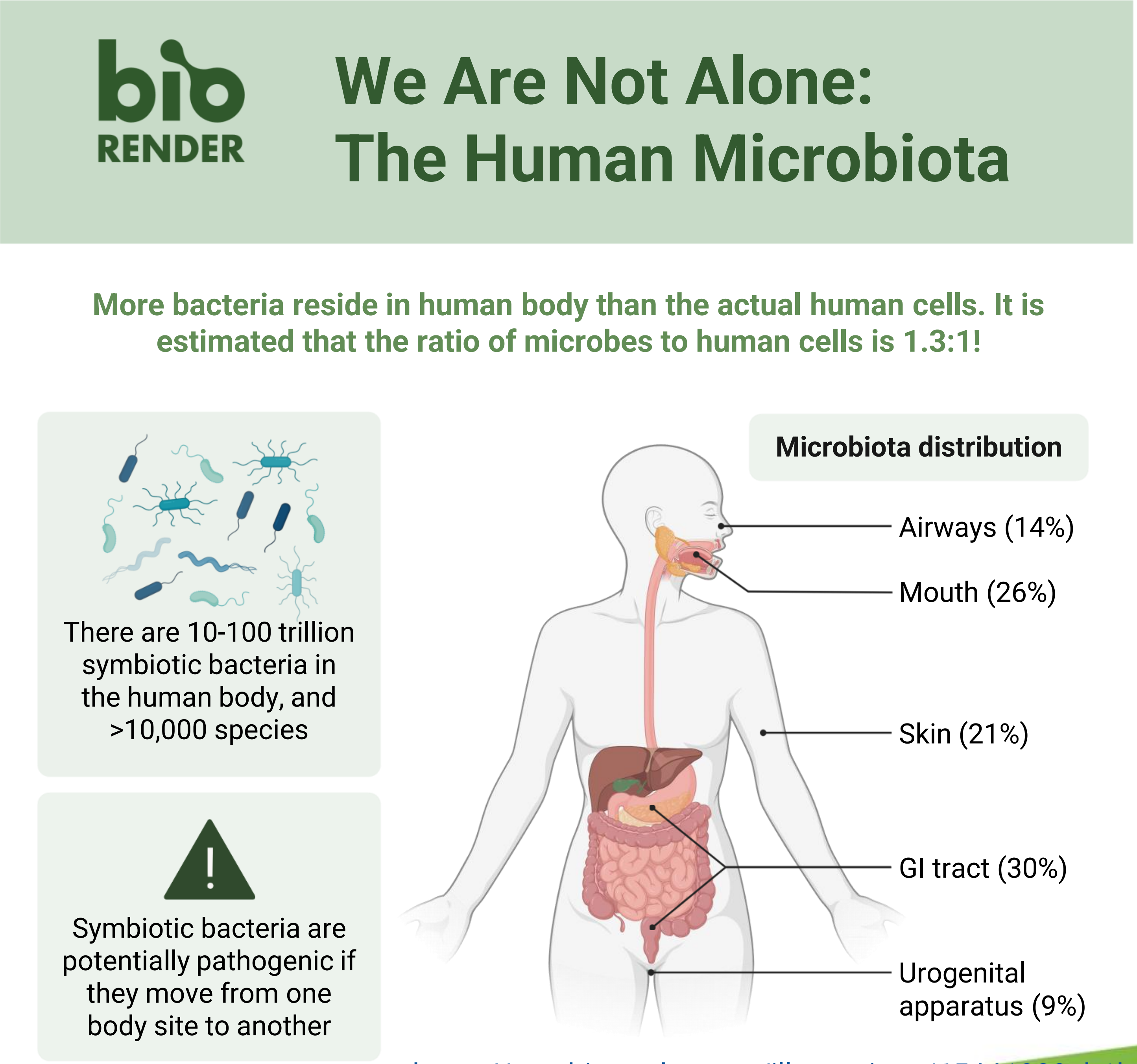
- *E. coli* diagnostic in the laboratory – incl. notifiable results (Cecilia Jernberg, ECDC)
- Which types of infections are caused by *E. coli,* and why?
- What do ‘virulence genes’ cover, and how can we use them to understand the differences in pathogenicity of E. coli?

Microbiota

Normal microbiota
- commensals
- Symbiotics =
  - Opportunists

Pathogens
- Primary pathogens
- Opportunistic pathogens

Difference?
- Virulence factors



We Are Not Alone: The Human Microbiota

More bacteria reside in human body than the actual human cells. It is estimated that the ratio of microbes to human cells is 1.3:1!

There are 10-100 trillion symbiotic bacteria in the human body, and >10,000 species

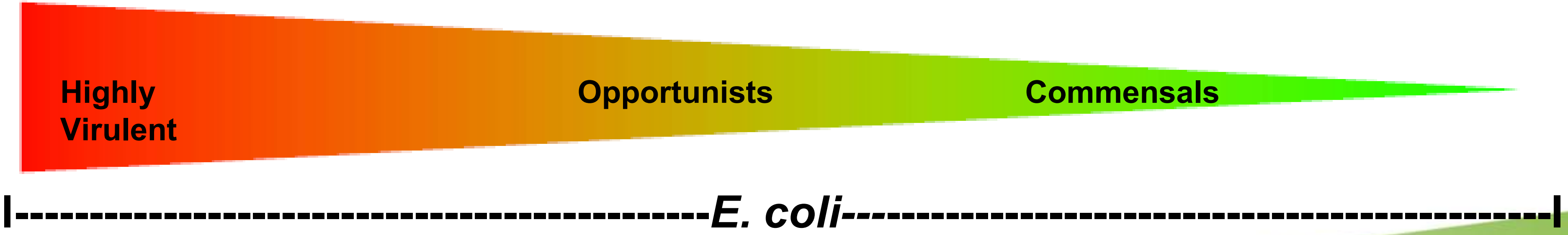
Symbiotic bacteria are potentially pathogenic if they move from one body site to another

https://app.biorender.com/illustrations/65441839cb1beca3631c4457

Pathogenicity and virulence

- Pathogenicity:
  - the ability of a microbe to cause disease in a host
- Virulence:
  - the relative capacity of a microbe to cause disease/symptoms in a host
  - degree of pathogenicity, the severity of disease
- Virulence factor:
  - a constituent of the pathogen that damages the host (e.g. a toxin) or is essential for viability in the host (e.g. capsule or iron-uptake system)

- Not all bacteria are pathogenic

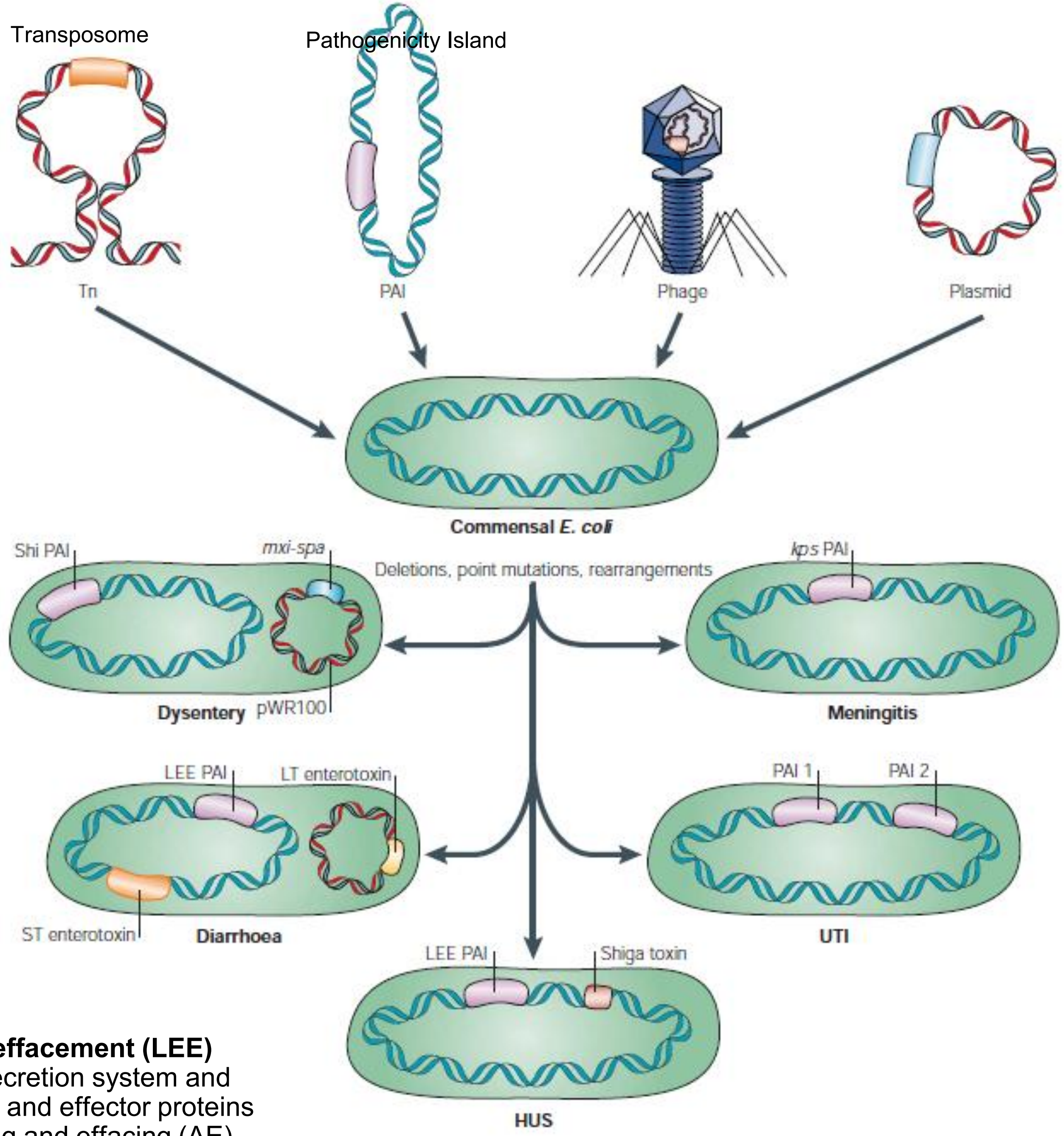


Highly Virulent **Opportunists** **Commensals**

|--*E. coli*---|

Escherichia coli infections

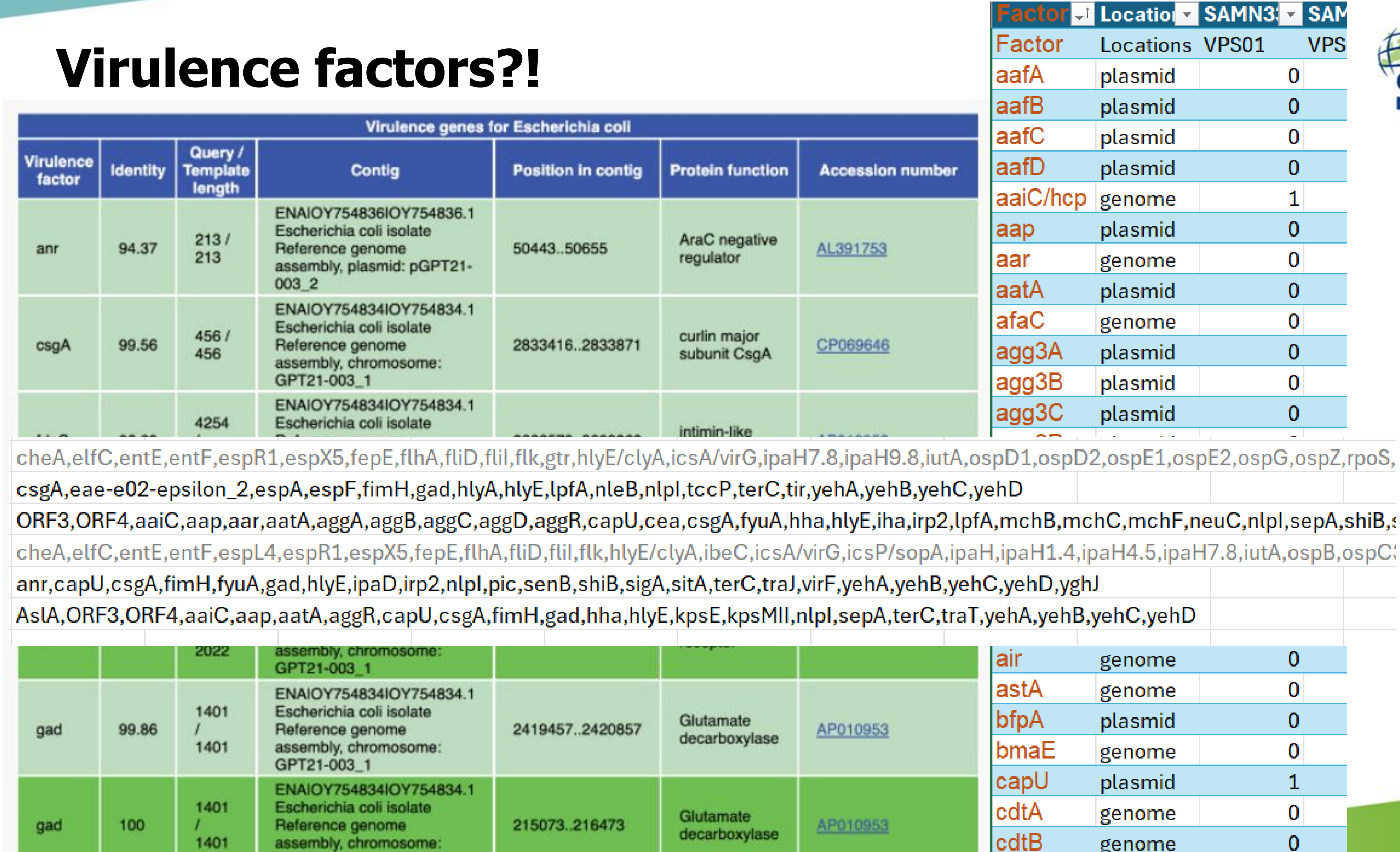
… all depending on the virulence factors



The **locus of enterocyte effacement (LEE)** encodes the Type III secretion system and associated chaperones and effector proteins responsible for attaching and effacing (AE) lesions in the large intestine

https://doi.org/10.1038/nrmicro818

Virulence factors?!



Virulence genes for Escherichia coli

Virulence factor	Identity	Query / Template length	Contig	Position in contig	Protein function	Accession number
anr	94.37	213 / 213	ENAIOY754836IOY754836.1 Escherichia coli isolate Reference genome assembly, plasmid: pGPT21-003_2	50443..50655	AraC negative regulator	AL391753
csgA	99.56	456 / 456	ENAIOY754834IOY754834.1 Escherichia coli isolate Reference genome assembly, chromosome: GPT21-003_1	2833416..2833871	curlin major subunit CsgA	CP069646
		4254	ENAIOY754834IOY754834.1 Escherichia coli isolate		intimin-like	
		2022	assembly, chromosome: GPT21-003_1			
gad	99.86	1401 / 1401	ENAIOY754834IOY754834.1 Escherichia coli isolate Reference genome assembly, chromosome: GPT21-003_1	2419457..2420857	Glutamate decarboxylase	AP010953
gad	100	1401 / 1401	ENAIOY754834IOY754834.1 Escherichia coli isolate Reference genome assembly, chromosome:	215073..216473	Glutamate decarboxylase	AP010953

Factor	Locatio	SAMN3	SAM
Factor	Locations	VPS01	VPS
aafA	plasmid	0	
aafB	plasmid	0	
aafC	plasmid	0	
aafD	plasmid	0	
aaiC/hcp	genome	1	
aap	plasmid	0	
aar	genome	0	
aatA	plasmid	0	
afaC	genome	0	
agg3A	plasmid	0	
agg3B	plasmid	0	
agg3C	plasmid	0	
air	genome	0	
astA	genome	0	
bfpA	plasmid	0	
bmaE	genome	0	
capU	plasmid	1	
cdtA	genome	0	
cdtB	genome	0	

cheA,elfC,entE,entF,espR1,espX5,fepE,flhA,fliD,flil,flk,gtr,hlyE/clyA,icsA/virG,ipaH7.8,ipaH9.8,iutA,ospD1,ospD2,ospE1,ospE2,ospG,ospZ,rpoS,

csgA,eae-e02-epsilon_2,espA,espF,fimH,gad,hlyA,hlyE,lpfA,nleB,nlpl,tccP,terC,tir,yehA,yehB,yehC,yehD

ORF3,ORF4,aaiC,aap,aar,aatA,aggA,aggB,aggC,aggD,aggR,capU,cea,csgA,fyuA,hha,hlyE,iha,irp2,lpfA,mchB,mchC,mchF,neuC,nlpl,sepA,shiB,

cheA,elfC,entE,entF,espL4,espR1,espX5,fepE,flhA,fliD,flil,flk,hlyE/clyA,ibeC,icsA/virG,icsP/sopA,ipaH,ipaH1.4,ipaH4.5,ipaH7.8,iutA,ospB,ospC

anr,capU,csgA,fimH,fyuA,gad,hlyE,ipaD,irp2,nlpl,pic,senB,shiB,sigA,sitA,terC,traJ,virF,yehA,yehB,yehC,yehD,yghJ

AslA,ORF3,ORF4,aaiC,aap,aatA,aggR,capU,csgA,fimH,gad,hha,hlyE,kpsE,kpsMII,nlpl,sepA,terC,traT,yehA,yehB,yehC,yehD

Escherichia coli groups

Two major groups;

- Diarrhoeagenic *E. coli* (DEC)

ETEC - enterotoxigenic *E. coli*
EPEC – enteropathogenic *E. coli*
EHEC (Incl. STEC) - enterohaemorrhagic *E. coli*
EIEC – enteroinvasive *E. coli* – Incl. *Shigella*
EAEC - enteroaggregative *E. coli*
DAEC – diffusely adhering *E. coli*

- Extraintestinal *E. coli* (ExPEC)
  - Uropathogenic (UPEC)
    - Urinary tract infections (UTI’s, sepsis)
  - Neonatal meningitidis (NMEC) *E. coli*

Do not edit
How to change the design

Have you been working with subtyping of E. coli before, incl. serotyping, biotyping and pathotyping?

ⓘ The Slido app must be installed on every computer you’re presenting from

slido

Diarrhoeagenic *E. coli* pathotypes

ETEC - enterotoxigenic *E. coli*

- causing travelers' diarrhea (watery, can be severe)

EPEC – enteropathogenic *E. coli*

- mild to severe (watery) diarrhea, vomiting (in children <2 y; LMIC countries)

EHEC (Incl. **STEC**) - enterohaemorrhagic *E. coli*/Shiga toxin producing *E. coli*

- bloody diarrhea and HUS (kidney failure)
- intense inflammatory response

EIEC – enteroinvasive *E. coli* – Incl. *Shigella*

- Enterocolitis or mild form of bacillary dysentery, similar to that caused by *Shigella* spp
- abdominal cramps, diarrhea (blood and mucus), vomiting, fever

EAEC/EAggEC - enteroaggregative *E. coli*

- persistent diarrhea in children <2 y

DAEC – diffusely adhering *E. coli*

- mild to moderate diarrhea in children >2 y, adolescents

Additionally…
AIEC -Adherent-invasive *E. coli*
EAggSTEC – enteroaggregative shiga toxin producing *E. coli*
….

Virulence factors and infection

How to establish a microbial infection?

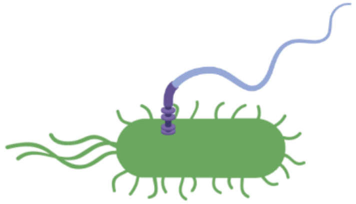
Multiple important factors:

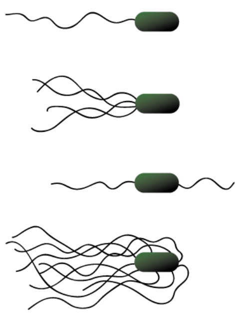
- Adhesion and motility
- Colonization/biofilm formation
- Avoidance of host defence mechanisms
- Nutritional factors
- Invasion
- Toxin or other aggressin production
- Secretion or delivery systems

Adhesion and motility (Flagella, fimbriae and pili)

Filamentous surface structures

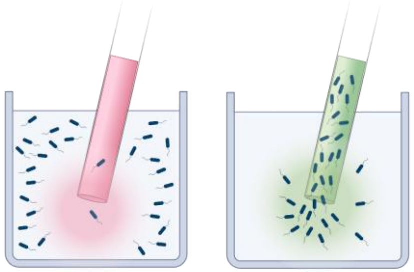
- Help *E. coli* interact with its environment, host tissues, and other bacteria
- Differ in structure, size, composition, and function





Flagella:

- Surface structure; long, thin filament(s); H-antigen
- Used for motility
- Situated in the outer membrane with a hook and a ‘motor’
- Enable bacteria to move toward nutrients or away from harmful substances (chemotaxis)



Common genes involved in flagella

Overview of some of the genes involved in expression of flagella:

Function	**Key Genes**	**Notes**
Master regulator	*flhDC*	Activates transcription of most flagellar operons
Basal body / hook	*fliE, fliF, fliG, fliM, fliN, flgB, flgC, flgD, flgE (hook), flgF, flgG, flgH, flgI, flgJ*	Core rotary motor + hook
Filament	***fliC****	Flagellin (main filament protein)
Motor proteins	*motA, motB*	Proton-driven rotary motor
Export apparatus	*flhA, flhB, fliP, fliQ, fliR*	Secretion system for exporting flagellar components
Regulation	*fliA (σ28), flgM*	*fliA* = flagellar sigma factor; *flgM* = anti-sigma factor

E. coli* H-antigen (flagellar) serotyping primarily relies on identifying variations in the ***fliC* gene.

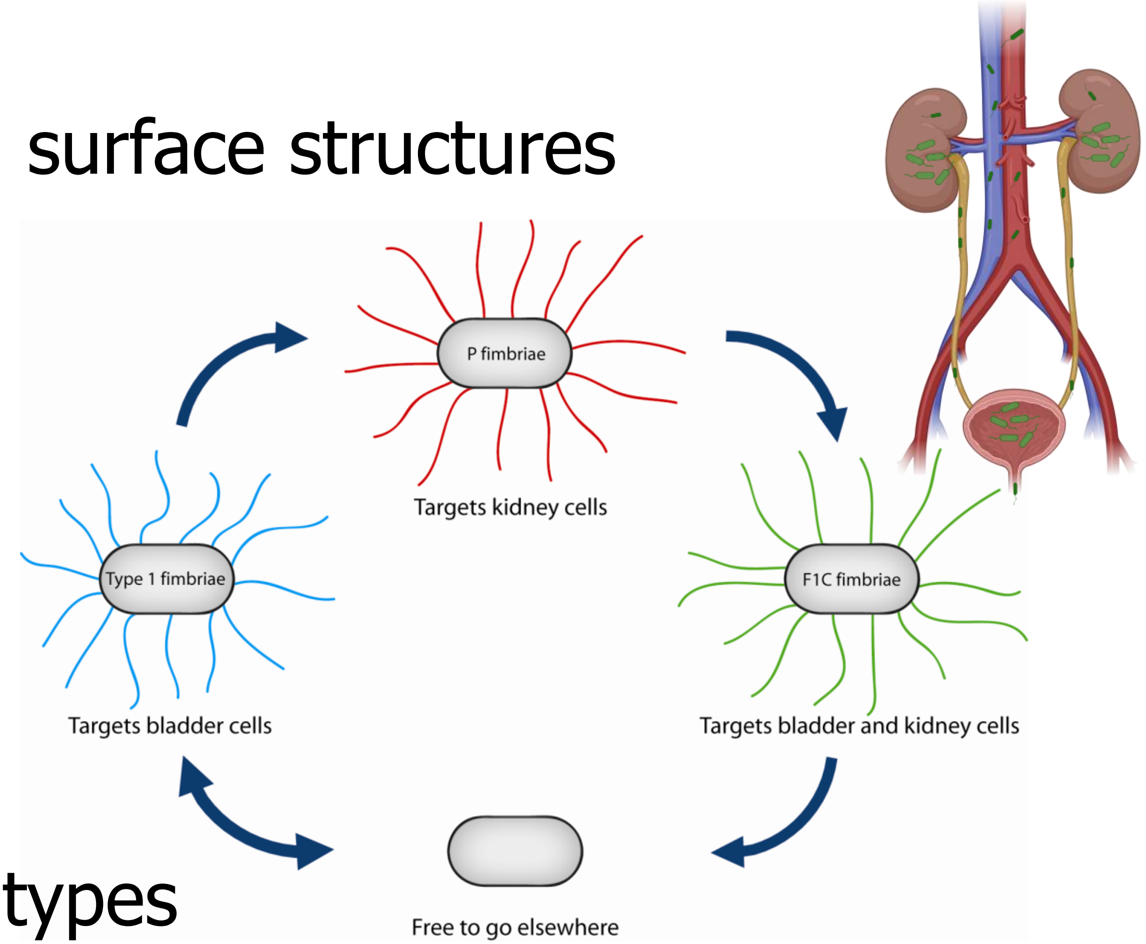
Other alternate flagellar genes like ***flkA***, ***flmA***, and ***fllA*** are used when *fliC* is not expressed.

Fimbriae and pili

Fimbriae and pili are hair-like protein surface structures

They play a crucial role in:

- Adherence to host cells
- Colonization of specific tissues
- Biofilm formation
- Host specificity
- Virulence in different *E. coli* pathotypes



Fimbriae in urinary tract infections

Common genes in fimbriae/pili expression

Pathotype	Key fimbriae / Pili	Major gene clusters	Role in pathogenesis
ETEC	CFA/I, CFA/II, CFA/IV; animal fimbriae (F4/K88, F5/K99, F6/987P, F18)	*cfaA–D, cs1–cs3, fae, fan, fas, fed*	Adherence to small intestine; essential for colonization and delivery of LT/ST enterotoxins
EPEC	Bundle-forming pilus (BFP)	*bfpA–P (*EPEC adherence factor plasmid)	Localized adherence (LA), microcolony formation; early step in A/E lesion development
EHEC/ STEC	Long polar fimbriae (LPF), curli, other adhesins (Iha, Saa)	*lpfA–D, csgA–G, iha, saa*	Colonization of large intestine, biofilm formation, support of A/E lesions (with *eae*/*tir*)
UPEC	Type 1 fimbriae, P fimbriae, S fimbriae, Dr adhesins	*fimA–H, papA–K, sfa, dra/afa*	Adherence to bladder (type 1 fimbriae) and kidney epithelium (P fimbriae), central in cystitis and pyelonephritis
EAEC	Aggregative adherence fimbriae (AAF/I–V)	*aggA–D, aafA–D, agg3, agg4, agg5*	Aggregative (“stacked-brick”) adherence, mucus induction, persistent diarrhea
DAEC	Dr adhesins	*afa/dra, daa*	Diffuse adherence on epithelial cells; associated with UTIs and some pediatric diarrhea
NMEC	Type 1 fimbriae, S fimbriae	*fimA–H, sfa*	Invasion of brain microvascular endothelial cells; contributes to meningitis
ALL	YHD/Yeh fimbriael cluster	*yehA-D*	The yeh cluster is widely distributed in *E. coli*

Avoidance of host defence mechanisms

- Survival through stomach acid
  - Glutamate decarboxylase enzymes *gad*-genes (*gadA-C* + *E*)
  - Arginine decarboxylase (*adiA+C*)
  - Lysine decarboxylase (*cadA*+B)

- Increased serum survival
  - Mainly in ExPEC: *iss*-gene

- Outer membrane antigens, LPS (O-type – somatic)

- Capsule formation (K-antigen)
  - Extracellular, polysaccharide layer (biofilm formation)
  - Cover antigens, prevents recognition
  - Protect against phagocytosis and antibiotics
  - especially important in ExPEC (invasive strains)

Atlas.sund.ku.dk

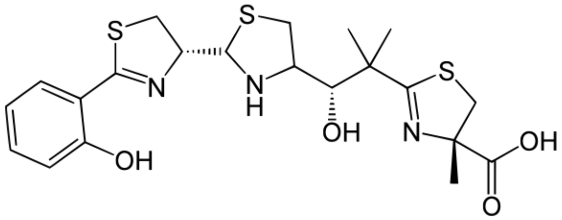
Common capsule and antigen genes in *E. coli*

Capsules are often strain specific and not a defining virulence feature

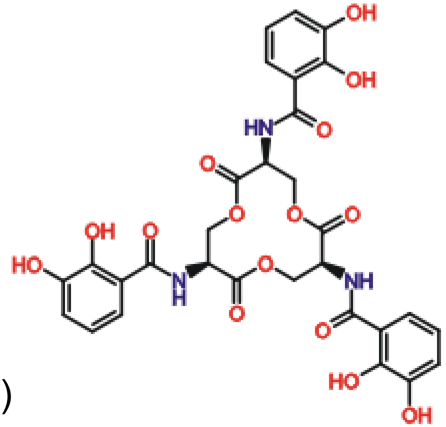
Capsule Group	Main Gene Clusters	Capsule Type Examples	Notes
Group 1	*cpsA–C*, *wzx*, *wzy*	O-antigen	
Group 2	*kpsDMTE*, *kpsF, kpsS, kpsU* + variable region	K1, K5, K12, K100	Most common in pathogenic strains
Group 3	*kpsDMTE*, *kpsCSU*	K10, K54	Often plasmid-encoded
Group 4	*gfcA–D*, *wza*, *wzb*, *wzc*	O-antigen	Found in commensals (EHEC)

Nutritional factors - iron

- Iron is essential for all microbes – and man = competition for iron is fierce
- Inside a host, iron is scarce, poorly soluble or tightly bound by host proteins
- Mechanisms used for iron uptake:
  - Siderophores – iron binding complexes
  - Direct iron (ferrous) uptake
  - Haeme uptake
  - Haemolysin - enzymatic degradation
  - ….

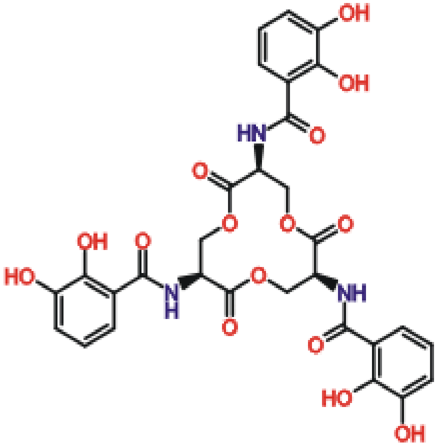

Yersiniabactin of *E. coli*


Haemolysis (Atlas.sund.ku.dk)


enterobactin of *E. coli*

Nutritional factors - Genes for Iron acquisition

- Siderophores (UPEC, EHEC)
  - Specific membrane-located receptor proteins used to transport iron-siderophores into the cell
    - Enterobactin synthesis *entA-F*
      - uses *tonB–exbB–exbD* energy system, *febA-D/G for* transport and *fes* for release
    - Salmochelin (*iro*-genes); Aerobactin (*iuc*-genes, *iutA*);
    - Yersiniabactin (*ybt / irp / fyu* –genes)
- Non-sideophore systems
  - Iron ion transport
    - *feoA/B/C*; *fecA-E/I/R; fur*
- Haemolysin
    - *hlyA* (UPEC + others)
    - *hlyE, clyA* (EIEC/Shigella + others)
    - *ehxA* (EHEC/STEC)



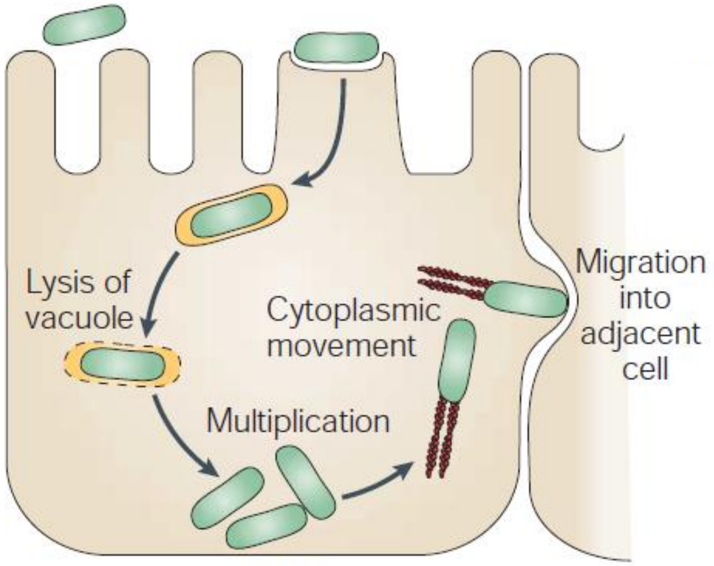
enterobactin of *E. coli*

Invasion/infection – EIEC/*Shigella*

EIEC/Shigella key invasion genes (on the pINV plasmid):

- Mediates active invasion of epithelial cells, allows intracellular replication and spread

- *ipaA, ipaB, ipaC, ipaD* – Invasion plasmid antigens
- ***ipaH*** gene for a Type III-secreted (T3SS) effector protein
- *mxi* genes – T3SS components
- *spa* genes – T3SS secretion apparatus
- ***virF,*** *virB* – transcriptional regulators
- *icsA, virG* – actin-based motility; intracellular spread



EIEC

Endocytic uptake of bacteria and disruption of endocytic vacuoles (phagocytosis)

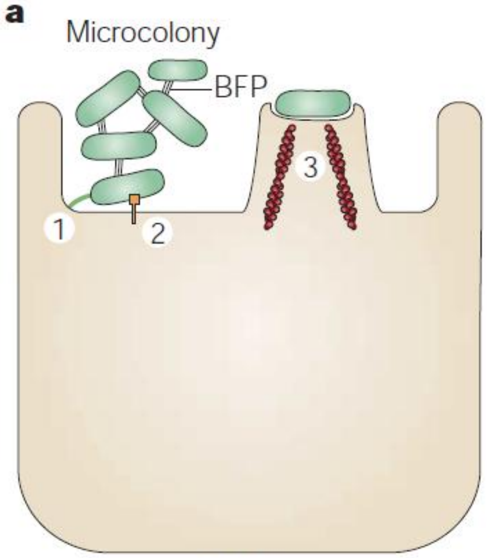
doi.org/10.1038/nrmicro818

Invasion/infection - EPEC

EPEC does not invade deeply like EIEC but performs limited effacement-associated invasion.

Destroy the normal microvillar structure, inducing the characteristic attaching and effacing (A/E) lesion

- EPEC attach via pili (*bfp*), insert intimin-receptor *tir* (1)
- *tir* is responsible for anchoring to the host cell via intimin *eae* and Type 3 secretion system (T3SS) (2)
- Arrangement of host cell cytoskeleton (3), > active ion secretion, > increased permeability



EPEC

doi.org/10.1038/nrmicro818

Invasion/infection - EPEC

- Key factors are on the Locus of Enterocyte Effacement (LEE) pathogenicity island (35-kb PAI)
    - ***eae*** – Intimin (attaching-effacing adhesion)
    - *tir* – Translocated intimin receptor (inserted into host membrane)
    - *esc/sep* genes – T3SS structural proteins
    - *espF, espG, espH, espZ, espB, espD* – effector proteins
    - *espC* - enterotoxin
    - *map* – mitochondrial-associated protein
- *nleA-C* genes – non-LEE encoded effectors
- Regulation of *espC*– *ler*
    - and in EPEC *per* is regulating *ler*

Invasion/infection – EHEC/STEC

- EHEC shares some LEE genes with EPEC, + produces Shiga toxins
- Key factors on the LEE pathogenicity island
  - ***eae*** – Intimin (attaching-effacing adhesion)
  - *tir* – Translocated intimin receptor (inserted into host membrane)
  - *esc/sep* genes – T3SS structural proteins
  - *espF, espG, espH, espZ, espB, espD* – effector proteins
  - Shiga toxins; ***stx1, stx2* – *several variants***

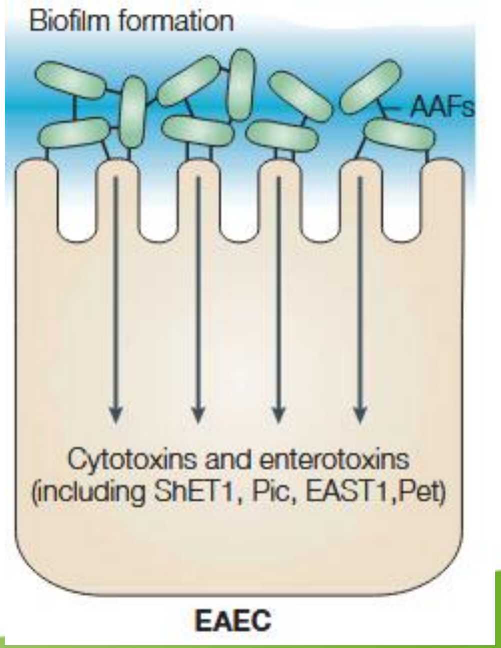
  - *nle* genes – non-LEE encoded effectors
  - *saa* (in some LEE-negative STEC) – autoagglutinating adhesin
  - *toxB* – adherence and epithelial interaction
  - ***ehxA* -** enterohaemolysin

Invasion/infection - EAEC

EAEC are not very invasive, but causes persistent mucosal colonisation, biofilm and inflammation.

The virulence is centered on adhesion, mucus interaction, inflammation and secretion of toxins (entero- and cytotoxins.)

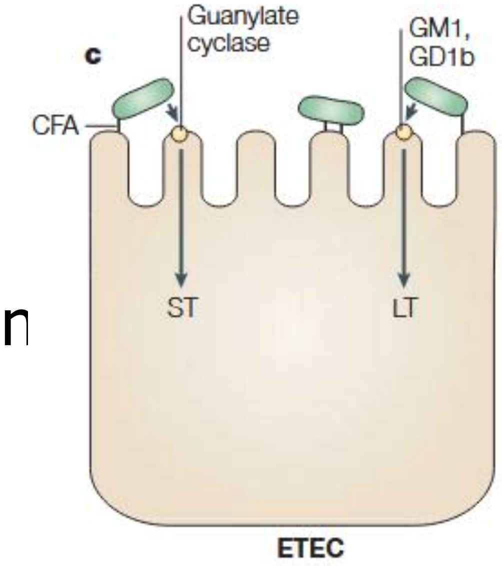
- Aggregative fimbriae - *aggA, aafA, agg3A, agg4A, agg5A*
- Dispersin – facilitating adhesion –*aap/aat, capU*
- Enterotoxin: Enteroaggregative *E. coli* ST (EAST1/*astA*)
- Autotransporter toxins –*pet, pic*
- AggR regulon; *aaiA–aaiV*-genes
- Haemolysin E - *hlyE*



Invasion/infection - ETEC

ETEC shows only limited invasion, relying on toxins (ST, LT)

Possible contributing genes:
tia - Toxigenic invasion locus A
tibA – autotransporter associated with adhesion/invasion
eatA – mucin-degrading protease facilitating epithelial interaction



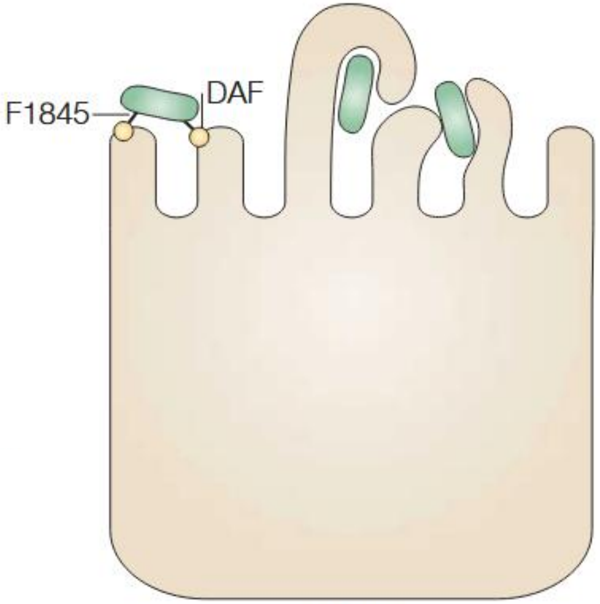
Toxins:
Heat-labile toxin: LT – *eltA, eltB*
Heat-stable toxins: STh and STp – *estA* (STa), *estB* (STb) - (*estIa* in VF)

Invasion/infection - DAEC

Presents a diffuse pattern of adherence > elongated microvilli > long finger-like cellular projections, wrapping the bacteria

Possible contributing genes:

- Dr fimbria: *afa/dra, daaE*
- F1845 is common



DAEC

Invasion/infection - NMEC

Neonatal Meningitis *E. coli* have special ability to survive in macrophages and serum, and to cross the intestinal - and blood–brain barrier

Key specific virulence genes:

- *ibeA, ibeB, ibeC* – Invasion of brain endothelium
- *ompA* – mediates binding to brain endothelial receptors
- *cnf1* – cytoskeletal manipulation
- *kps* genes, *neuC* (K1 capsule) – aids serum resistance and intracellular survival
- *fimH* – adherence to epithelial cells
- *vat* - vacuolating autotransporter toxin

Overview of common pathotype markers

E. coli pathotype	Key virulence markers	Virulence function
EPEC	*eae*, *bfpA*, EAF	*eae* = intimin; *bfpA* + EAF plasmid
ETEC	*elt* (LT), *est* (ST), CF genes (*cfa*, *cs1*–*cs7*)	LT and ST enterotoxins; CF = colonization factors
EIEC	*ipaH*, *ial*, *virF*	*ipaH* is multicopy → high sensitivity; similar to *Shigella*
EHEC/STEC	*stx1*, *stx2*, *eae*, *ehxA*, *lpf*	STEC = *stx*+; EHEC = *stx* + *eae*; *ehxA* = enterohemolysin, long polar fimbriae
EAEC	*aggR*, *aatA/aap*, *aaiC*	*aggR* = master regulator; EAEC is genetically diverse
DAEC	*daa*, *afa/dra* genes	Adhesins associated with diffuse adherence
AIEC	No single marker; sometimes *lpfA*, *fimH* variants used	Requires phenotypic confirmation (adhesion + macrophage survival)

References

Kasper, Nataro and Mobley, 2004; Pathogenic *Eschericia coli* (https://doi.org/10.1038/nrmicro818)

Michelacci V, Tozzoli R, Arancia S, D'Angelo A, Boni A, Knijn A, et al. Tracing Back the Evolutionary Route of Enteroinvasive Escherichia coli (EIEC) and Shigella Through the Example of the Highly Pathogenic O96:H19 EIEC Clone. Front Cell Infect Microbiol. 2020;10:260. https://doi.org/10.3389/fcimb.2020.00260

Torpdahl M, White ED, Schjørring S, Søby M, Engberg J, Engsbro AL, Holt H, Lemming L, Lützen L, Olesen B, Coia JE, Kjelsø C, Müller L. Imported spring onions related to the first recorded outbreak of enteroinvasive Escherichia coli in Denmark, November to December 2021. Euro Surveill. 2023;28(15):pii=2200572. https://doi.org/10.2807/1560-7917.ES.2023.28.15.2200572

Mudau KL, Ntobeng LR, Kalu CM, Tekere M, Pathogenicity and virulence factors of Escherichia coli discovered using next generation sequencing technologies and proteomics. 2025. Frontiers in Bacteriology 4. https://doi.org/10.3389/fbrio.2025.1677775

Robins-Browne RM, Holt KE, Ingle DJ, Hocking DM, 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. doi: 10.3389/fcimb.2016.00141

Virulence profiling

- Not an easy task with out good, curated resources or tools
- Typical markers will give a good indication
- Consider which genes and plasmid are important, not only toxins but also genes for adhesion/invasion
- Consider also regulatory genes
- Remember there is a big genetic overlap between some pathotypes
- We will try it in the exercise

Acknowledgements

The creation of this training material was commissioned by ECDC to the Technical University of Denmark with the direct involvement of Jette Kjeldgaard. Figures were made with BioRender