Verocytotoxin producing *Escherichia coli* (VTEC) diagnostics

Verocytotoxin-producing *E. coli* (VTEC)/ Shiga toxin-producing *E. coli* (STEC)

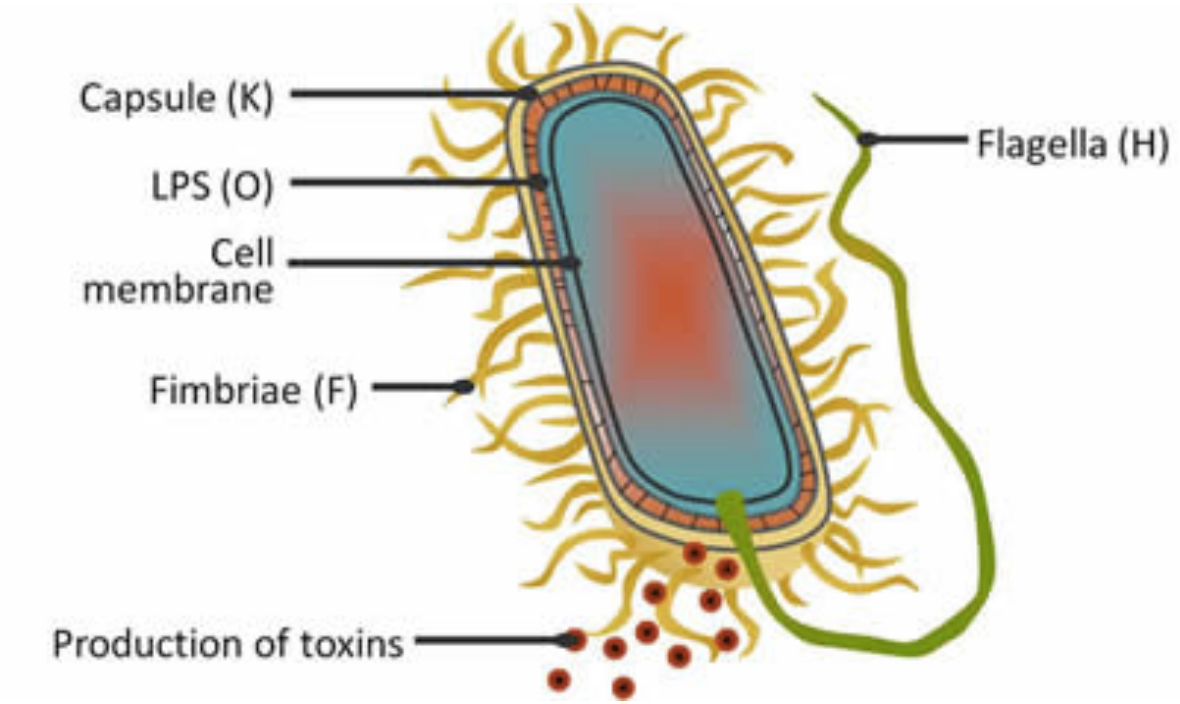
- A gastrointestinal pathogen
- Produces verocytotoxins as well as other virulence factors
- May cause bloody diarrhoea and occasionally Haemolytic Uremic Syndrome (HUS)
- Contracted by ingestion of contaminated food or water or through person-person contact
- In USA, 276,000 VTEC infections occur each year

Previous routine typing of VTEC by the Danish State Serum Institute

- Serotyping
- Examined for..
  - .. β-glucuronidase activity
  - .. haemolysin production
  - .. Verocytotoxin production
  - .. Presence of specific virulence factors, most importantly, verotoxin 1 (*vtx1*), verotoxin 2 (*vtx2*) and intimin (*eae*), which were detected by DNA hybridization
- Further subtyping of the verocytotoxins was carried out by PCR
- Isolates with same serotype and toxin profile compared by PFGE

Serotyping

In serotyping a group of closely related microorganisms are distinguished by a characteristic set of antigens by the use of antibodies.



SerotypeFinder O:H serotyping of *E. coli* using WGS data

O-type is based on the O-processing genes:
wzx, wzy, wzm, and *wzt*, and *wzx/wzy* or *wzm/wzt* pairs

H-type is based on the detection of flagellin genes:
fliC, flkA, fllA, flmA, flnA

Joensen KG. et al. 2015. J.Clin.Microbiol. 53(8):2410-2426.

Evaluating the performance of SerotypeFinder

	Number of genomes		
	for validation	**with detected genes**	**with consistent WGS -and conventional results**
O-typing	601	569 (~95%)	560 (~98%)
H-typing	509	508 (~100%)	503 (~99%)

Center for Genomic Epidemiology

Username
Password
New Reset Login

Home | Services | Instructions | Output | Article abstract

SerotypeFinder 1.1

SerotypeFinder identifies the serotype in total or partial sequenced isolates of E. coli.
Fasta file with test sequence: Test_sequence

View the version history of this server.

The database is curated by:
Flemming Scheutz, SSI
(click to contact)

Select organism
Select multiple items, with Ctrl-Click (or Cmd-Click on Mac)
E. coli

Select threshold for %ID
85 %

Select minimum length
The minimum length is the number of nucleotides a sequence must overlap a serotype gene to count as a hit for that gene. Here represented as a percentage of the total serotype gene length.
60 %

Select type of your reads
Assembled Genome/Contigs*

Isolate File

Name	Size	Progress	Status

Upload Remove

VirulenceFinder

JCM
Journals.ASM.org

Real-Time Whole-Genome Sequencing for Routine Typing, Surveillance, and Outbreak Detection of Verotoxigenic *Escherichia coli*

Katrine Grimstrup Joensen,[a,b] Flemming Scheutz,[b] Ole Lund,[c] Henrik Hasman,[a] Rolf S. Kaas,[a,c] Eva M. Nielsen,[b] Frank M. Aarestrup[a]

National Food Institute, Division for Epidemiology and Microbial Genomics, Technical University of Denmark, Kongens Lyngby, Denmark[a]; Department of Microbiology and Infection Control, Statens Serum Institut, Copenhagen, Denmark[b]; Center for Biological Sequence Analysis, Department of System Biology, Technical University of Denmark, Kongens Lyngby, Denmark[c]

Fast and accurate identification and typing of pathogens are essential for effective surveillance and outbreak detection. The current routine procedure is based on a variety of techniques, making the procedure laborious, time-consuming, and expensive. With whole-genome sequencing (WGS) becoming cheaper, it has huge potential in both diagnostics and routine surveillance. The aim of this study was to perform a real-time evaluation of WGS for routine typing and surveillance of verocytotoxin-producing *Escherichia coli* (VTEC). In Denmark, the Statens Serum Institut (SSI) routinely receives all suspected VTEC isolates. During a 7-week period in the fall of 2012, all incoming isolates were concurrently subjected to WGS using IonTorrent PGM. Real-time

Joensen et al 2014. J. Clin. Microbiol. 52(5): 1501-1510

Center for Genomic Epidemiology

Username
Password
New Reset Login

Home | Services | Instructions | Output | Article abstract

VirulenceFinder 1.5

View the version history of this server.

The database is curated by:
Flemming Scheutz, SSI
(click to contact)

Select species

- E. coli
- Enterococcus
- S. aureus

Select threshold for %ID

90 %

Select minimum length

60 %

Select type of your reads

Assembled Genome/Contigs*

Isolate File

Name	Size	Progress	Status

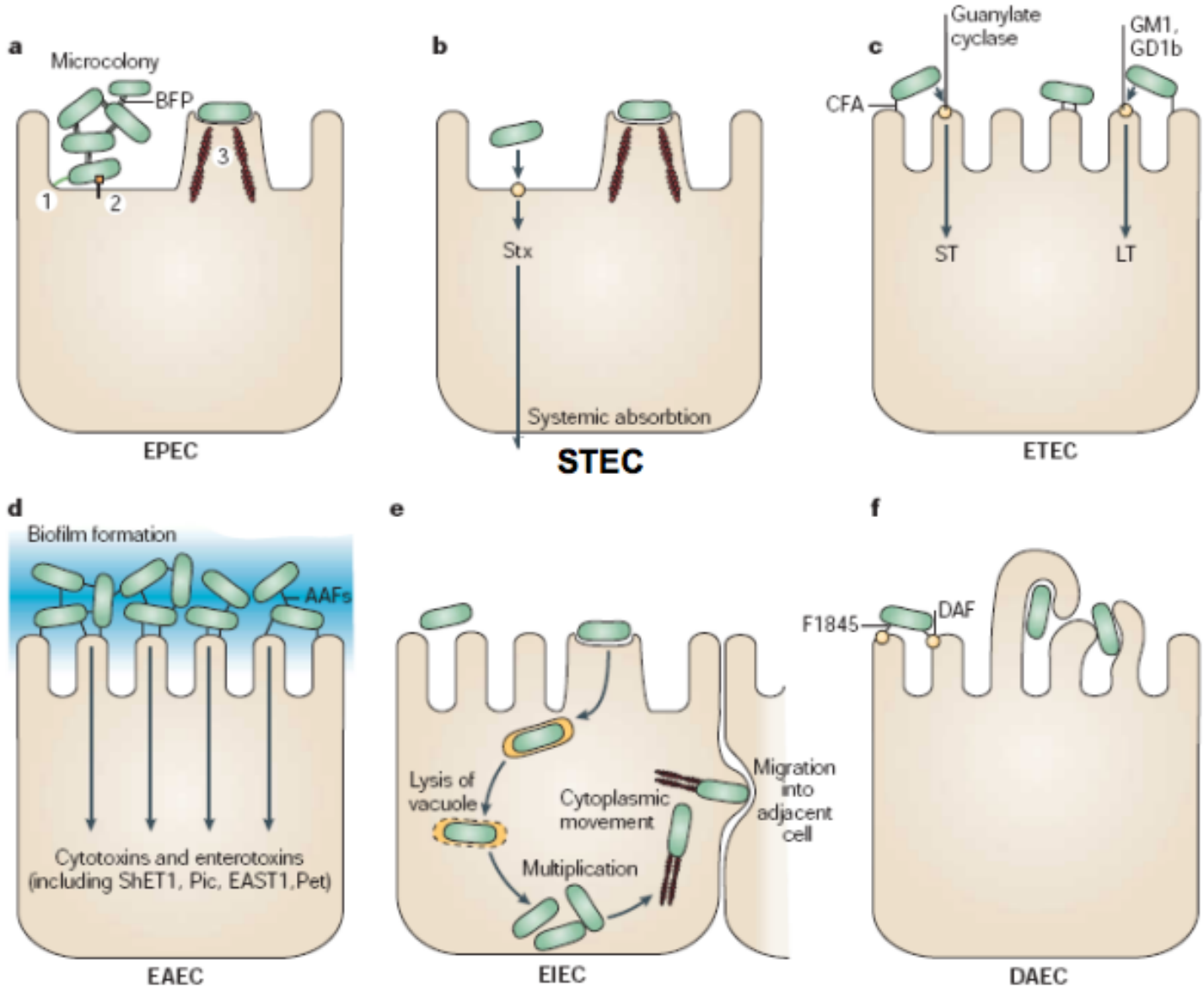
Upload Remove

Database content

	# Genes	# Alleles
E. coli	102	957
S. aureus	164	230
Enterococcus	26	104

https://cge.cbs.dtu.dk/services/data.php

Classification of *E. coli* causing diarrhoea, DEC



Kaper et al. 2004. Nature Reviews Microbiology 2, 123-140

Simplified categorisation that we will use during exercise

- Non-pathogenic: Few or no obvious virulence factors.

- UPEC: Adhesion factors to avoid being flushed away from the urinary tract, siderophore proteins (*iroN*). Presence of Pap (P) fimbriae (*papG* adhesion) can be a sign of increased virulence.

- VTEC, STEC (EHEC): A hall-mark virulence factor is the Shiga toxin (*stx2A* and *stx2B*).

- EAEC: Many different virulence factors, especially aggregative adherence fimbriae (AFFs) located on a 100-kb pAA plasmid, mycolycins such as those encoded by the *pic* gene and toxins (*pet* and *astA*). The regulator encoded by *aggR* also located on the pAA plasmid is coordinating the virulence factors.

- ETEC: Production of Heat-Stabile Toxin (ST) or Heat-Labile Toxin (LT). The former can be encoded by the *sta1* or *stb* genes and the latter by the *elt* or *ltcA* genes.

BLAST-based CGE finder tools

Draft assembly

```
>Contig_1
TGGATTCTATAGAAATTGATGAGGACTTGTG
TCAAGTCACCCAAAAAGTCGTGAAGCCTTTT
CAGAATATAAAAGTTATCCATACGGATATTC
TGAAATTTAACTTCCCCAAAAACAAAGACTA
TAAAATATTTGGTAATATCCCCTTTAACATC
AGTACTGATATTGTCAAAAAAATTGCTTTTG
AAAGCAATTCGAAATATAGTTAAATATAAAA
TATTCTCAAAACTTTTTAACGAGTGAAAAAG
TACTCAACCAAATAATAAAACAATTGAATTT
AAAAGAAACCGATACCGTTTACGAAATTGGA
ACAGGTAAAGGGCATTTAACGACGAAACTGG
CTAAAATAAGTAAACAGGTAACGTCTATTGA
ATTAGACAGTCATCTATTCAACTTATCGTCA
GAAAAAATGAAAAAAAATATAAAATATTCTC
AAAACTTTTTAACGAATGAAAAGGTACTCAA
CCAAATAATAAAACAATTGAATTTAAAAGAA
ACCGATACCGTTTACGAAATTGGAACAGGTA
AAGGGCATTTAACGACGAAACTGGCTAAAAT
AAGTAAACAGGTAACGTCTATTGAATTAGAC
AGTCATCTATTCAACTTATCTTCAGAAAAA
>Contig_12
AGTGTGTGCGAGAGAGCGAGCGTGCGTGCGA
GAGAGTTTCGCGCGCGCTTTAGAGAGCGTGC
GAGCGAGCGAGCGTGTTTGTGCCCCAGAGAA
ATGGTAATATCCCCTTTAACATCAGTACTGA
TATTGTCAAAAAAATTGCTTTTGAAAGCAAT
TCGAAATATAGTTAAATATAAAATATTCTCA
AAACTTTTTAACGAGTGAAAAAGTACTCAAC
CAAATAATAAAACAATTGAATTTAAAAGAAA
CCGATACCGTTTACGAAATTGGAACAGGTAA
```

Database

```
>Gene_1
GATAATGTAATAGAAATTGGATCAGGAAAAG
GTCATTTTACCAAAGAACTTGTCAAAATGAG
TCGGTGGGTGGATTCTATAGAAATTGATGAG
GACTTGTGTCAAGTCACCCAAAAAGTCGTGA
AGCCTTTTCAGAATATAAAAGTTATCCATAC
GGATATTCTGAAATTTAACTTCCCCAAAAAC
AAAGACTATAAAATATTTGGTAATATCCCCT
TTAACATCAGTACTGATATTGTCAAAAAAAT
TGCTTTTGAAAGCAATTCGAAATATAGTTA
>Gene_2a
ATGAACAAAAATATAAAATATTCTCAAAACT
TTTTAACGAGTGAAAAAGTACTCAACCAAAT
AATAAAACAATTGAATTTAAAAGAAACCGAT
ACCGTTTACGAAATTGGAACAGGTAAAGGGC
ATTTAACGACGAAACTGGCTAAAATAAGTAA
ACAGGTAACGTCTATTGAATTAGACAGTCAT
CTATTCAACTTATCGTCAGAAAAA
Gene_2b
ATGAAAAAAAATATAAAATATTCTCAAAACT
TTTTAACGAATGAAAAGGTACTCAACCAAAT
AATAAAACAATTGAATTTAAAAGAAACCGAT
ACCGTTTACGAAATTGGAACAGGTAAAGGGC
ATTTAACGACGAAACTGGCTAAAATAAGTAA
ACAGGTAACGTCTATTGAATTAGACAGTCAT
CTATTCAACTTATCTTCAGAAAAA
```

BLAST-based CGE finder tools

- SpeciesFinder (not KmerFinder)
- MLST
- ResFinder
- PlasmidFinder
- pMLST
- SerotypeFinder
- VirulenceFinder

Generating your own database

- Collect **DNA** sequences of the genes you wish to search for
- Save the DNA sequences in FASTA format

```
>Seq1
ACTCGCGATCCGCATAGCGCATCGCATG
>Seq2 optional comment
ATGAAAACAATGATTTATCCTCACCAATATAATTATATCAGATCGGTTATT
TATGCGGCAATGATTTATCCTCACCAATGATGAGAGAGCAGATACTCTTTG
AACAAAGAAATTGAAGCAATACTTAATAAATTT
```

- Take care to provide each gene with a unique identifier (everything from the > to the first space)
- Save the FASTA file as pure text

Congratulations - you have made your own database!

MyDBFinder

Use your own database with any gene(s) of interest