

Genomic insights and virulence in soil-persistent *E. coli*

Nicholas Waters

April 9, 2018

Department of Microbiology
School of Natural Sciences
National University of Ireland, Galway

A Brief History of Soil-persistent *E. coli*



-
- | | |
|------|---|
| 1886 | Escherich: Discovery of <i>E. coli</i> |
| 1948 | Bardsley: Soil may act as reservoir for <i>E. coli</i> |
| 1963 | W. and J. Boyd: Cold persistence observed |
| 1988 | Fujioka and Shizumura: Alternative indicators suggested |
| 1995 | R. Sjogren: soil persistence across time and depth |
| 2003 | Byappanahalli, et al: Soil persistence is widespread |
| 2010 | Brennan, et al: Persistence in maritime temperate soils |
-

Research Questions



- What types of *E. coli* are able to persist in soil?

Research Questions



- o What types of *E. coli* are able to persist in soil?
- o What virulence factors are harboured by these strains?

Research Questions



- o What types of *E. coli* are able to persist in soil?
- o What virulence factors are harboured by these strains?
- o What can we infer about adaptation from these?

Research Questions

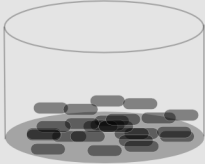


- What types of *E. coli* are able to persist in soil?
- What virulence factors are harboured by these strains?
- What can we infer about adaptation from these?
- Can we differentiate soil-persistent *E. coli* from recent contamination?

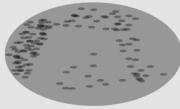
Workflow



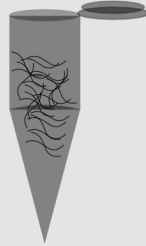
Filter



Culture



Extract DNA



ATCGATGCATAGCGAAGCTAGAGCTAGACG
GACTAGGAGCTAGAGACGTCTTGTGCTG
AGCTAGAGCTAGACGCCCCGTGTCTGTG
AAAACTCGCCTGATCACGTGCTCTCGTG
AAGCTATTTTCGAGCAACGTCTGTCAGTACA

Analyze



Sequence

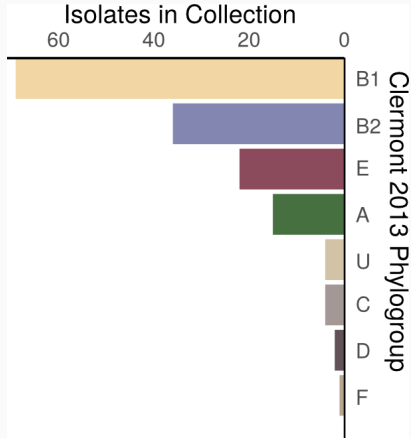
The data



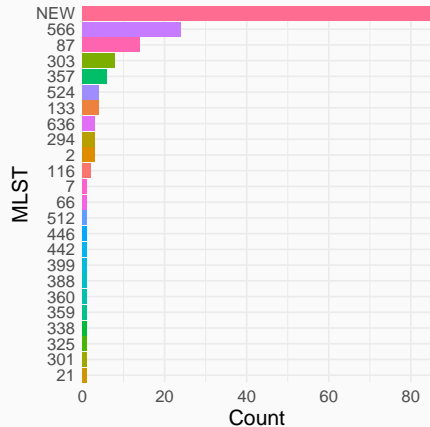
- 171 isolates sequenced
- 22 failed assembly QC or ANI threshold
- 149 true *E. coli* passed ANI screening

Sequence Typing

Clermont 2013



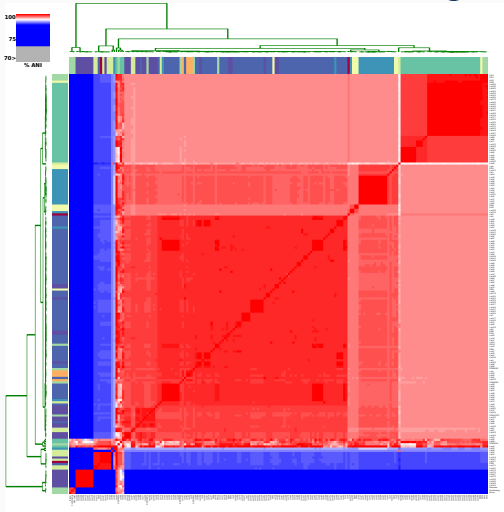
Achman 7 gene MLST



Average Nucleotide Identity



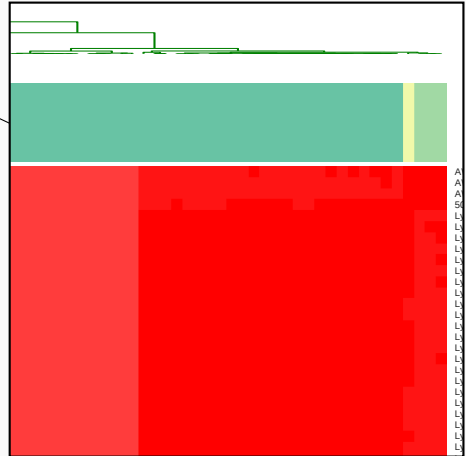
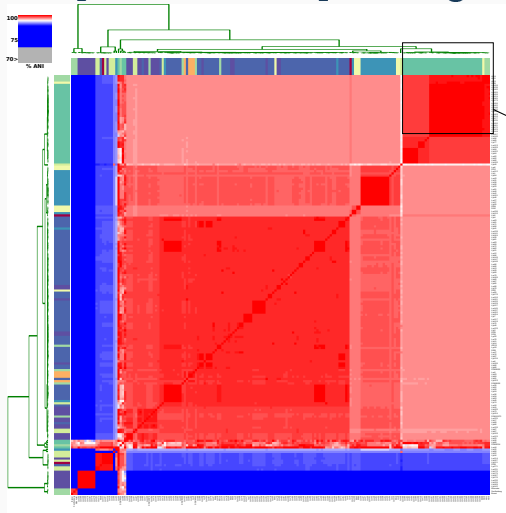
<http://widdowquinn.github.io/pyani/>



Average Nucleotide Identity



<http://widdowquinn.github.io/pyani/>

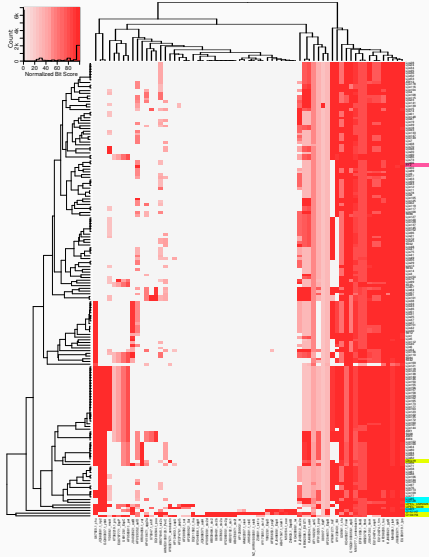


Virulence

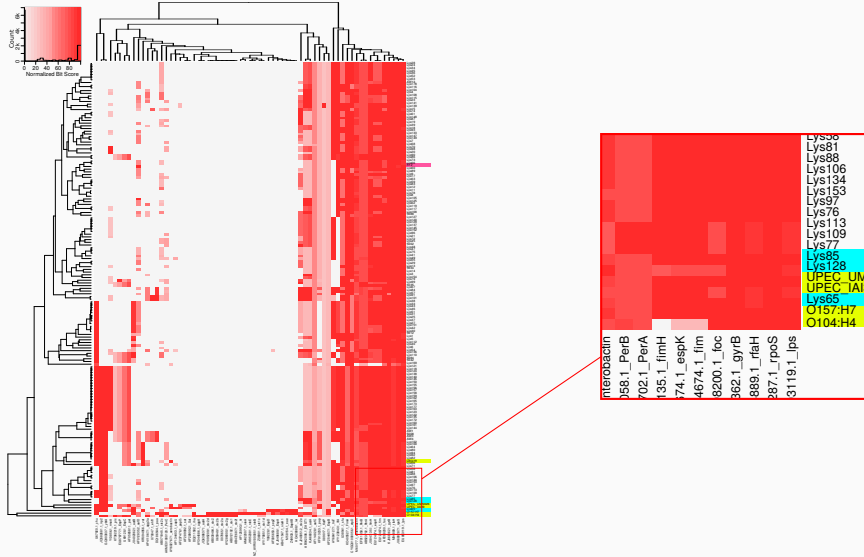


- o Search literature for genes implicated in virulence
- o Select representative sequences for 65 virulence factors
- o Use reciprocal translated blast to find occurrences
- o Filter results, visualize

Virulence Results



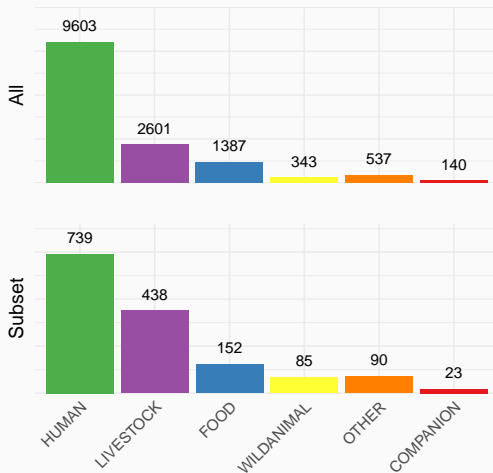
Virulence Results



Enterobase comparison strains



- One isolate from each Achman 7 MLST
- Total: 1185

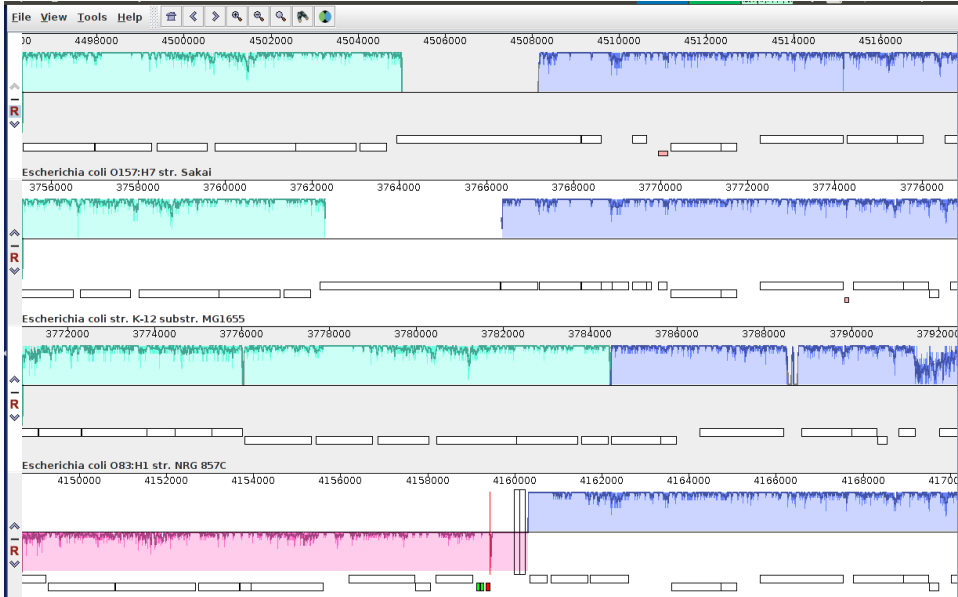


Pangenome Analysis



	N	Core >99%	Soft Core 95-99	Shell 15-95	Cloud 0-15	total
Soil	149	2662	505	2664	15,831	21,662
Enterobase	1193	1822	514	3346	73,606	79,288
All	1342	1806	535	3406	78,121	83,868

Assessing Assembly



1. Create reference pangenome
2. Find genes next to contig borders
3. Blast against pangenome
4. Reject hits $< 90\%$ of CDS length

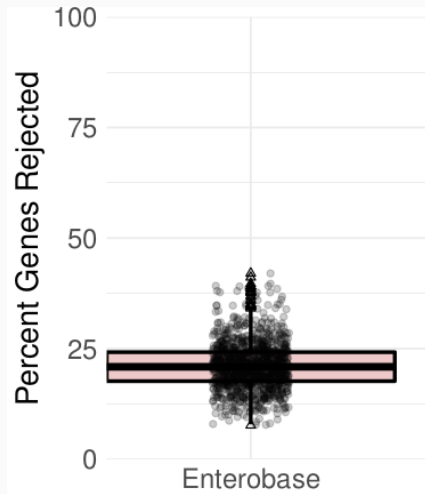
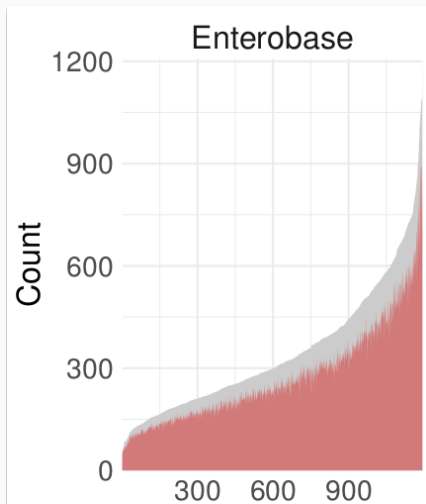
1. Create reference pangenome
2. Find genes next to contig borders
3. Blast against pangenome
4. Reject hits $< 90\%$ of CDS length



annofilt performance



<https://nickp60.github.io/annofilt/>



Interested in Genome Assembly?



Come ask about our tool **riboSeed** to assemble through rDNA repeats!

- o Poster 466 zone D (upper gallery)
- o Tuesday and Wednesday Evening

LEV CT URE
AGE BA IAL ARCHITE
ER
riboSeed
LEVER AGE BACT ER IAL ARCHITE CTURE

Detecting differential presence/absence



Scary
microbial pan-GWAS

Future work



- subpangenomes
- virulence pathways
- AMR
- mobile elements

Conclusions



- Soil *E. coli* represent diverse lineages
- Soil *E. coli* possess a range of virulence genes
- Pangenome analysis is very sensitive to annotation

Acknowledgments



OÉ Gaillimh
NUI Galway

NUIG Microbiology

- ☐ Dr. Fiona Brennan
- ☐ Dr. Florence Abram
- ☐ Soil and Environmental
Microbiology Research Group
- ☐ Functional Environmental
Microbiology Group



James Hutton Institute, Dundee

- ☐ Dr. Leighton Pritchard
- ☐ Dr. Ashleigh Holmes

Questions?