Data Notes

Only have Shigella sonnei atm in dropbox.

Also have data core aln and tree for Shigella flexneri as well

Shigella sonnei ref - NC\_007384 – standard for bug

Masked phage / repeats and then recombination with Gubbins

2 aln

one from gubbins – includes masked sites with -. The final.tre is from gubbins.

Second – used SNPsites to extract ATCGs only from gubbins aln. Treefile.tree from iqtree with constant sites provided.

Plasmid - LN624486 – associated with multidrug (MDR) and extremely drug resistant (XDR). There are different forms of the plasmid – can show some papers on this if helpful.

Have done plasmid aln for Sonnei, Flexneri, Flexneri + Sonnei combinations at >50, >80 and >90 threshold.

2 aln for moment

All sonnei isolates with >90% coverage of the plasmid. SNP aln only. Not masked for recomb.

All sonnei isolates with >50% coverage of the plasmid. SNP aln only. Not masked for recomb.

Fig Notes

Phandango – shows recombination blocks – limited in Shigella sonnei

AMR profile in genotype by year (samples span five years). Genotype scheme from Hawkey et al 2021 Nat Comms – useful nomenclature

Plasmid profile by genotype and year

Plasmid – at >50 and >90 coverage in Sonnei population. AMR and Genotype profile.

Plasmid presence by two species

**Update 22/08/2022**

Genotype + year + month txt file. Genotype definied by specific SNPs in core – used for surveillance – corresponds to tree topology. Some are linked to epi factors like MSM and others geo – like Cipro resistance in South Asia.

Quick plots of Sonnei + genotype with chr and 4 plasmids. Vir plasmid looks to roughly match topology of chr – this is expected – found a paper from 2013 that looked at this as well. Should co-evolve.

Small spA plasmid. There isn’t much in the literature about this plasmid. It is in the sonnei population a lot – but looks to be both gained/lost – especially in the broad lineage 3 (genotype 3.XXXX). lineage 3 associated with global expansion.

Aln – at 70% for the large MDR plasmid. Flex, sonnei and FlexSon (both together)

Have AMR plots sort of worked up. Need to refine some more – some genes are associated with specific mobile elements that may be chr or plasmid mediated – to make more accessible to non bacterial AMR ppl.

Broadly – there are some genes eg that should be on the spA plasmid (sul2, strAB) on a transposon. Chance these could be integrated into chromosome. Think some heatmaps of chr + plasmid + specific AMR genes will address this question\

23/08/2022

Aln of Sonnei and Flex to same ref (chr and vir plasmid).

In the Aln\_Trees/FlexSon\_Ss046

* Three groups in the chr tree – sonnei (big one) and the flex – splits out into two well known groups. Smaller one is called flex 6. No recomb detection undertaken
* Plasmid – dropped threshold ot 50% plasmid reference covered by reads. Small group on long branch – flex 6. This has been previously characterised as having a different isoform pINV A of the Vir plasmid (based off 3 virulence genes).
* Soneni and flex in the bigger group cluster by spp (quick look). Both have been previously characterised as having pINV A (this is shigella specific nomenclature).

26/08/2022

Think have nearly everything uploaded aln wise.

Just need the flex only – running recomb detection atm

Will update the excel spreadsheet that summarises the different aln.

Have also dot pointed methods of what I did in prep for writing up properly