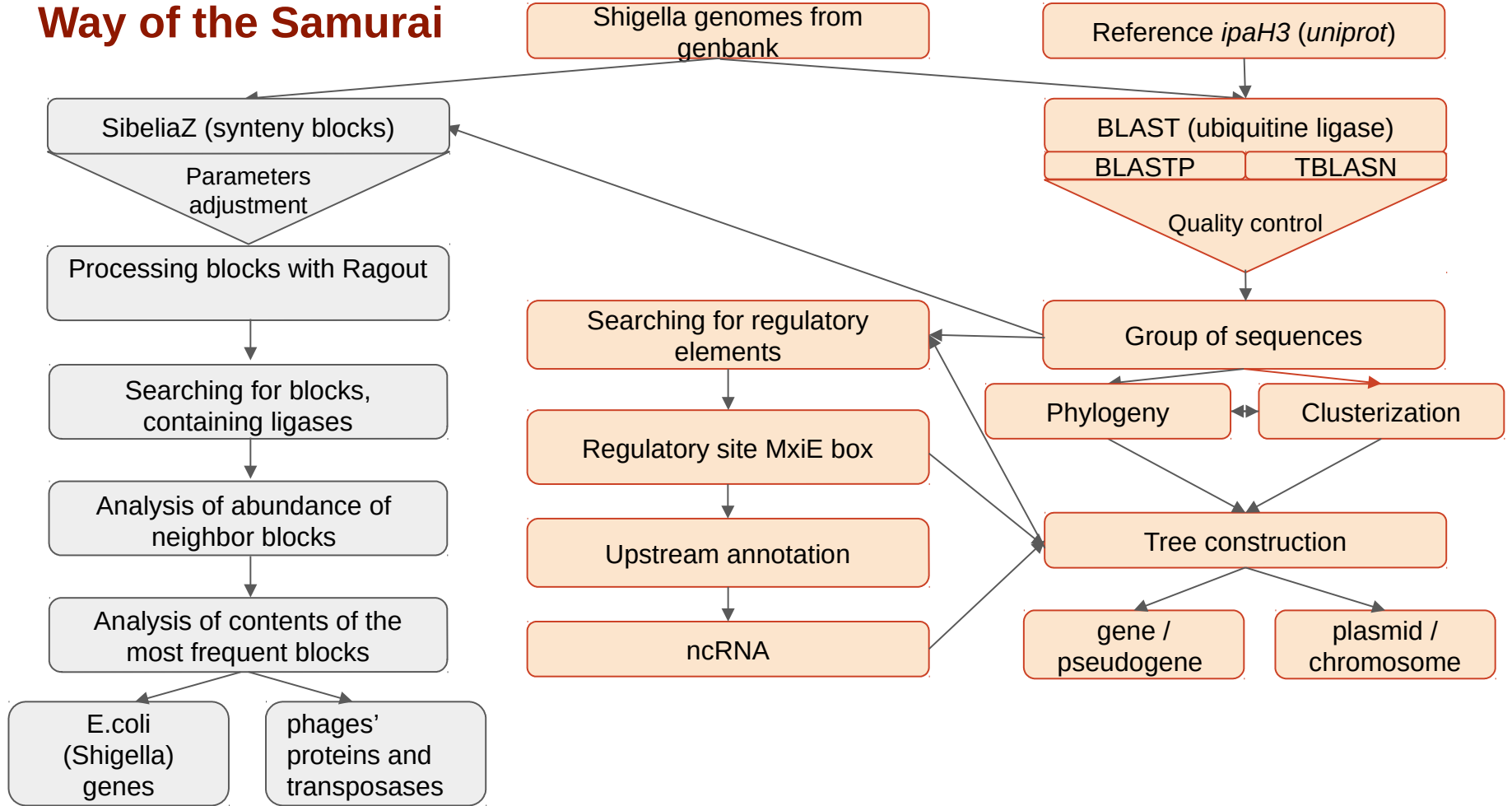


Way of the Samurai



ipaH, why do you need 400 conservative letters?

Conserved fragment: 1...1184
organism = '*Shigella boydii* CDC3083-94'

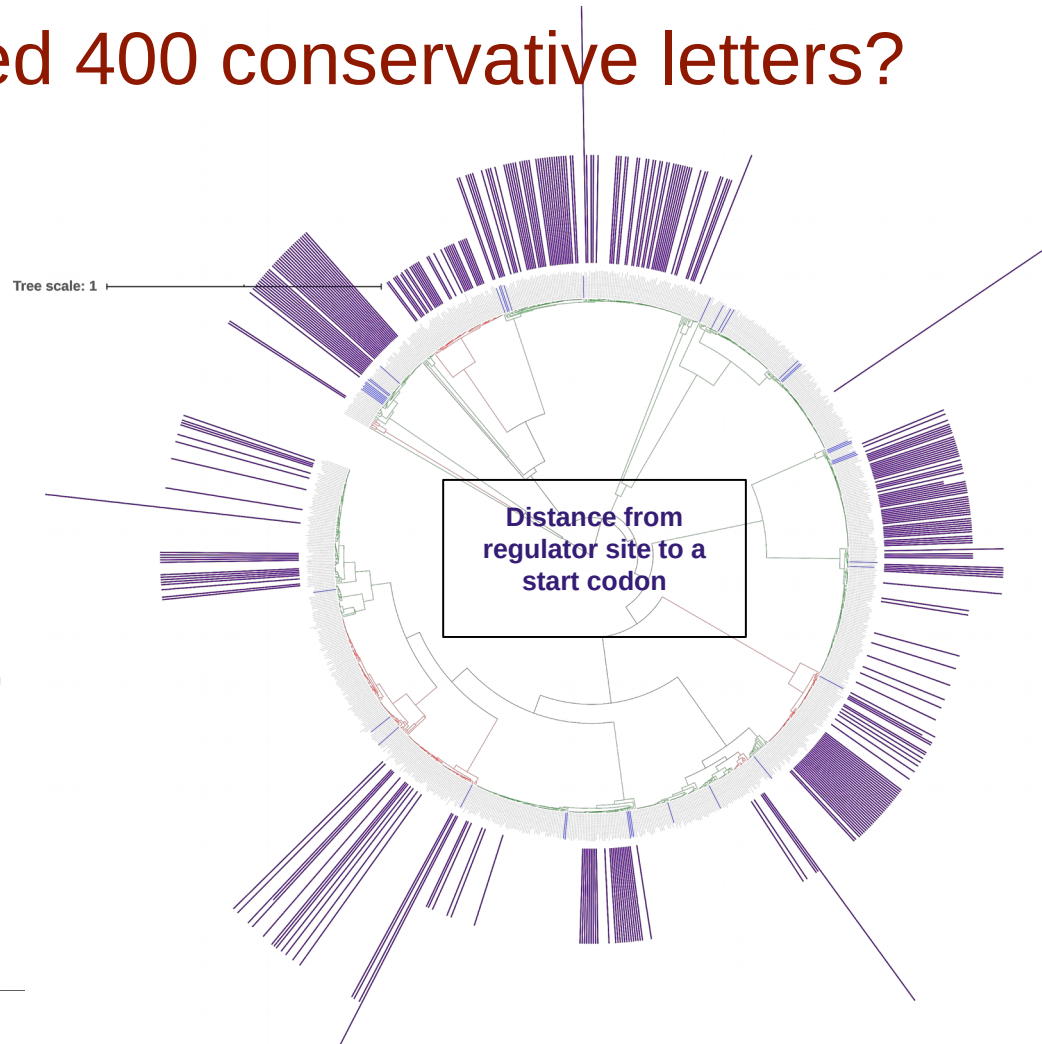
gene: 406 ... >1184
/product = 'ubiquitin ligase'

Sites of transcription factors??

One more protein?!

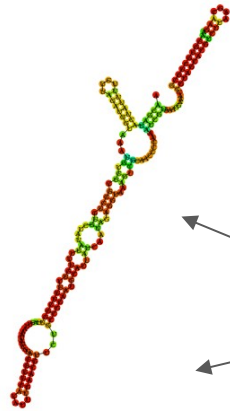
location of MxiE box site in reference *Shigella* strain

	MxiE box	TATAAT	ATG
VP genes			
<i>ospB</i>	ttttt GTC CGTTTTTTTAAAT	catagtatacaagaagc..	TtgcAT actatat - 57 ATG
<i>ospC1</i>	agaaa GTATCG TTTTTTTACAG	taaaattccattgctgttcaa	TaaAAT t gatata - 229 ATC
<i>ospE1</i>	aaaaa GTATCG TTTTTTTACAG	taaacttcatttagcogac.	TATAAT gtaaaaa - 38 ATG
<i>ospE2</i>	aaaaa GTATCG TTTTTTTACAG	taaacttcatttagcogac.	TATAAT gtaaaaa - 36 ATG
<i>ospF</i>	ataat GTATCG TTTTTATaAaAG	atgataaacaatcaatataa	gATAAT atateta - 20 ATG
<i>virA</i>	gaaat GTATCG TTTTTcTTAaAG	agaagaataaacattccatllt	TATLAT gtlccca - 56 ATG
<i>ipaH4</i>	ctcta GgAT cGTTTTTTTAAAG	actttctcgttttatattgc.	atTAAT agaccaa - 25 ATG
<i>ipaH9</i>	aaaca GTATCG TTTTTTTACAG	ccaattttgttttatccctat	TATAAT aaaaaa - 96 ATG
<i>ipaH7</i>	aatgt GTATCG TTTTTTTACAG	taatttttaattgtttatlc	TATAAT aggaata - 271 ATG
Chromosomal genes			
<i>ipaHa</i>	ttaca Ggcg CGTTTTTTTAAAG	aatcttcactccatttgcac	gAatAT attaata - 25 ATG
<i>ipaHb</i>	agaaT GTAc cGTTTTTTTAAAG	aaaaaacagttacattgaga	TATgAT tttgatc - 357 ATG
<i>ipaHc</i>	aaaca GTATCG TTTTTTTACAG	ccaattttgttttgcoctt.	TATAAT aaaaaaa - 88 ATG
<i>ipaHd</i>	aaaaa GTATCG TTTTTTTACAL	taaaccgatccagtttagga	TAgTgT aaagcag - 163 ATG
<i>ipaHe</i>	aaata GTATCG TTTTTTTACAG	ttaaatcaacatcacttct	TAAAT gaaaaa - 85 ATG
MxiE box	GTATCG TTTTTTTAAAG	----- 20 n -----	TATAAT

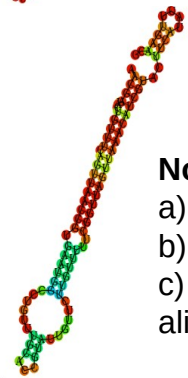


ncRNAs were found in the ipaH upstreams for three different clades of the tree

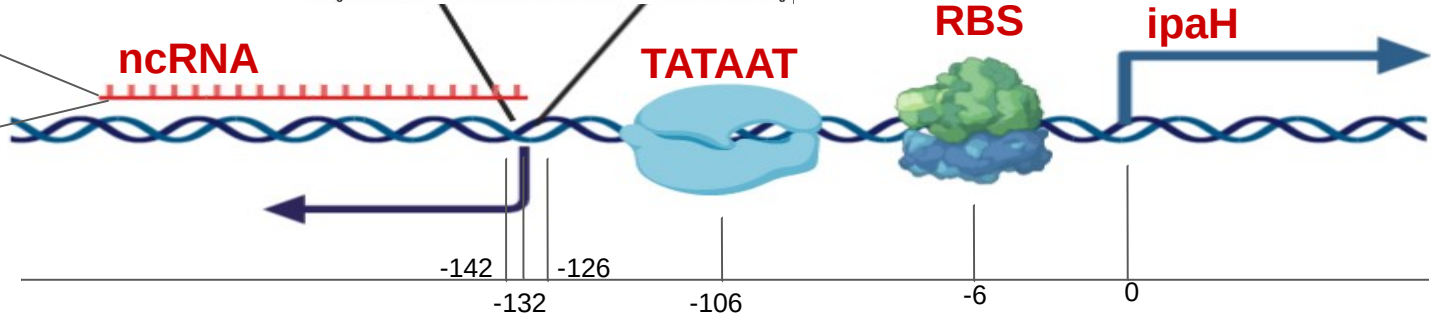
a)



b)



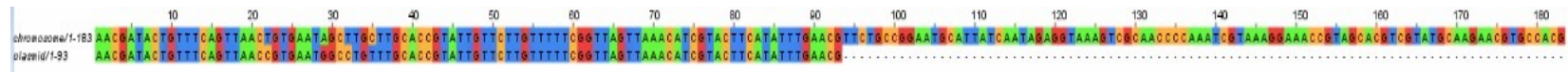
Non-coding RNA
 a) on chromosome
 b) on plasmid
 c) sequecne alignment



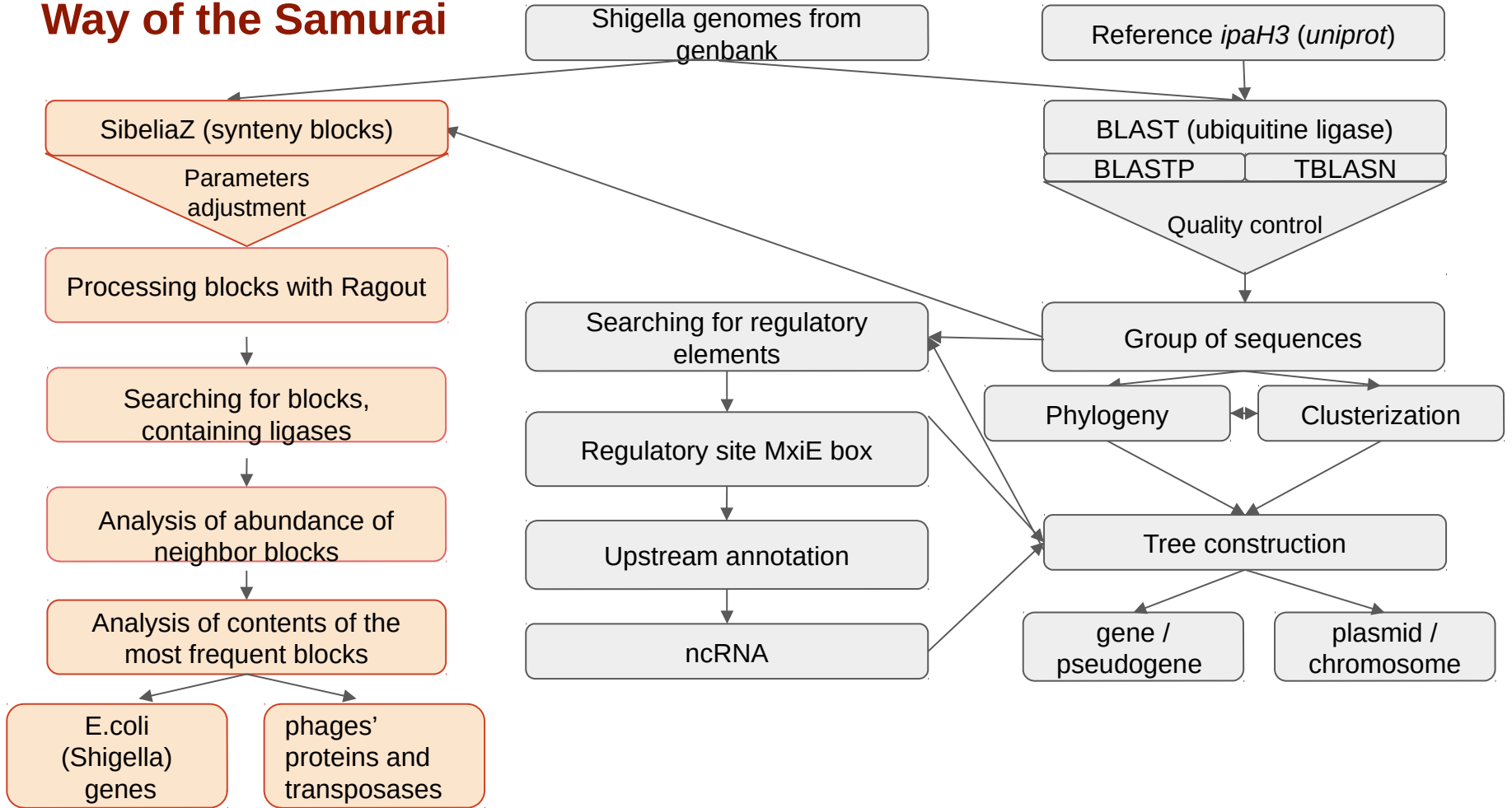
Conclusions

- Transcription factor MxiE's sites are presented in both chromosomes and plasmids
- ncRNAs are presented in ligase upstreams of both chromosomes and plasmids
- Genome may have multiple ubiquitine ligases, which upstreams have hcRNA

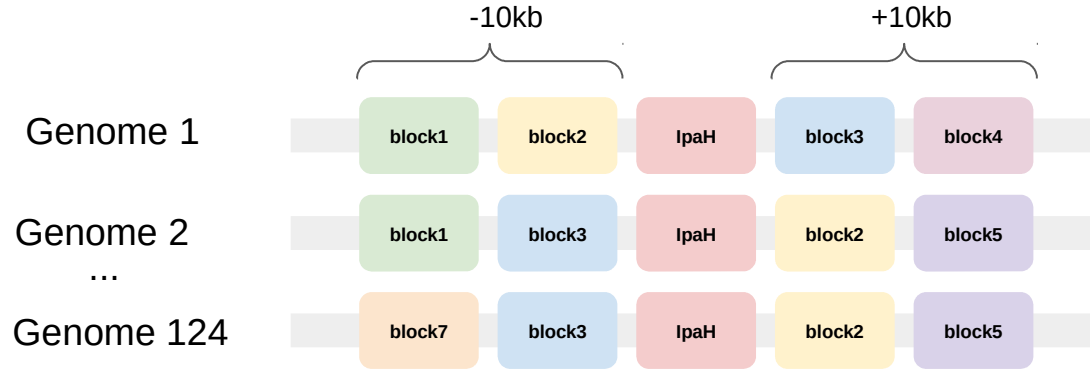
c)



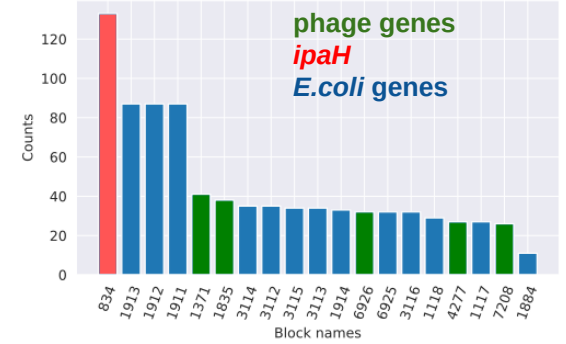
Way of the Samurai



Chromosomal ligases have specific locus of genomic island integration

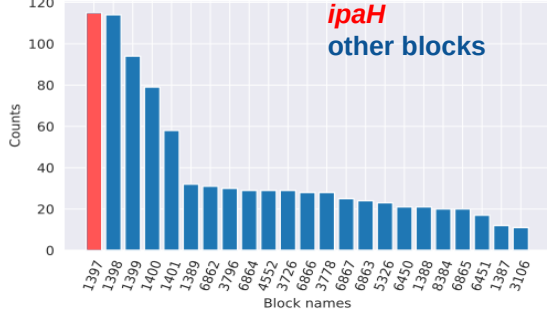


Distribution of neighbour blocks across the *Shigella* genomes

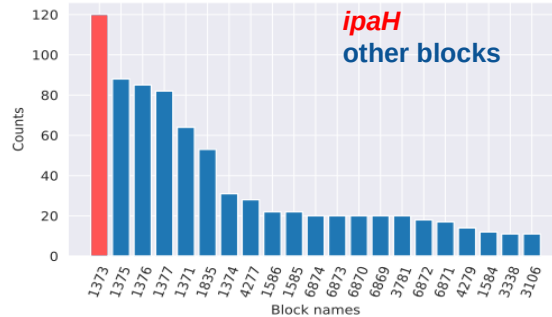


Count frequency: 2 2 3 2 1 1 2

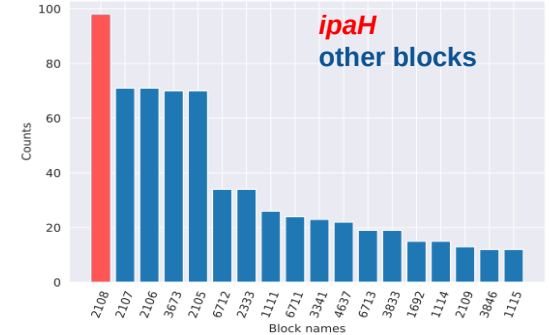
Distribution of neighbour blocks across the *Shigella* genomes



Distribution of neighbour blocks across the *Shigella* genomes

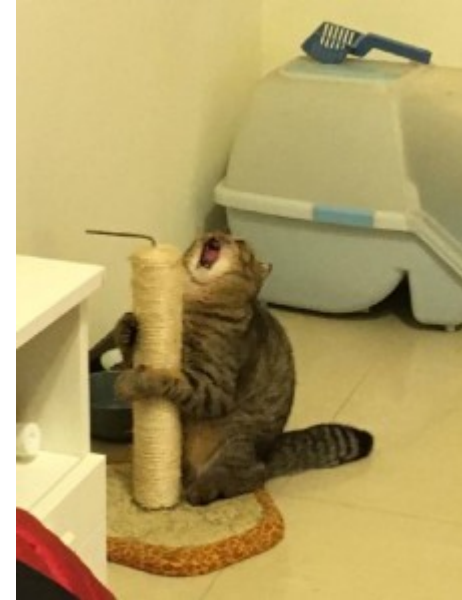


Distribution of neighbour blocks across the *Shigella* genomes



Conclusions

- I. Chromosomal and plasmid ligases are different; we have not identified significant cases of transfer.
- II. Ligases of all types, including chromosomal and plasmid, are regulated by MxiE
- III. Only in one cluster (# 3) we did not identify a single pseudogene, perhaps this indicates the importance of this type of ligases (*maybe we could not find them?*)
- IV. In the upstreams of two types of chromosomal ligases, we found ncRNA overlapping the site of the transcription factor. In the upstream of one of the plasmid types, we also found ncRNA, but shorter (*perhaps ncRNA regulates ipaH?*)
- V. We found that ligases surrounded by many phage genes, but their strong diversity does not allow us to unambiguously reconstruct the genomic island.
- VI. The same bacterial genes of *E. coli* was found with ligases of the same type, that indicates the specificity of the genomic island insertion loci.



to be continued...