Plasmid 114 (2021) 102562

Contents lists available at ScienceDirect

Plasmid

journal homepage: www.elsevier.com/locate/yplas

Insights into the individual evolutionary origins of *Yersinia* virulence factor effector proteins

Veronica R. Moorman [1,*], James I. Cohen [1,2]

Kettering University, 1700 University Ave, Flint, MI 48504, USA

ARTICLE INFO

Keywords:
Virulence factors
Yop genes
Yersinia
Phylogenetics
pYV plasmid
Mosaic plasmid

ABSTRACT

Pathogenic *Yersinia* bacteria, including *Y. pseudotubuclosis Y. enterocolitica*, and *Y. pestis*, contain the mosaic plasmid pYV that encodes for, among other things, a number of proteinaceous virulence factors. While the evolutionary histories of many of the biovars and strains of pathogenic *Yersinia* species are well documented, the origins of many of the individual virulence factors have not been comprehensively examined. Here, the evolutionary origins of the genes coding for a set of *Yersinia* outer protein (Yop) virulence factors were investigated through phylogenetic reconstruction and subsequence analysis. It was found that many of these genes had only a few sequenced homologs and none of the resolved phylogenies recovered the same relationships as was resolved from chromosomal analyses. Many of the evolutionary relationships differ greatly among genes on the plasmid, and variation is also found across different domains of the same gene, which provides evidence of the mosaic nature of the plasmid as well as multiple genes on the plasmid. This mosaic aspect also relates to patterns of selection, which vary among the studied domains.

1. Introduction

Yersinia is a genus of facultatively anaerobic Gram-negative coccobacilli. Out of the more than a dozen accepted species of *Yersinia*, three are known to cause disease in humans: *Y. pseudotuberculosis* (Far East scarlet-like fever), *Y. enterocolitica* (Yersiniosis), and *Y. pestis* (bubonic, septicemic, and pneumonic plagues). While *Y. pestis* also contains other virulence plasmids (pFra and pPla) leading to specific pathogenesis, all three pathogenic *Yersinia* species contain a common virulence plasmid of approximately 70 kilobases (known as pYV, pCD1, pLCR, or pIB1) that encodes several genes including those that code some bacterial virulence factors known as Yops (*Yersinia* outer membrane proteins), as illustrated in Fig. 1A–B. The original acquisition of these virulence plasmids is thought to have likely formed through either a common ancestor progenitor for all three of the pathogenic *Yersinia* species (Howard et al., 2009) or through independent acquisition by *Y. enterocolitica* and *Y. pseudotuberculosis*/*Y. pestis* (Reuter et al., 2014) and likely shares a common ancestor with virulence plasmids in *Shigella* and enteroinvasive *E. coli* (EIEC) (Pilla and Tang, 2018).

Many of these Yop proteins are secreted into human host cells through a type three secretion system (ttss or t3ss) using the YopB/D translocon. These proteins are known as Yop effectors, as they affect the host cells directly. A subset of these effectors are known to be required for an infection, including YopH (a protein phosphatase which disrupts host signaling), YopE (a Rho GTPase), and YopM (recruits other proteins and translocates into the host nucleus) (Rimpilainen et al., 1992). Others, like YopT (cysteine protease), YopO (a serine/threonine protein kinase also known as YpkA), and YopJ or YopP (an acetyltransferase that increases host apoptosis) are not necessarily required, yet aid in the infection process. Expression of the *yop* genes which code for these Yop proteins is correlated with the capacity of the *Yersinia* bacteria to avoid host defense mechanisms (Cornelis et al., 1998).

Additionally, some of the secreted Yop effector proteins require a corresponding Syc chaperone protein to help them translocate into the host cell, as illustrated in Fig. 1C. These include YopE, YopH, and YopT which require SycE (Wattiau and Cornelis, 1993), SycH (Woestyn et al., 1996), and SycT (Iriarte and Cornelis, 1998), respectively. Other Yop effector proteins, such as YopO, YopP/YopJ, and YopM, do not require a chaperone and can travel alone through the YopB/YopD translocator (Trulzsch et al., 2003). In the present study, the evolutionary history of

* Corresponding author.
E-mail address: vmoorman@kettering.edu (V.R. Moorman).
[1] These authors contributed equally.
[2] Present address: Weber State University, 1415 Edvalson St Ogden UT 84408.

https://doi.org/10.1016/j.plasmid.2021.102562
Received 23 June 2020; Received in revised form 31 December 2020; Accepted 14 January 2021
Available online 23 January 2021
0147-619X/

each of the genes that code for these Yops and Sycs were compared to each other as well as to the 16S rRNA gene, which is encoded as part of the genome, and the phylogeny resolved from Pathosystems Resource Integration Center (PATRIC) (Davis et al., 2020).

Interestingly, much of what is known regarding the evolution of *Yersinia* virulence is confined to tracing the origins of the virulence plasmids within the *Yersinia* genus itself (Tan et al., 2016; Duan et al., 2014; Achtman et al., 1999; Demeure et al., 2019; Bochkareva et al., 2018). The plasmid pYV, however, is a mosaic plasmid, and thus individual genetic elements have distinct sources (Pesesky et al., 2019). Consequently, the evolutionary origins of many of these individual virulence factors remain unexplored and the similarities and differences among the evolutionary histories of these virulence factors have yet to be investigated in detail. Here, we report our investigation into the evolutionary origins of *Yersinia* virulence as examined through phylogenetic analyses of these virulence genes on pYV.

2. Materials and methods

2.1. Sequences and matrix construction

The identification of homologous sequences was initiated using the KIM5 strain of *Yersinia pestis*. Specifically, sequences of genes located on the pYV plasmid were obtained from AF053946.1 (Hu et al., 1998) (*yopE*, *yopH*, *yopT*, *sycT*, *sycE*, *sycH*, *yopJ*, *yopO*, and *yopM*); each gene was annotated except for *sycT* and *yopT*, which were found by comparing the full plasmid sequence to that of *Y. pestis* strain CO92 where they were annotated (NC_003131.1). Alternatively, the 16S rRNA gene was obtained from NZ_CP009836.1. These sequences, as well as their individual protein domain sequences (when applicable), were BLASTed using The National Center for Biotechnology Information (NCBI)'s tblastn function which involves a Basic Local Alignment Search Tool (BLAST) algorithm to search across the nr/nt nucleotide databases using a protein query. Default settings were employed, except for increasing the maximum number of sequences to 20,000. Because most of the matrices included multiple individuals for each species, a separate matrix that only included one arbitrarily selected representative from each species was constructed from the full matrix. These matrices are referred to as full and strict respectively.

2.2. Phylogenetic analyses

The 23 DNA regions, inclusive of both full genes and domain sections, of the full and strict sets of species were aligned using TranslatorX (Abascal et al., 2010), with MAFFT (Katoh and Standley, 2013) as the alignment program and guessing the most likely reading frame to ensure a codon-based alignment for the gene regions. For 16S, a non-coding region, MAFFT alone was used for the alignment. Phylogenetic analyses were conducted for all alignments using RAxML (Stamatakis, 2014) on the Kettering University High-Performance Computing cluster (KUHPC) or the CIPRES web server (www.phylo.org). For these analyses, a GTR + G + I model (CAT for 16S due to the large number of species) was employed to resolve the best scoring Maximum Likelihood (ML) tree and to conduct 1000 or 10,000 rapid bootstrap analyses (100 for 16S due to its large size), with the number of bootstrap analyses depending on the size of the dataset. The resulting trees were examined and compared, and the phylogeny of *Yersinia* and relatives from the Pathosystems Resource Integration Center (PATRIC) (Davis et al., 2020) was also utilized for a comparison given a different, non-plasmid set of a genes employed for the reconstruction of the phylogeny, which includes multiple species overlapping with those in the present study. Phylogenetic analyses were not conducted for the strict sets of the YopH linker, YopE N-terminal, and YopM C-terminal as these datasets included only one, two, and three species, respectively. Alignments and results from RAxML are available at Dryad (doi:https://doi.org/10.5061/dryad.905qftthh).

2.3. Patterns of selection and ancestral gene reconstruction

In addition to reconstructing phylogenetic relationships, patterns of selection for the genes involved in virulence in *Yersinia* – *sycE*, *sycT*, *sycH*, *yopE*, *yopT*, *yopH*, *yopM*, *yopJ*, and *yopO* – were examined with PAML (Yang, 2007; Yang, 1997) and Selecton (Stern et al., 2007; Doron-Faigenboim et al., 2005) using the strict datasets for these genes. Additionally, the various domains of *yopH*, *yopM*, and *yopO* were independently examined with PAML. Some of the strict datasets had a large number of species (hundreds) making these types of analyses computationally challenging without providing additional information concerning evolutionary patterns among *Yersinia* and relatives. Therefore, for these datasets, the number of species was reduced to 100, with only *Yersinia* and close relatives included (i.e., 100-species datasets). In PAML, codeml was employed with the following parameters: a codon frequency based on the average nucleotide frequencies at each of the codon positions was used (CodonFreq = 2), a free-ratio model was employed (model = 1), and kappa (initially 2) and omega (initially 0.4) were estimated, alpha was fixed at 0, and a clock model was not used. Ka/Ks values were identified among branches of *Yersinia* and relatives. These analyses were undertaken on the KUHPC.

For Selecton, the codon alignment, ML tree, and protein structure (based on PDB ID) were used, and analyses were run for each structure in

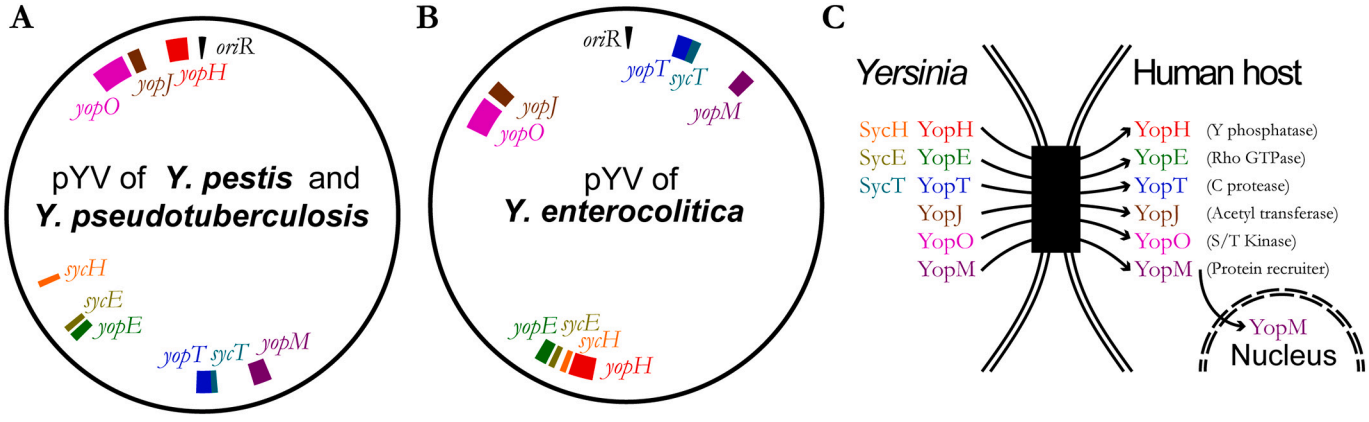

Fig. 1. Locations of genes and their protein products investigated in this study. A) Plasmid map showing genes of interest in this study of *Y. pestis* and *Y. pseudotuberculosis*. B) Plasmid map showing genes of interest in this study of *Y. enterocolitica*. Both plasmid maps were created using Plasmidomics v 0.2 (Winkler, 2020) and data from references (Hu et al., 1998; Snellings et al., 2001). C) Graphical depiction of the locations and roles of the proteins of study in *Yersinia* infection. All panels use the same colour scheme to.

the Protein Data Bank. For each gene, five models (M8 beta + w ≥1, M8a beta + w = 1, M7 beta, M5 gamma, and Mechanistic Empirical Combination [MEC]) were examined, with 8 categories for the distribution. Genes with codons under strong selection were identified and compared. These analyses were conducted on the Selecton server (http://selecton.tau.ac.il).

3. Results

3.1. Phylogenetic analyses

A total of ten different *Yersinia pestis* KIM5 genes were analyzed in the present study. In addition to investigating evolutionary patterns of full-length sequences for the six genes coding for multidomain proteins (YopH, YopE, YopJ, YopM, YopO, and YopT), individual domains were analyzed separately as was the linker region in YopH. In total, there were 23 genes or gene fragments analyzed, ranging from 138 to 8130 base pairs in aligned length and with 66 to 22,537 sequences per region. The statistics for each gene or gene fragment are included in Table 1, and phylogenies of some genes are included in Figs. 2 and 3.

For most of the genes analyzed, only one origin was identified in the genus *Yersinia* (Table 1). YopE (92% bootstrap support [BS]), YopO (100% BS), SycH (100% BS, Fig. 2C), and SycT (100% BS) resolve similar relationships, with *Yersinia* being sister to *Edwardsiella* and *Aeromonas, Vibrio* and *Aeromonas*, *Edwardsiella* and *Aeromonas*, or *Photorhabdus* and *Vibrio*, respectively, with these being the only genera present for each of these genes. For the aforementioned genes, only the three pathogenic species of *Yersinia* appear to have the gene, (*Y. enterocolitica*, *Y. pestis*, and *Y. pseudotuberculosis*), and while each species is not monophyletic, the entire genus is in each case. Specifically, it was resolved that the phylogenetic relationships in the C-terminal domain of YopE are the same as those with the entire gene, but only *Yersinia* species appear to have sequences homologous to YopE's N-terminus and, therefore, are included in the resolved phylogeny. The middle and C-terminal domains of YopO have similar relationships to those from the entire gene, with the recovered sister groups differing slightly among them. The N-terminus resolves a clade of *Vibrio* and *Oenococcus* as sister to species of *Yersinia*; the C-terminus just includes *Vibrio* as sister to *Yersinia*; and analyses of the entire *yopO* gene reconstructs *Aeromonas* in that position. It should be noted that *Oenococcus* is on a long branch. The C-terminal domain only includes species of *Yersinia*, *Vibrio*, and *Aeromonas*, while the N-terminal domain and the entire gene have many more taxa including some eukaryotes. In SycH, *Yersinia ruckeri* is sister to a clade that includes *Providencia alcalifaciens* and many individuals of *Salmonella enterica*, and a clade of *Y. enterocolitica*, *Y. rohdei*, *Y. alecksicae*, and *Arsenophonus nasoniae* is sister to this clade. The clade of *Edwardsiella*, *Aeromonas*, and the pathogenic *Yersinia* species are sister to this large clade.

In YopJ, YopT, and SycE, two origins of each gene are resolved in *Yersinia* (Fig. 2B). In YopJ, a clade consisting of *Y. enterocolitica*, *Y. pestis*, and *Y. pseudotuberculosis* is sister to one composed of species of *Edwardsiella* (93% BS), while *Y. ruckeri* is a member of a separate clade that includes species of *Pandoraea*, *Ralstonia*, *Vibrio*, and others (23% BS). A similar result is recovered with just the YopJ N-terminal domain, although many more species are included in the analysis of this region. The clade of the three infectious species of *Yersinia* is sister to *Escherichia marmotae*, and this clade is sister to one that includes *Edwardsiella* and *Pseudomonas*. *Y. ruckeri* is in another, separate region of the phylogenetic tree as is one sample of *Y. pestis*. In YopT, most species of *Yersinia* are in a clade sister to species of *Candidatus*, *Photorhabdus*, *Vibrio*, and others (49% BS), while *Y. entomophaga* is a member of a separate clade that is sister to a species of *Janthinobacterium* (97% BS). These results are similar to those from the C-terminal domain of this gene; although, the three infectious species of *Yersinia* are sister to a clade of only species of *Candidatus*. For the N-terminal domain, the three infectious species of *Yersinia* are sister to species of *Photorhabdus*, which are the only other species included in the analysis. In SycE, *Y. enterocolitica*, *Y. pestis*, and *Y. pseudotuberculosis* are all members of one clade sister to *Drosophila*,

Table 1
Statistics for the sequences analyzed in this study, including information about the full data sets and strict data sets that were used.

Gene Information			Full Data Set				Strict set
Gene protein product (domain)[a]	GenBank Accession #	NA #ing from AF053946 (Hu et al., 1998)	# of Homologous Species	Aligned length (base pairs)	Computed Tree's Likelihood Score	# of origins in *Yersinia*	# of Homologous Species
YopH	AAC62606	68,243–69,649	14,777	4089	−835,521.01	5	479
YopH (N)	AAC62606	68,243–68,632	121	393	−1876.42	2	7
YopH (C)	AAC62606	68,731–69,649	16,627	3705	−905,336.29	4	495
YopH (L)	AAC62606	68,633–68,730	70	138	−453.86	1	–
SycH	AAC62591	48,188-48,613	66	432	−2459.33	1	6
YopE	AAC62587	44,845-44,186	77	669	−3114.26	1	7
YopE (N)	AAC62587	44,845–44,591	76	255	−442.02	1	2
YopE (C)	AAC62587	44,578–44,186	92	417	−2115.93	1	8
SycE	AAC62588	45,039-45,431	161	462	−5532.40	1–2 (1 in *Y. pestis* and relatives)	18
YopJ/YopP	AAC62603	65,694-66,560	569	1251	−25,739.25	2	27
YopJ/YopP (N)	AAC62603	65,769–66,287	848	618	−19,952.69	3 (2 in *Y. pestis*)	31
YopT	AAC62582	35,828–34,860	199	1209	−11,447.70	2 (1 in *Y. pestis*)	14
YopT (N)	AAC62582	35,828–35,529	97	300	−1158.82	1	5
YopT (C)	AAC62582	35,522–34,863	326	1092	−22,355.14	2 (1 in *Y. pestis*)	31
SycT	AF053946[b]	34,860–34,465	74	417	−2015.68	1	5
YopO/YpkA	AAC62602	63,100-65,298	16,000	8130	−1,191,907.40	1	1152
YopO/YpkA (M)	AAC62602	63,541–64,314	22,537	5823	−329,018.62	1	1438
YopO/YpkA (C)	AAC62602	64,411–65,286	92	903	−4065.01	1	5
YopM	AAC62580	30,873-32,102	9285	7215	−198,457.68	7	121
YopM (N)	AAC62580	30,894–31,061	187	171	−989.15	1	7
YopM (M)	AAC62580	31,398–32,009	6608	5415	−184,436.70	18	138
YopM (C)	AAC62580	32,028–32,099	74	72	−22,355.14	1	3
16S	NZ_CP009836 (Johnson et al., 2015)	n/a	11,813	3704	−1,436,759.077	1	n/a

[a] N, C, and M stand for the N-terminal, C-terminal, or Middle domain of an identified protein while L stands for a linker region between known domains.
[b] The gene sequence was determined by comparing the full plasmid sequence to that of strain CO92.

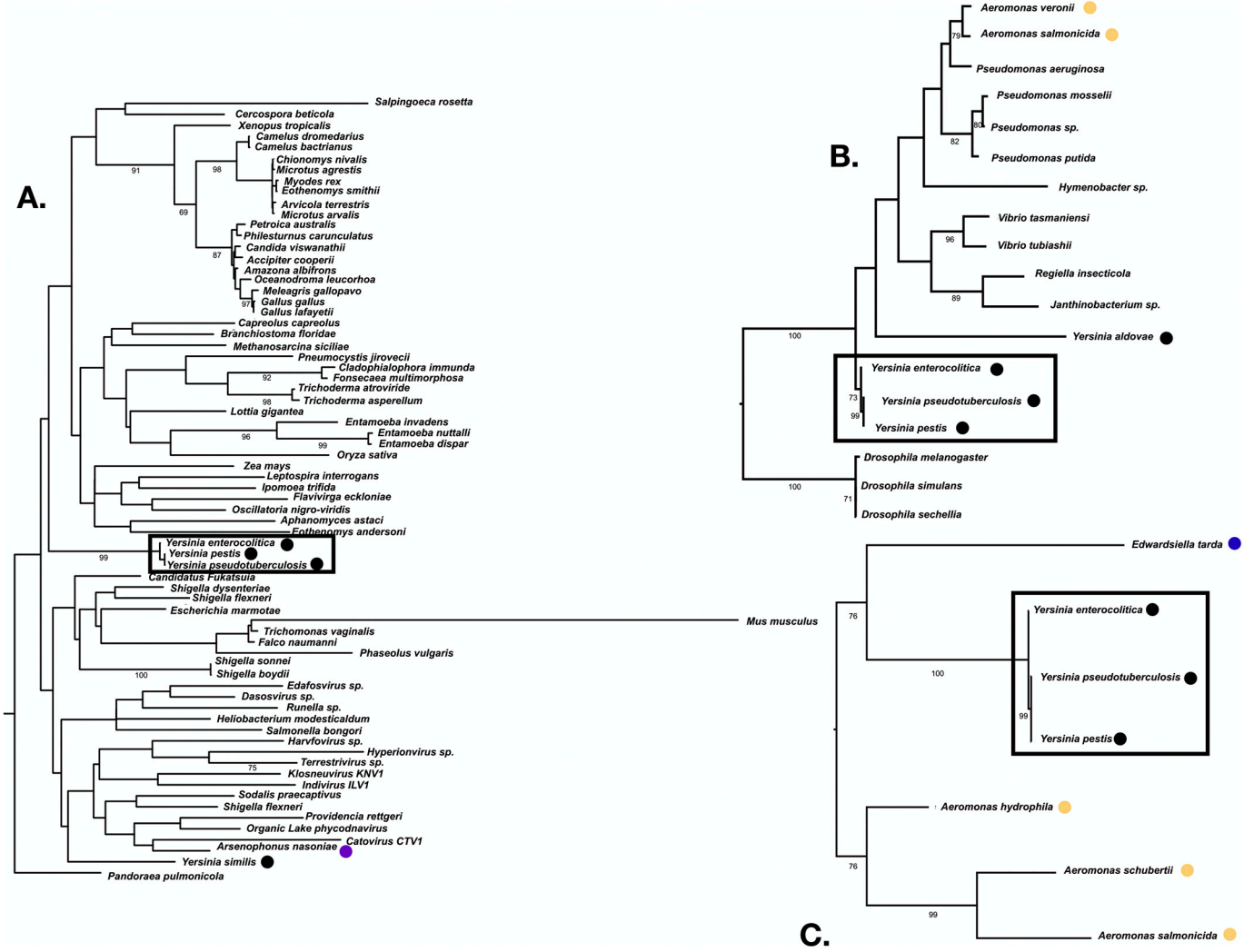

Fig. 2. Notable examples of phylogenetic trees of *Yersinia* and close relatives based on strict and 100-species datasets. These trees were based upon A) the full region of YopM, B) SycE, and C) and SycH. Bootstrap values greater than 70% under branches. Colored circles indicate common genera in multiple phylogenies (yellow is *Aeromonas*, red is *Erwinia*, purple is *Arsenophonus*, blue is *Edwardsiella*, and black is *Yersinia*). Black boxes highlight the three infectious species of *Yersinia*: *Y. pestis*, *Y. pseudotuberculosis*, and *Y. enterocolitica*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with *Y. enterocolitica* sister to this clade (66% BS). The other origin in SycE involves *Y. aldovae* as sister to this clade along with species of *Vibrio*, *Regiella*, and *Janthinobacterium* (92% BS).

Analysis of the YopM encoding gene resolved seven origins in *Yersinia*, and this is the largest number for any of the genes included in the present study. Three of these origins include multiple species of the genus, while others comprise one to a few species, such as one with *Y. aldovae* (91% BS), *Y. enterocolitica* and *Y. aldovae* (100% BS), or *Y. enterocolitica*, *Y. pestis*, and *Y. pseudotuberculosis* (100% BS, Fig. 2A). The sister relationships to all of these are dissimilar to those from the other studied genes. While analyses of the full-length YopM sequence resulted in seven origins of the gene in *Yersinia*; only one origin in *Yersinia* is identified for each of the N- and C-terminal domains of the gene. In analyses of the entire YopM gene, some origins are for monophyletic groups of *Yersinia* (e.g., 81% and 84% BS), but others resolve species of the genus in large clades with species of *Shigella* (e.g., 91% and 97% BS). This relationship is recovered in analyses of the N- and middle domains, but not the C-terminal domain, in which only species of *Yersinia* were included in the analyses. These relationships are not recovered in analyses of other plasmid genes.

In YopH, there are five origins of the gene for species of *Yersinia* among the many prokaryotes and eukaryotes that have homologs of this gene, with fewer, but still multiple origins, in the strict and 100-species datasets (Fig. 3A–C). Sampled individuals of *Y. enterocolitica*, *Y. pestis*, and *Y. pseudotuberculosis* are members of two clades (doi:https://doi.org/10.5061/dryad.905qftthh), with one origin for these species in the strict and 100-species datasets (Fig. 3A, B). One poorly supported clade is sister to species of *Protochlamydia* and *Waddlia* (23% BS), while the clade, which has high support, is sister to *Aeromonas*, with this clade sister to one that includes *Edwardsiella* (99% BS). Nonpathogenic species of *Yersinia* are members of other clades; although, *Y. enterocolitica* is included in these other clades as well. These phylogenetic relationships are similar, but not identical, to those of the C-terminal domain (Fig. 3B), which contains the active site. Notably, *Y. ruckeri* is allied with *Y. rohdei* and *Y. aleksiciae* in analyses of the entire gene (100% BS), but it is not closely related to other species of *Yersinia* in analyses of just the C-terminal domain. For both the full-length and the C-terminal domain, all genes present in the phylogenetic analyses (Fig. 3A and B) are known phosphatases.

The evolutionary relationships of the YopH N-terminal domain resolve *Y. enterocolitica*, *Y. pestis*, and *Y. pseudotuberculosis* in two clades, each of which is monophyletic, and species of *Aeromonas* are sister to the

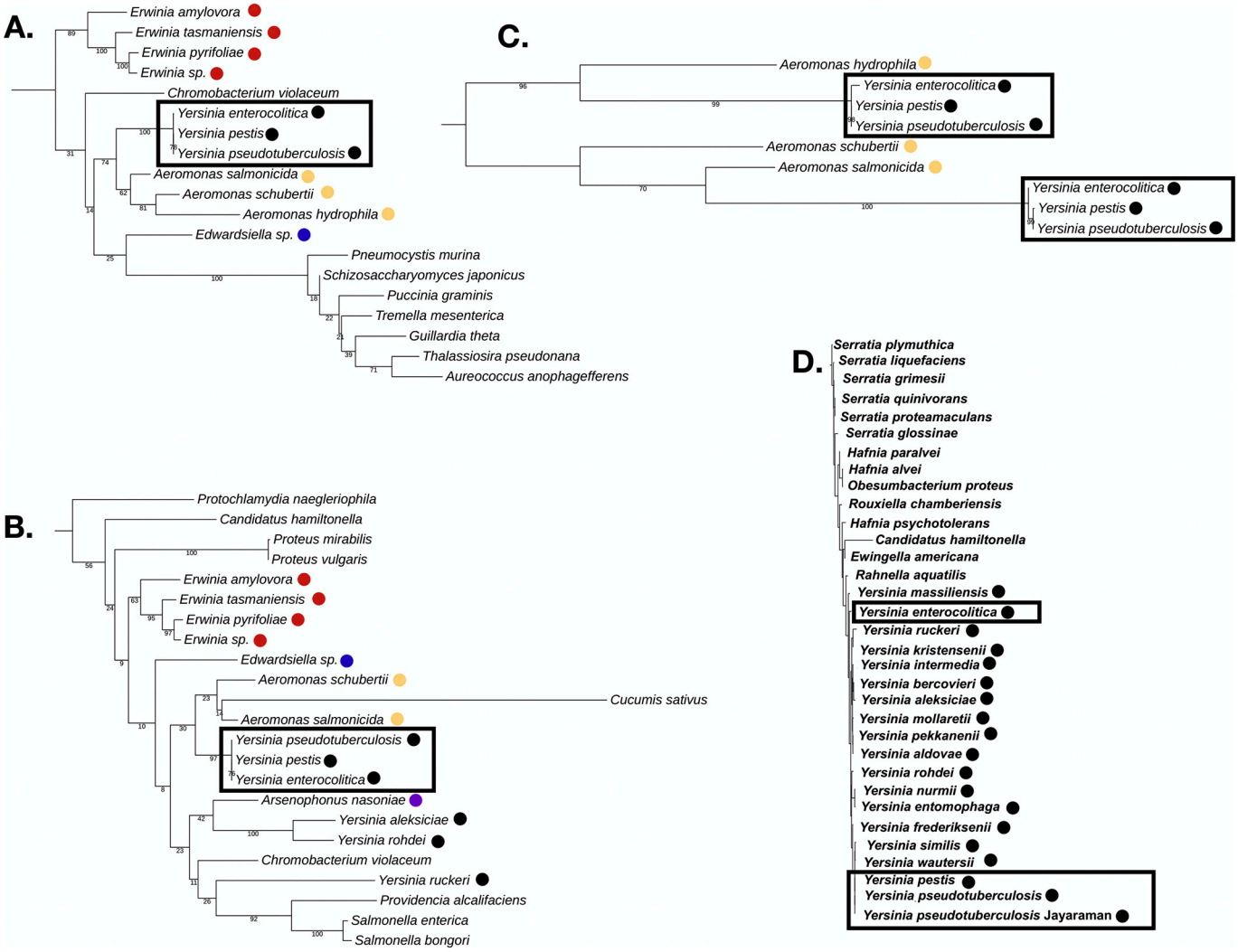

Fig. 3. YopH and 16S phylogenetic trees of *Yersinia* species and close relatives based on strict or 100-species datasets. These trees were based upon A) full region of YopH (based on 100 species), B) C-terminal (catalytic) domain of YopH (based on 100 species), C) N-terminal domain of YopH, and D) 16S rRNA region used for comparison. Bootstrap values greater than 70% under branches, but not 16S due to very short branch lengths. Colored circles indicate common genera in multiple phylogenies (yellow is *Aeromonas*, red is *Erwinia*, purple is *Arsenophonus*, blue is *Edwardsiella*, and black is *Yersinia*). Black boxes highlight the three infectious species of *Yersinia*: *Y. pestis*, *Y. pseudotuberculosis*, and *Y. enterocolitica*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

clades of pathogenic *Yersinia* (Fig. 3C). Interestingly, this second clade contains a known gene duplicate from pYV, YscM/LcrQ (Rimpilainen et al., 1992), a pattern that is not also recovered in analyses of the full-length sequence. The short linker region of the gene between the two domains resolves *Yersinia*, which only includes *Y. enterocolitica*, *Y. pestis*, and *Y. pseudotuberculosis*, as monophyletic and sister to a clade comprising *Thermococcus*, *Chrysochloris*, and *Camelus*.

The only DNA region analyzed as part of this study that is not located on the pYV plasmid was the rRNA 16S, and analyses of this region resolved one origin of *Yersinia* (71% BS, Fig. 3D). *Yersinia pestis* and *Y. pseudotuberculosis* are sisters, while *Y. enterocolitica* is sister to all species of the genus except *Y. massiliensis*. *Rahnella aquatilis* is sister to *Yersinia*, and *Ewingella*, *Hafnia*, *Rouxiella*, and *Serratia* are resolved as close relatives of *Yersinia*.

3.2. Patterns of selection

Analyses of selection with PAML and Selecton identified genes in which positive selection has occurred. In YopT, SycT, and SycH, the Ka/Ks ratio was less than one for clades of *Yersinia*; although, the two Syc genes had fewer than eight species per phylogeny. In analyses of YopH, YopJ, and YopM, the clade that included all species (most in YopJ) of *Yersinia* was under positive selection (Ka/Ks greater than one). Domains of YopH, YopM, and YopO yielded slightly different results, with only the C-domain of YopM having a Ka/Ks ratio greater than one for all species of *Yersinia*. For the N-domain of YopM, C-domain of YopO, and N- and C-domains of YopH, one species pair had a Ka/Ks ratio greater than one, while the genus values were less than one.

The strict or 100-species datasets of SycE, SycT, YopE, YopT, YopM, and YopH were all analyzed with Selecton, and the results of most analyses did not identify any amino acids that were under positive selection with all but one model of evolution. This was the case across all of the genes. The exception involved analyses with the Mechanistic Empirical Combination (MEC) model in which multiple amino acids of YopE, SycT, SycH, YopM, YopT, and YopH were under selection. This particularly was the case for YopE, and YopM in which ca. 10% of the residues were resolved as having high levels of selection as opposed to the other three genes in which only up to ca. 2% of the residues were under selection. For all genes, MEC was also determined, via Akaike Information Criterion (AIC), to be a preferred model compared to the

others tested with Selecton.

4. Discussion

4.1. Evolutionary history of the virulence effector genes

While there are a few dozen (Demeure et al., 2019) known virulence factors in *Yersinia pestis*, we studied all of the six virulence effectors that are common among *Y. pestis*, *Y. pseudotuberculosis*, and *Y. enterocolitica* and encoded on the mosaic plasmid pYV (YopM, YopH, YopE, YopO, YopT, and YopJ/P). Interestingly, of these six, only three are known to be necessary for virulence (Demeure et al., 2019). We were therefore intrigued by the possibility that different patterns of evolution were reconstructed between the necessary and not necessary virulence factors; however, the evolutionary relationships differed among studied genes. No taxon is resolved as sister or a close relative in more than half of the phylogenies of the *yop* genes, which suggests that the genes on the plasmid are the result of horizontal genetic transfer via multiple distinct lineages and origin events. Indeed, these relationships all differ from those resolved by 16S and by PATRIC. This provides evidence of the promiscuous nature concerning the construction of the pYV plasmid in the three infectious species of *Yersinia*, particularly as multiple *yop* genes are not present in other species of *Yersinia*. Because many of these genes are restricted to the three infectious species of *Yersinia*, one hypothesis is that these species form a monophyletic group, but that is not the case based on the 16S and PATRIC phylogenies. These results suggest that the plasmid was transferred between *Y. enterocolitica* and *Y. pestis* + *Y. pseudotuberculosis*, or, alternatively, that the plasmid was acquired early in the evolution of the genus and lost in most of the species; however, the former is a more parsimonious explanation than the latter. This is consistent with the scenario previously noted by other research groups using a different methodology (Howard et al., 2009).

One notable aspect of the results of the present study involves the absence of a particular trend: the stepwise origin of these suites of infectious genes. In the studied species of the genus, which involved broad searches of publicly available sequence data, the three Yop and Syc genes were restricted to *Y. pestis*, *Y. pseudotuberculosis*, and *Y. entercolitica*, with none of these genes present in other species of the genus. This provides evidence for the lack of a stepwise origin of the ability to be infectious in the genus (although, see section 4.4 for another explanation). Presently, this would suggest that other species of *Yersinia* will not become infectious due to the lack of any of the necessary and sufficient genes, but given the putative ability to acquire these genes rapidly, this may not be the case. The fact that on at least two occasions the infectious genes and their chaperones were appropriately combined on the plasmid, with multiple domains necessary for this to occur, continues to point to the role serendipity plays in the evolutionary process.

4.1.1. Evolutionary history of YopH

The gene *yopH* codes for the conventional tyrosine phosphatase YopH. It is known to be one of the most active phosphatases (Zhang et al., 1992) and while not sufficient for virulence in *Yersinia*, it is necessary (Bolin and Wolf-Watz, 1988). The evolutionary history of protein phosphatases in eukaryotes has been previously documented, and while it has been largely speculated that YopH and its bacterial homologs arose through horizontal gene transfer (Marks et al., 2009; Guan and Dixon, 1990), the precise origins of these prokaryotic genes remained elusive. To investigate this, the determined phylogenetic trees for the *yopH* gene as well as those from each individual domain are shown in Fig. 3A–C.

One line of evidence supporting this supposition is that while conventional phosphatases are common in eukaryotes, they are quite rare in prokaryotes. In addition to known YopH homologs (in *Salmonella* [SptP] (Kaniga et al., 1996), *Aeromonas* [AopH/AroH] (Najimi et al., 2009), *Erwinia* (Triplett et al., 2006)), only a few other prokaryotes have their own conventional protein-tyrosine phospahtases (Kennelly, 2002; Kennelly, 2001): *Nostoc commune* has IphP and *Mycobacterium tuberculosis* has MPtpB, while *Methanobacterium thermoautotrophicum*, *Methanococcus jannaschii*, *Pyrococcus hirokoshii*, *Bacillus subtilis*, *Deinococcus radiodurans*, *Escherichia coli* K-12, and *Vibrio cholerae* are all suspected to as well. In the present study, none of these suspected species are resolved as closely related to species of *Yersinia* based on the full-length YopH phylogeny, but the known homologs were more closely related in the reconstructed phylogeny for the C-terminal domain. Interestingly, however, species of *Aeromonas* and *Edwardsiella* are sister or a close relative, respectively, to the three infectious species of *Yersinia* (Fig. 3A–C). While these relationships have high support values, these genera have not previously been identified as having conventional protein phosphatases, which suggests that these species (and possibly others) may be harboring protein phosphatases across the genome and associated plasmids. While most genes and gene fragments included in the present study displayed Ka/Ks ratios of less than one, YopH was one of two with Ka/Ks ratios greater than one, indicating that positive selection was occurring that may have resulted in the gene producing such a potent phosphatase, at least in the three infectious species of the genus.

While YopH is best known from the three infectious species of *Yersinia*, nonpathogenic species of the genus are members of various clades in the YopH phylogenies, indicating that there is a conventional protein phosphatase homolog in the *Yersinia* genome (in contrast to on a virulence plasmid). YopH in the three infectious species certainly has a distinct origin from that of the phosphatases of other species in the genus, as evidenced by their distinct clades of the C-terminal domain (Fig. 3B) as well as only the three infectious species having an N-terminal domain (Fig. 3C). Additionally, the full phylogeny of full-length YopH incorporated eukaryotic phosphatase genes, including multiple human putative homologs, but these genes were not resolved as closely related to species of *Yersinia* (and thus are not included in Fig. 3). The well-known homologous nature of these genes compared with YopH are therefore not an indication of a close evolutionary relationship.

The N- and C- terminal domains of YopH are resolved to have independent origins (Fig. 3A–C). This is even though the known *Aeromonas* homolog (AopH/AroH) is sister in both the N-terminal and C-terminal phylogenies of YopH, demonstrating the possibility of various horizontal gene transfer scenarios, such as the entire gene transferred to *Yersinia* from *Aeromonas* (or vice-versa) or the domains transferred independently and appropriately combined to form the gene. While the former is more parsimonious than the latter, the amount of promiscuity appears sufficiently large among the plasmid genes (and their domains) to suggest that additional research should be undertaken to better understand which pattern may have occurred as well as the direction of the transfer. The known homolog from *Salmonella* (SptP), however, has been resolved such that its C-terminal domain shows sequence similarity to YopH's C-terminal domain, while its N-terminal domain shows sequence similarity to the N-terminal domain of *Yersinia*'s YopE, as has been identified by other groups (Kaniga et al., 1996).

4.1.2. Evolutionary history of YopO

The gene *yopO* codes for a serine/threonine (S/T) kinase that, while aiding in infection, is not typically cited as necessary for *Yersinia* virulence. Specifically, it is the middle portion of the gene (here denoted as the M-domain) that confirms kinase activity (Galyov et al., 1993), while the C-terminal domain acts in Rho GTPase–binding (Prehna et al., 2006). Like the conventional phosphatases described above, S/T kinases are common in eukaryotic organisms but rare in the prokaryotes, although less rare than the conventional phosphatases (Kennelly, 2002), and YopO has known homologs in eukaryotes (Galyov et al., 1993). Indeed, here both the full length and middle domain of YopO were identified as having homologs in eukaryotes, including *Drosophila melanogaster* and *Homo sapiens*. These relationships, however, are quite different for those resolved for the YopH C-terminal domain where only species of *Yersinia*, *Vibrio*, and *Aeromonas* were identified as having

genes with sequence similarity (Fig. 3B). Additionally, like *Aeromonas*'s *aopH* and *sycH* described above, the gene *aopO* (a *yopO* homolog) was identified as being present on *Aeromonas*'s pAsa5 plasmid (Tanaka et al., 2017). Interestingly, it is the kinase domain without eukaryotic homologs, as kinases are canonically eukaryotic, serine/threonine kinases likely existed before the divergence between prokaryotes and eukaryotes (Han and Zhang, 2001), and this activity may have been evolutionarily maintained between *Yersinia* and eukaryotes (Navarro et al., 2007) or have been transferred horizontally between prokaryotes and eukaryotes (Whitehead et al., 2013).

4.1.3. Evolutionary history of YopM

The gene *yopM* codes for the protein YopM, which translocates to the host without the aid of a chaperone, subsequently recruiting other proteins and further translocating to the host nucleus suppressing infection-induced activation (Chung et al., 2016). Molecular details of its role, however, are less known than those of most other effectors. Structurally, it contains three known domains: an N-terminal translocation domain (which does not require a chaperone), a C-terminal E3 ligase domain, and middle leucine-rich region. Interestingly, it is known that there are chromosomal paralogs to the YopM virulence protein within *Yersinia* genomes (Strauch et al., 2000), for which evidence was also seen here (Fig. 2A), with species of the genus not forming a monophyletic group. The evolutionary relationship between these *Yersinia* YopMs has been previously reported (Hu et al., 2016), but not any evolutionary relationship with non-*Yersinia* proteins. Here, however, it was found that the N- and M- domains showed sequence similarities with *Shigella*, although the C-terminal E3 ligase domain only resolved *Yersinia* paralogs. Like YopH, YopM displayed Ka/Ks ratios of greater than one providing evidence of positive selection for the gene, although this is primarily driven by the N-terminal domain, which could suggest that the ability to translocate is important for effective virulence of this protein. Additionally, the phylogenies of YopM and its domains provide evidence that while each terminal domain has separate homologs, each has may have been acquired from different ancestors (or, given the lack of homologs outside of *Yersinia* for the C-terminal domain, originated de novo), probably multiple times given the reconstruction of the middle domain. The distinct origins of the various domains likely resulted in the particular phylogenetic patterns of paraphyly for the gene itself while the N- and C-terminal domains are resolved as monophyletic for *Yersinia*. These results collectively suggests that the origin of this gene may follow a mix-and-match approach, particularly related to the middle domain. Indeed, it could be that multiple times the middle domain has linked together the N- and C-terminal domains to result in the present formation of the *yopM* gene among species of *Yersinia*.

4.2. Evolutionary history of the plasmid-encoded Yop chaperones (Syc) and their associated Yop binding domains

Some of the virulence factor Yops (specifically YopH, YopT, and YopE) require the use of chaperones (SycH, SycT, and SycE, respectively) to cross through the type three secretion system translocon and, therefore, into the host cell. Each of these chaperone-requiring Yops has an N-terminal signal sequence like all of the excreted Yops, but also has full N-terminal domains that bind to their respective chaperones. Consequently, we were interested in looking at patterns of co-evolution between Yops and their Syc chaperones among these three pairs.

The chaperones (SycT, SycH, and SycE) are only present in the three infectious species of *Yersinia*, *Y. enterocolitica*, *Y. pestis*, and *Y. pseudotuberculosis* species, and these three species form a monophyletic clade in phylogenies for each gene (although one sequence from *Y. aldovae* is present in SycE, although not part of the same clade) (Fig. 2B, C). While the sister to the clade of infectious *Yersinia* differs depending on the gene, only two to five genera include homologs of each of these chaperones (i.e., *Aeromonas*, *Edwardsiella*, *Janthinobacterium*, *Photorhabdus*, *Pseudomonas*, *Regiella*, and *Vibrio*). The N-terminal (chaperone-binding) domains of the relevant virulence factors were also analyzed and compared versus the chaperones themselves.

The N-terminal domain of YopE is only present in the three infectious species of *Yersinia*, even though the entire gene is present in these species as well as a few others in *Vibrio* and *Aeromonas*. Similar relationships are recovered for SycE, the chaperone protein, but this chaperone gene appears more widespread across Eubacteria, being present in the same genera as YopE as well as *Pseudomonas*, *Regiella*, and a species of *Janthinobacterium* (Fig. 2B). While this provides evidence for similar relationships between a Yop and its chaperone protein, the phylogenies of the two are not identical, suggesting that not only can genomic novelties arise in parts of genes (Alberts, 2002) but that Yop chaperones and their binding domains can each have distinct evolutionary histories. Indeed, because of this latter difference, it may be that virulence is only bestowed on individuals that can successfully acquire both via horizontal gene transfer.

SycH and the YopH N-terminal domain also resolve similar evolutionary relationships, although SycH is present in *Edwardsiella* and *Aeromonas* while the YopH N-terminal domain only includes individuals from the latter (Figs. 2C, 3C). The YopH N-terminal domain does appear, however, to have arisen via gene duplication, given the similarity, both in sequence and phylogenetic relationships, between this part of the gene and YscM, another gene on the pYV plasmid (Michiels et al., 1991). While SycH and the YopH N-terminal domain are only present in a small number of taxa (less than 10), the full-length YopH gene is homologous to species across the tree of life. Additionally, *Yersinia* is resolved in multiple distinct clades in analyses of the full-length YopH gene, with one clade having the relationship reconstructed as part of the N-terminal domain and others resolving *Erwinia*, *Arsenophonus*, *Candidatus*, or *Providencia* as close relatives. These taxa are not identified as having sequences homologous to the N-terminal domain. These results provide evidence that the N-terminal domain has an evolutionary history distinct from that of the C-terminal domain and the linker, as well as from the SycH chaperone.

In both SycT and YopT, a clade that includes *Photorhabdus* and another taxon is sister to a clade comprising the three infectious species of *Yersinia*, and these three species are monophyletic. This suggests that there is one common origin for each of these genes in the genus. Across prokaryotes, YopT is identified to have a greater number of homologs than SycT. In the N-terminal domain of YopT, only *Photorhabdus* is resolved as sister (100% BS), and it is the only other taxon that has homologous sequences to the three infectious species of *Yersinia*. This suggests that the N-terminal domain may have originated in one of the genera and was subsequently transferred to the other, but given the lack of additional taxa with homologs, the direction of transfer remains unclear.

Overall, YopE, YopT, SycE, SycH, SycT, and the N-terminal domain of YopH, have similar phylogenetic relationships that resolve the three infectious species of *Yersinia* as monophyletic, and these genes are not widespread across the prokaryotes. Indeed, most are only present in *Yersinia* and distant relatives, such as *Photorhabdus* and *Edwardsiella*. In contrast, the C-terminal domain of YopH is widespread across the three domains as well as has multiple origins of this gene in *Yersinia*. Additionally, YopH is the only one of these three genes in which the individual domains resolve distinct phylogenetic relationships.

4.3. Comparisons of plasmid-encoded gene phylogenies with 16S and PATRIC

In the present study, the 16S gene was used as an indicator of the evolutionary history of the organism's chromosome, which is reflective of the organism's phylogenetic history, and, despite its shortcomings, this chromosomal region is widely used for such purposes (Rajendhran and Gunasekaran, 2011). The reconstructed 16S tree is largely consistent with other phylogenies of *Yersinia* and relatives, such as that which is located at Pathosystems Resource Integration Center (PATRIC) (Davis

et al., 2020), in which *Yersinia* is monophyletic and species are resolved to be most closely related to *Candidatus* Regiella insecticola as well as *Serratia* and *Rahnella* species. Taxa that are closely related in plasmid-based phylogenies reconstructed in the present study, such as *Edwardsiella*, however, are not resolved in similar relationships in the tree based on 16S sequence data or from PATRIC. The *Yersinia* chromosome and the pYV plasmid genes do not share a common evolutionary history.

4.4. Homology identification bias

Each of the genes (and domain fragments) studied recovers an evolutionary history different than that of the chromosome. These data, however, are based on sequences identified via the similarity criteria of BLAST searches. There is therefore a question of whether a lack of identified homologs in other species is due to a lack of homologous sequences or a lack of sufficient genomic data. Interestingly, all three major *Yersinia pestis* biovars have had their entire genome sequenced (Chain et al., 2006; Parkhill et al., 2001; Deng et al., 2002) and so have at least 8 other species of the genus (*Y. aldovae* (Johnson et al., *2015*), *Y. aleksiciae* (Sprague et al., 2015), *Y. frederiksenii* (Johnson et al., *2015*), *Y. intermedia* (Johnson et al., *2015*), *Y. kristensenii* (Johnson et al., 2015), *Y. rohdei* (Johnson et al., *2015*), *Y. ruckeri* (Johnson et al., 2015), *Y. similis* (Sprague and Neubauer, 2014)). Additionally, the genera resolved to be most closely related based on the 16S chromosomal gene and via PATRIC, including *Serratia*, *Candidatus*, and *Rahnella*, have all had their complete genomes sequenced (see (Khan et al., 2017; Degnan et al., 2010; Martinez et al., 2012) for example). This means that a lack of identity with genes in these species is indicative of a lack of homologous regions of the genome. At the same time, plasmids are not usually included as part of sequenced and reported genomes (Gan et al., 2020), which could result in genes that are frequently on plasmids excluded from NCBI and, consequently, the BLAST searches undertaken for the present study. This potential for bias could skew the results herein presented and, therefore, provides encouragement to the research community to assemble not just genomes but also related plasmids to better understand patterns of evolution for important genes located on accessory chromosomes. Future work on the phylogenetics of *Yersinia* and YopH will focus on mapping genome sequences to plasmids as well as de novo assembly of plasmids for phylogeny reconstruction.

5. Conclusion

The pYV plasmid provides an intriguing genomic region for phylogenetic studies, and this is due to the mosaic nature of the plasmid itself as well as the genes of which it is composed. Examinations of the evolutionary history of genes of the plasmid and the various domains of the genes demonstrate that the gene combinations and synteny appear to be the result of promiscuity and fortuity. The former culminates in the various genes and domains combined, and the latter results in these being some of the necessary and sufficient genes for not just virulence but particularly effective pathogenicity of species of *Yersinia*. Additionally, the present study demonstrates the importance of sequencing and assembling plasmids so the biological community can examine the intriguing patterns of genes present on them.

Funding

The research was partially funded by a Major Research Instrumentation grant from the National Science Foundation (USA), MRI-1725938. This research did not receive any additional grants from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

None.

Acknowledgments

The authors wish to acknowledge the help of multiple undergraduate students in the beginning, exploratory phases of this project.

References

Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 38 (Web Server issue):W7-13.
Achtman, M., Zurth, K., Morelli, G., Torrea, G., Guiyoule, A., Carniel, E., 1999. Yersinia pestis, the cause of plague, is a recently emerged clone of Yersinia pseudotuberculosis. Proc. Natl. Acad. Sci. U. S. A. 96 (24), 14043–14048.
Alberts, B., 2002. Molecular Biology of the Cell, 4th ed. Garland Science, New York.
Bochkareva, O.O., Dranenko, N.O., Ocheredko, E.S., et al., 2018. Genome rearrangements and phylogeny reconstruction in Yersinia pestis. PeerJ. 6, e4545.
Bolin, I., Wolf-Watz, H., 1988. The plasmid-encoded Yop2b protein of Yersinia pseudotuberculosis is a virulence determinant regulated by calcium and temperature at the level of transcription. Mol. Microbiol. 2 (2), 237–245.
Chain, P.S.G., Hu, P., Malfatti, S.A., et al., 2006. Complete genome sequence of Yersinia pestis strains Antiqua and Nepal516: evidence of gene reduction in an emerging pathogen. J. Bacteriol. 188 (12), 4453–4463.
Chung, L.K., Park, Y.H., Zheng, Y., et al., 2016. The Yersinia virulence factor YopM hijacks host kinases to inhibit type III effector-triggered activation of the pyrin inflammasome. Cell Host Microbe 20 (3), 296–306.
Cornelis, G.R., Boland, A., Boyd, A.P., et al., 1998. The virulence plasmid of Yersinia, an antihost genome. MMBR. 62 (4), 1315–1352.
Davis, J.J., Wattam, A.R., Aziz, R.K., et al., 2020. The PATRIC bioinformatics resource center: expanding data and analysis capabilities. Nucleic Acids Res. 48 (D1), D606–D612.
Degnan, P.H., Leonardo, T.E., Cass, B.N., et al., 2010. Dynamics of genome evolution in facultative symbionts of aphids. Environ. Microbiol. 12 (8), 2060–2069.
Demeure, C.E., Dussurget, O., Fiol, G.M., Le Guern, A.S., Savin, C., Pizarro-Cerda, J., 2019. Yersinia pestis and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. Genes Immun. 20 (5), 357–370.
Deng, W., Burland, V., Plunkett 3rd, G., et al., 2002. Genome sequence of Yersinia pestis KIM. J. Bacteriol. 184 (16), 4601–4611.
Doron-Faigenboim, A., Stern, A., Mayrose, I., Bacharach, E., Pupko, T., 2005. Selecton: a server for detecting evolutionary forces at a single amino-acid site. Bioinformatics. 21 (9), 2101–2103.
Duan, R., Liang, J., Shi, G., et al., 2014. Homology analysis of pathogenic Yersinia species Yersinia enterocolitica, Yersinia pseudotuberculosis, and Yersinia pestis based on multilocus sequence typing. J. Clin. Microbiol. 52 (1), 20–29.
Galyov, E.E., Hakansson, S., Forsberg, A., Wolf-Watz, H., 1993. A secreted protein kinase of Yersinia pseudotuberculosis is an indispensable virulence determinant. Nature. 361 (6414), 730–732.
Gan, Y.Q., Zhang, T., Gan, Y.Q., Zhao, Z., Zhu, B., 2020. Complete genome sequences of two enterococcus faecium strains and comparative genomic analysis. Exp. Ther. Med. 19 (3), 2019–2028.
Guan, K.L., Dixon, J.E., 1990. Protein tyrosine phosphatase activity of an essential virulence determinant in Yersinia. Science. 249 (4968), 553–556.
Han, G., Zhang, C.C., 2001. On the origin of Ser/Thr kinases in a prokaryote. FEMS Microbiol. Lett. 200 (1), 79–84.
Howard, S.L., Strong, P.C.R., Wren, B.W., 2009. The three bears and virulence-associated plasmids in the genus Yersinia. In: Schwartz, E. (Ed.), Microbiology Monographs. Springer, Berlin, pp. 171–186.
Hu, P., Elliott, J., McCready, P., et al., 1998. Structural organization of virulence-associated plasmids of Yersinia pestis. J. Bacteriol. 180 (19), 5192–5202.
Hu, Y., Huang, H., Hui, X., et al., 2016. Distribution and evolution of Yersinia leucine-rich repeat proteins. Infect. Immun. 84 (8), 2243–2254.
Iriarte, M., Cornelis, G.R., 1998. YopT, a new Yersinia Yop effector protein, affects the cytoskeleton of host cells. Mol. Microbiol. 29 (3), 915–929.
Johnson, S.L., Daligault, H.E., Davenport, K.W., et al., 2015. Thirty-two complete genome assemblies of nine Yersinia species, including Y. pestis, Y. pseudotuberculosis, and Y. enterocolitica. Genom. Announc. 3 (2).
Kaniga, K., Uralil, J., Bliska, J.B., Galan, J.E., 1996. A secreted protein tyrosine phosphatase with modular effector domains in the bacterial pathogen Salmonella typhimurium. Mol. Microbiol. 21 (3), 633–641.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30 (4), 772–780.
Kennelly, P.J., 2001. Protein phosphatases–a phylogenetic perspective. Chem. Rev. 101 (8), 2291–2312.
Kennelly, P.J., 2002. Protein kinases and protein phosphatases in prokaryotes: a genomic perspective. FEMS Microbiol. Lett. 206 (1), 1–8.
Khan, A.R., Park, G.S., Asaf, S., Hong, S.J., Jung, B.K., Shin, J.H., 2017. Complete genome analysis of Serratia marcescens RSC-14: a plant growth-promoting bacterium that alleviates cadmium stress in host plants. PLoS One 12 (2), e0171534.
Marks, F., Klingmüller, U., Müller-Decker, K., 2009. Cellular Signal Processing : An Introduction to the Molecular Mechanisms of Signal Transduction. Garland Science, New York.
Martinez, R.J., Bruce, D., Detter, C., et al., 2012. Complete genome sequence of Rahnella aquatilis CIP 78.65. J. Bacteriol. 194 (11), 3020–3021.

Michiels, T., Vanooteghem, J.C., Lambert de Rouvroit, C., et al., 1991. Analysis of virC, an operon involved in the secretion of Yop proteins by Yersinia enterocolitica. J. Bacteriol. 173 (16), 4994–5009.
Najimi, M., Balado, M., Lemos, M.L., Osorio, C.R., 2009. Genetic characterization of pAsa6, a new plasmid from Aeromonas salmonicida subsp salmonicida that encodes a type III effector protein AopH homolog. Plasmid. 61 (3), 176–181.
Navarro, L., Koller, A., Nordfelth, R., Wolf-Watz, H., Taylor, S., Dixon, J.E., 2007. Identification of a molecular target for the Yersinia protein kinase a. Mol. Cell 26 (4), 465–477.
Parkhill, J., Wren, B.W., Thomson, N.R., et al., 2001. Genome sequence of Yersinia pestis, the causative agent of plague. Nature. 413 (6855), 523–527.
Pesesky, M.W., Tilley, R., Beck, D.A.C., 2019. Mosaic plasmids are abundant and unevenly distributed across prokaryotic taxa. Plasmid. 102, 10–18.
Pilla, G., Tang, C.M., 2018. Going around in circles: virulence plasmids in enteric pathogens. Nat. Rev. Microbiol. 16 (8), 484–495.
Prehna, G., Ivanov, M.I., Bliska, J.B., Stebbins, C.E., 2006. Yersinia virulence depends on mimicry of host rho-family nucleotide dissociation inhibitors. Cell. 126 (5), 869–880.
Rajendhran, J., Gunasekaran, P., 2011. Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. Microbiol. Res. 166 (2), 99–110.
Reuter, S., Connor, T.R., Barquist, L., et al., 2014. Parallel independent evolution of pathogenicity within the genus Yersinia. Proc. Natl. Acad. Sci. U. S. A. 111 (18), 6768–6773.
Rimpilainen, M., Forsberg, A., Wolfwatz, H., 1992. A novel protein, Lcrq, involved in the low-calcium response of Yersinia-pseudotuberculosis shows extensive homology to YopH. J. Bacteriol. 174 (10), 3355–3363.
Snellings, N.J., Popek, M., Lindler, L.E., 2001. Complete DNA sequence of Yersinia enterocolitica serotype 0 : 8 low-calcium-response plasmid reveals a new virulence plasmid-associated replicon. Infect. Immun. 69 (7), 4627–4638.
Sprague, L.D., Neubauer, H., 2014. Genome sequence of Yersinia similis Y228T, a member of the Yersinia pseudotuberculosis complex. Genom. Announc. 2 (2).
Sprague, L.D., Neubauer, H., et al., 2015. Genom. Announc. 3 (5).
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30 (9), 1312–1313.
Stern, A., Doron-Faigenboim, A., Erez, E., Martz, E., Bacharach, E., Pupko, T., 2007. Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. Nucleic Acids Res. 35 (Web Server issue):W506-511.
Strauch, E., Hoffmann, B., Heins, G., Appel, B., 2000. Isolation of a new insertion element of Yersinia intermedia closely related to remnants of mobile genetic elements present on Yersinia plasmids harboring the Yop virulon. FEMS Microbiol. Lett. 193 (1), 37–44.
Tan, S.Y., Tan, I.K.P., Tan, M.F., Dutta, A., Choo, S.W., 2016. Evolutionary study of Yersinia genomes deciphers emergence of human pathogenic species. Sci. Rep. 6.
Tanaka, K.H., Vincent, A.T., Emond-Rheault, J.G., Adamczuk, M., Frenette, M., Charette, S.J., 2017. Plasmid composition in *Aeromonas salmonicida subsp. salmonicida* 01-B526 unravels unsuspected type three secretion system loss patterns. BMC Genomics 18 (1), 528.
Triplett, L.R., Zhao, Y., Sundin, G.W., 2006. Genetic differences between blight-causing Erwinia species with differing host specificities, identified by suppression subtractive hybridization. Appl. Environ. Microbiol. 72 (11), 7359–7364.
Trulzsch, K., Roggenkamp, A., Aepfelbacher, M., Wilharm, G., Ruckdeschel, K., Heesemann, J., 2003. Analysis of chaperone-dependent Yop secretion/translocation and effector function using a mini-virulence plasmid of Yersinia enterocolitica. Int. J. Med. Microbiol. Suppl. 293 (2–3), 167–177.
Wattiau, P., Cornelis, G.R., 1993. Syce, a chaperone-like protein of Yersinia-enterocolitica involved in the secretion of YopE. Mol. Microbiol. 8 (1), 123–131.
Whitehead, M.P., Hooley, P., MR WB., 2013. Horizontal transfer of bacterial polyphosphate kinases to eukaryotes: implications for the ice age and land colonisation. BMC Res. Notes. 6, 221.
Winkler, R., 2020. Plasmid Drawing Program: Plasmidomics 0.2. http://www.ira.cinvestav.mx/Investigaci%C3%B3n/Biotecnolog%C3%ADayBioqu%C3%ADmica/ProfesoresInvestigadores/DrRobertWinkler/Software/PlasmidDrawingProgramPlasmidomics02/tabid/778/language/en-US/Default.aspx. Accessed.
Woestyn, S., Sory, M.P., Boland, A., Lequenne, O., Cornelis, G.R., 1996. The cytosolic SycE and SycH chaperones of Yersinia protect the region of YopE and YopH involved in translocation across eukaryotic cell membranes. Mol. Microbiol. 20 (6), 1261–1271.
Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13 (5), 555–556.
Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24 (8), 1586–1591.
Zhang, Z.Y., Clemens, J.C., Schubert, H.L., et al., 1992. Expression, purification, and physicochemical characterization of a recombinant Yersinia protein tyrosine phosphatase. J. Biol. Chem. 267 (33), 23759–23766.