‘Add, stir and reduce’: *Yersinia* spp. as model bacteria for pathogen evolution

Alan McNally[1], Nicholas R. Thomson[2], Sandra Reuter[3] and Brendan W. Wren[4]

Abstract | Pathogenic species in the *Yersinia* genus have historically been targets for research aimed at understanding how bacteria evolve into mammalian pathogens. The advent of large-scale population genomic studies has greatly accelerated the progress in this field, and *Yersinia pestis*, *Yersinia pseudotuberculosis* and *Yersinia enterocolitica* have once again acted as model organisms to help shape our understanding of the evolutionary processes involved in pathogenesis. In this Review, we highlight the gene gain, gene loss and genome rearrangement events that have been identified by genomic studies in pathogenic *Yersinia* species, and we discuss how these findings are changing our understanding of pathogen evolution. Finally, as these traits are also found in the genomes of other species in the Enterobacteriaceae, we suggest that they provide a blueprint for the evolution of enteropathogenic bacteria.

The ever-expanding number of sequenced microbial genomes, combined with an increased understanding of the pathogenesis and ecology of bacterial pathogens, is illuminating the dynamics that have shaped the evolution of mammalian pathogenesis in these organisms. Many bacterial pathogens infect humans only incidentally and often produce virulence factors that are effective against non-mammalian organisms as diverse as insects, protozoa, nematodes, predatory bacteria and phages[1–3]. These natural hosts provide a considerable driving force for the evolution of bacterial pathogens that also infect humans, enabling us to re-evaluate human–pathogen interactions in the light of the eco–evo perspective; that is, a perspective of bacterial ecology and evolution in which organisms are evaluated broadly in the light of evolution and ecology, rather than the narrow constraints of their behaviour during human infection. When viewed from the eco–evo perspective, it is not surprising that genes encoding factors that promote virulence in mammalian infections are also found in bacteria that infect or colonize non-mammalian hosts[4,5]. Conversely, as bacteria evolve to inhabit new hosts, or new niches within the same host, genomics can reveal remnants of genes and gene clusters that may have provided an adaptive advantage in the past, but are no longer selected for and so have been degraded into non-functional pseudogenes.

Therefore, many bacteria seem to be ‘pre-armed’ with virulence factors as they enter new environments and niches, and their evolution is not solely shaped by the pathogenic interactions with human, or mammalian, hosts[6,7]. The eco–evo perspective is highly pertinent for the evolution of the *Yersinia* genus, as it includes both non-pathogenic and pathogenic species with diverse phenotypes. Furthermore, all 18 species in the genus have recently been fully sequenced and metabolically phenotyped, which has enabled comparative analyses to be carried out between pathogenic and non-pathogenic species[8]. As such, *Yersinia* spp. represent a key model for understanding the forces that shape the evolution of pathogenic bacteria. Our current understanding of the genus encompasses both long-term evolution — how pathogenic species with a broad host range, such as *Yersinia enterocolitica*, evolve over millions of years from a non-pathogenic ancestor[9] — and short-term evolution — how recent evolutionary bottlenecks have led to the rapid emergence of highly pathogenic clones, such as the emergence of *Yersinia pestis* from *Yersinia pseudotuberculosis*[10].

The evolution of pathogenic *Yersinia* species shows remarkable similarities to the evolution of species in the *Salmonella* genus. Similarly to the *Yersinia* genus, serovars in the *Salmonella* genus vary in host restriction and have varying degrees of pathogenic potential. These include *Salmonella enterica* subsp. *enterica* serovar Typhimurium, which has a broad host range and causes self-limiting diarrhoea, and the human-restricted *Salmonella enterica* subsp. *enterica* serovar Typhi, which can cause acute invasive disease. These similarities made pathogenic *Yersinia* species the organisms of choice for much of the pioneering work that used classical genetics to investigate mechanisms of bacterial pathogenesis. For example, studies using *Y. enterocolitica* spearheaded research into bacterial invasion of host epithelial cells[11], studies using *Y. enterocolitica* and *Y. pestis* advanced our understanding of the role of plasmids in

[1]*Pathogen Research Group, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK.*
[2]*Pathogen Genomics, Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK.*
[3]*Department of Medicine, University of Cambridge, Box 157 Addenbrooke’s Hospital, Hills Road, Cambridge CB2 2QQ, UK.*
[4]*Department of Pathogen Molecular Biology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK.*
Correspondence to A.M.
alan.mcnally@ntu.ac.uk

doi:10.1038/nrmicro.2015.29
Published online 15 Feb 2016

bacterial virulence[12,13], and *Y. pseudotuberculosis* was used to identify and characterize type III secretion systems (T3SSs) and the role of these systems in host subversion[14]. These findings have led to advances in our understanding of how bacteria evade host responses during infection[15–17].

In this Review, we examine how the eco–evo perspective, informed by whole-genome sequencing and population genomic and ecological studies, has challenged previous hypotheses on the evolution of mammalian pathogenesis in the *Yersinia* genus. Furthermore, we consider how these studies are revealing common principles in the emergence of pathogenesis in other species in the Enterobacteriaceae. Finally, we show why including all members of a genus in population genomic studies, rather than pathogenic species alone, can provide greater insight into, and resolution for, understanding the evolution of mammalian pathogenesis.

The *Yersinia* genus

Yersinia is a genus in the Enterobacteriaceae family of Gram-negative bacteria. The nomenclature of the genus is based on classical systematics and biochemical species-classification methods, which resulted in the description of the genus as being a highly diverse group of bacteria with 18 defined species[18–26]. The best characterized of these are the three species that cause disease in mammals (including humans), *Y. enterocolitica*, *Y. pseudotuberculosis* and *Y. pestis*. Both *Y. enterocolitica* and *Y. pseudotuberculosis* are zoonotic pathogens that cause self-limiting gastroenteritis in humans, whereas *Y. pestis*, the causative agent of plague, is a pathogen in rodents and fleas that is occasionally transmitted to humans[27]. The remaining 15 known species of the genus are commonly found in soil and aquatic environments and are generally non-pathogenic in mammals (although some are pathogenic in other hosts, as *Yersinia ruckeri* causes red mouth disease in salmonids[21] and *Yersinia entomophaga* has insecticidal activity)[24]. Recent phylogenomic analysis of the entire *Yersinia* genus has enabled an accurate, whole-genome single-nucleotide-polymorphism (SNP)-based assessment of its population structure[8], which has shown that the genus contains 14 distinct species clusters. This robust sequenced-based taxonomy differed from the standing taxonomic description of the genus, which was largely formulated on the basis of biochemical differences and 16S ribosomal gene phylogeny (BOX 1). It also showed that the mammalian pathogens *Y. enterocolitica* and the *Y. pseudotuberculosis*–*Y. pestis* species complex, which belongs to a species cluster that also comprises the non-pathogenic *Yersinia similis* and *Yersinia wautersii*, form separate branches at opposite ends of the *Yersinia* evolutionary tree and do not branch together as was previously thought[8].

Early microbial population genetics studies suggested that *Y. pestis* is a recently emerged clone of *Y. pseudotuberculosis,* even though it has a markedly different lifestyle to *Y. pseudotuberculosis* and causes a much more acute disease[28,29]. These studies inferred a history for *Y. pestis* in which it evolved from *Y. pseudotuberculosis* in an evolutionary 'blink of an eye', estimated to be between 2,000–10,000 years ago[28,30], through a combination of gene gains, gene losses and genome rearrangements. The evolutionary events that led to pathogenesis were further explored in early microbial comparative genomics studies that used the first available sequenced genomes. These revealed that *Y. pestis* had emerged from *Y. pseudotuberculosis* through a mixture of gene gain events — in the form of the acquisition of species-specific plasmids, bacteriophage, integrons and genomic islands — and gene loss events, in which the loss of complete or partial genes led to a reduction in genomic flexibility and metabolic streamlining. These gains and losses of genomic features are associated with a change in lifestyle and, ultimately, niche restriction[10,31–33]. The relationship between *Y. enterocolitica* and the *Y. pseudotuberculosis*–*Y. pestis* species complex was until recently poorly defined. In 2006, a comparison of the available genomes from all three species suggested that they shared a distant but common pathogenic ancestor that had acquired key determinants of pathogenicity before splitting into distinct lineages[9,27]. These determinants of pathogenicity included the essential virulence plasmid pYV, which encodes the Ysc T3SS (BOX 2). However, the availability of just a single *Y. enterocolitica* genome sequence at this time, and a lack of genome sequences for non-pathogenic *Yersinia* species, limited the interpretations that could be made from comparative genomic analyses. Further genomes were subsequently published for *Y. enterocolitica*, in 2011 and 2013 (REFS 34–37), as well as single genomes for type strains of environmental *Yersinia* species, in 2010 (REF. 38). However, our understanding of the evolution of the genus as a whole, and the evolutionary position of *Y. enterocolitica* within it, was still limited by the lack of comprehensive whole-genome data. Nonetheless, even with this limited genomic data, a combination of gene gain and gene loss in the lineages of the three human pathogenic *Yersinia* species was evident and highlighted the complexity of the evolution of mammalian pathogenicity in these bacteria, which showed that pathogens can evolve by a process of 'add DNA, stir, and reduce' (REF. 27).

The study of the evolution of pathogenesis in the *Yersinia* genus has also suggested an inexorable relationship between the loss of metabolic function by genome decay, a reduction in ecological flexibility, and the evolution towards niche-restricted highly pathogenic lineages[27]. As population genomic studies have become an established tool in microbiology, the level of resolution at which we can study these evolutionary events has improved considerably. By sequencing the genomes of entire bacterial genera, it is possible to determine the key gene acquisition and loss events in the evolution of bacterial pathogens, thereby shedding light on the evolutionary forces and patterns that underpin these processes, and identify the common features of this evolution in diverse lineages.

The role of gene gain

The evolutionary path from environmental or commensal bacterium to human pathogen has long been associated with the gain of mobile genetic elements by horizontal gene transfer, which occurs through phage transduction, plasmid uptake, natural transformation or DNA conjugation. The most important of these gene gain events in *Yersinia* spp. seems to have been the acquisition

Type III secretion systems (T3SSs). Needle-like protein complexes used by some bacteria to export proteins. T3SSs are often responsible for the translocation of bacterial effector proteins from pathogenic or symbiotic bacteria directly into the cytoplasm of a host cell.

Salmonids Fish from the *Salmonidae* family, which includes salmon and trout.

Phylogenomic analysis The analysis of evolutionary trees created from whole-genome sequences. In contrast to the analysis of single-gene phylogenetic markers, such as the ribosomal 16S gene, phylogenomic analysis can infer high-resolution evolutionary patterns.

Integrons Cassettes of genes incorporated into loci encoding a site-specific recombinase, a recombination recognition site and a promoter. Integrons are often found in conjunction with other genes, such as antibiotic resistance genes.

Genomic islands Large genetic regions, acquired by horizontal gene transfer, that encode one or more functional groups of genes. They are frequently associated with tRNA genes, are flanked by repeat structures and contain mobility genes encoding integrases or transposases that are required for chromosomal integration and excision.

Horizontal gene transfer The transfer of DNA, frequently cassettes of genes, between organisms.

Transduction The phage-mediated transfer of DNA, frequently cassettes of genes, between organisms.

Natural transformation The direct uptake of DNA from the environment and the incorporation of this genetic material into the chromosome by competent cells.

Conjugation The transfer of DNA — usually plasmids — between organisms through direct cell-to-cell contact or through a bridge that forms between cells.

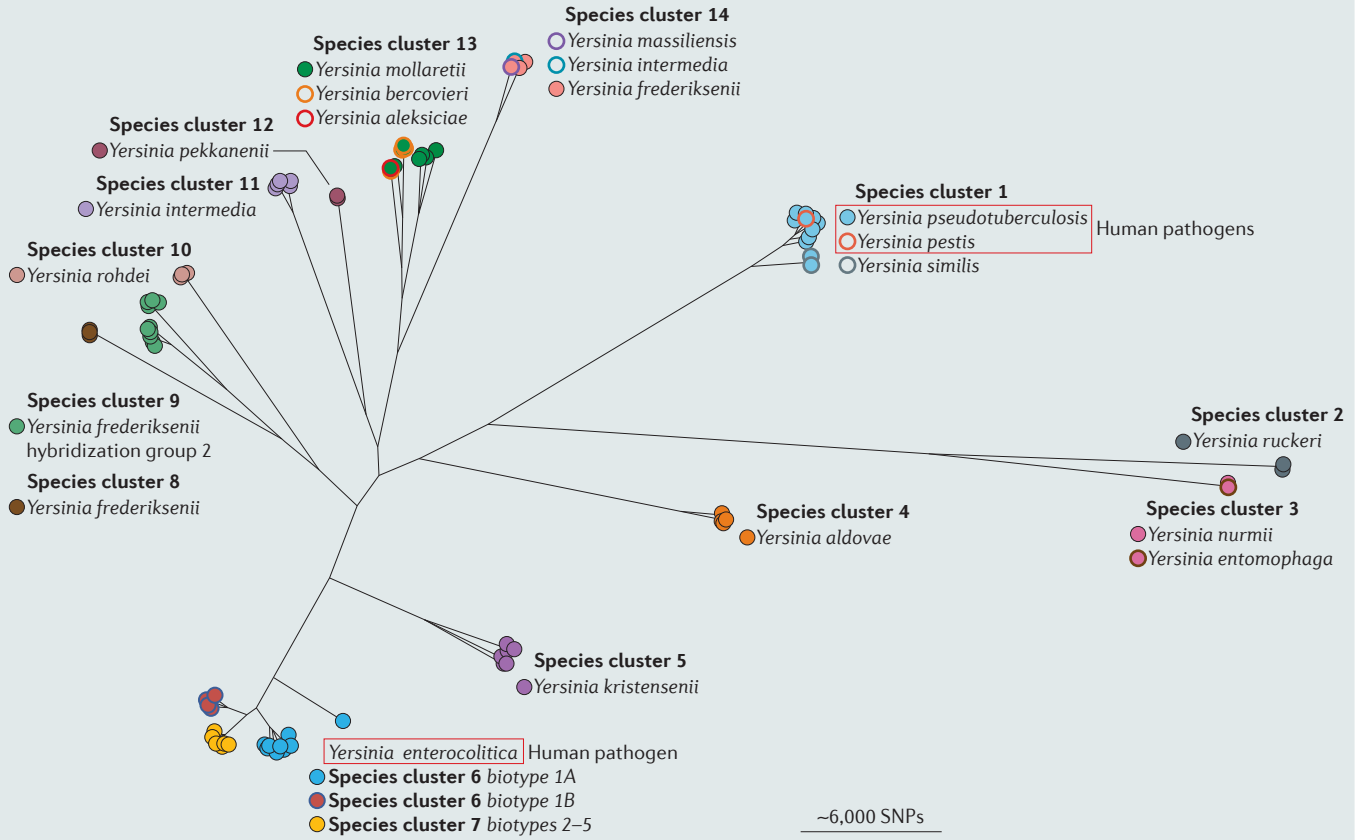
of the family of pYV virulence plasmids that encode the Ysc T3SS[8,12,13,39]. The Ysc T3SS delivers *Yersinia* outer proteins (Yops) directly into host cells on contact, a process termed the '*Yersinia* deadly kiss' (REF. 40). Delivery of Yops into host immune cells inhibits actin polymerization, thereby preventing phagocytosis of bacterial cells, and suppresses the transcription of genes of the innate immune response. The downregulation of the production of pro-inflammatory cytokines enables the proliferation of the infection[40–42].

Box 1 | **Phylogenomics of *Yersinia* species: approaches and main findings**

Traditionally, bacterial species are identified using limited biochemical tests that are based on the phenotypic expression or non-expression of a trait. However, the formulation of standardized methods to measure phenotypes is often problematic, and bacteria can be incorrectly identified when a mutation arises that alters the phenotype. Genomic information is less ambiguous than biochemical tests as a basis for species identification, and variations within genes under neutral selection can be used to infer evolutionary relationships, which provide a robust framework that is independent of phenotypic variation. For the genus *Yersinia*, 84 identified housekeeping genes show between 10–25% divergence in single-nucleotide polymorphisms (SNPs) between *Yersinia pestis* and *Yersinia enterocolitica*[8]. This level of SNP divergence meets the criteria used for the design of multi-locus sequence typing (MLST)[112], and the application of MLST can survey the accumulation of variation in the genus to determine its evolutionary history and to estimate a maximum likelihood phylogeny. Furthermore, Bayesian statistical analysis of the population structure can be used to identify and define species clusters[8] (see the figure). In this analysis, statistical probabilities are calculated to describe the sequence variation present within the population and the set of housekeeping genes are not only considered as a whole but also as separate gene entities. Some of the species clusters that have been identified by Bayesian statistical analysis contain several strains that conventional techniques had previously defined as separate species, such as species cluster 13 and species cluster 14. In other cases, strains that were previously defined as a single species are distributed over several species clusters; for example, *Yersinia frederiksenii* strains are found in species cluster 8, species cluster 9 and species cluster 14. These incongruities between species classifications using conventional techniques and the species clusters defined by Bayesian analyses highlight the limitations of biochemical tests that fail to accurately reflect the relationships between different species. Bayesian analysis positions *Y. enterocolitica* at the diametrically opposite end of the tree to the other two human pathogenic *Yersinia* species, *Yersinia pseudotuberculosis* and *Y. pestis*. The separated positions of the *Y. pseudotuberculosis*–*Y. pestis* species complex and *Y. enterocolitica* support the hypothesis that pathogenic potential has evolved independently at least twice in the *Yersinia* genus.

However, it must be noted that purely genomic evidence might not always be sufficient to define species boundaries, as *Y. pestis* and *Y. pseudotuberculosis* are highly similar on a genomic level but exhibit markedly different mechanisms of disease as well as niche preference and lifestyle.

Phylogenomics has also helped to clarify the position of the fish pathogen *Yersinia ruckeri*, which forms an outlying species cluster (species cluster 2). Even though no classical root of the tree is given, *Y. ruckeri* might be considered an outgroup. Indeed, it has been debated whether *Y. ruckeri* is truly part of the *Yersinia* genus[21]; however, the identification of the newly classified lineages *Yersinia nurmii* and *Yersinia entomophaga* (which together form species cluster 3) strengthens the case for including *Y. ruckeri* in the *Yersinia* genus.



Yersinia outer proteins (Yops). A set of effector proteins secreted by the Ysc type III secretion system (T3SS). Yops are injected into phagocytic cells, in which they inhibit the production of pro-inflammatory cytokines and induce apoptosis of the infected cell.

Multi-locus sequence typing (MLST). A strain typing technique based on the allele profiling of seven housekeeping genes conserved in a given species.

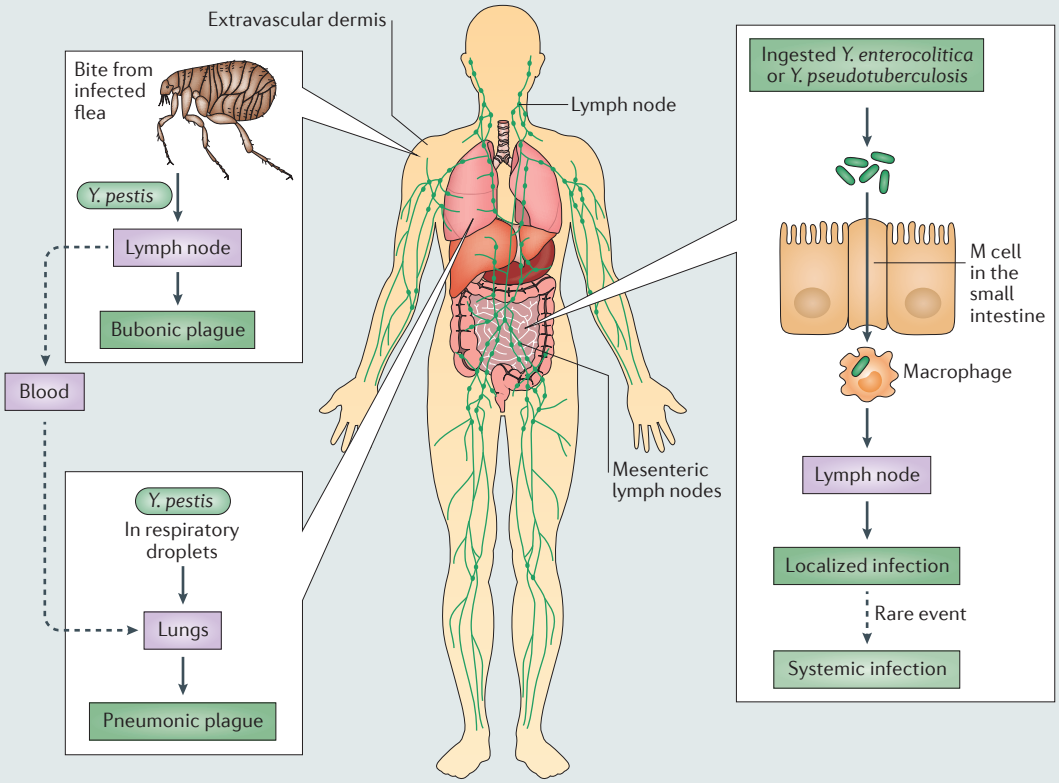
The advantage of sequencing broadly across a genus and deeply (that is, many isolates) within each species is that the origins and distribution of each gene described as a *Yersinia* spp. virulence factor, or indeed any *Yersinia* spp. gene, can be ascertained[8]. Such an analysis confirmed that pYV was one of only two virulence-associated genetic loci present in all three *Yersinia* spp. lineages that are commonly associated with

Box 2 | **Mechanisms of pathogenesis in *Yersinia* spp.**

The three species of *Yersinia* that are pathogenic in humans — *Yersinia pestis*, *Yersinia pseudotuberculosis* and *Yersinia enterocolitica* — share a large virulence plasmid, designated pYV (plasmid of *Yersinia* virulence), that is key to their ability to cause disease. Although commonly thought of as a single entity, pYV is highly variable between species and strains of species, containing different origins of replication and having variable genetic architecture[8,12]. However, the large genetic locus encoding the Ysc type III secretion system (T3SS) is present in pYV in all pathogenic strains. The Ysc T3SS, which was the first T3SS to be fully characterized, mediates the targeted delivery of *Yersinia* outer protein (Yop) effector proteins into host cells. Yops are the key virulence factors in all three pathogenic species; on contact with host macrophages, the injection of Yops results in the silencing of the pro-inflammatory cytokine response and in apoptotic death of the infected macrophages[40].

Y. enterocolitica and *Y. pseudotuberculosis* are zoonotic infectious pathogens that cause inflammatory intestinal disease following the ingestion of undercooked or contaminated pork products and vegetables[113]. The bacterial cells enter the small intestine, where they attach to the intestinal epithelium. Numerous virulence factors have been implicated in this process, including fimbriae, flagellal proteins and adhesins (YadA and attachment invasion locus protein (Ail))[18]. The invasin protein then triggers the internalization of the bacteria, which translocate across the epithelium[47]. Bacterial cells that are replicating in the intestinal tract can also disseminate to lymphatic tissues through infection and silencing of phagocytic immune cells[114]. As a result, this infection is often diagnosed as pseudoappendicitis and can lead to septicaemia[115] (see the figure).

Unlike *Y. enterocolitica* and *Y. pseudotuberculosis*, *Y. pestis* does not have an intestinal phase in its infection cycle. Instead, *Y. pestis* forms a biofilm in the foregut of the rodent flea vector[64]. When the flea next takes a blood meal, the biofilm is regurgitated into the host. If injected into the bloodstream, the bacterial cells immediately encounter phagocytes from the host immune system, triggering the activation of the Ysc T3SS and the silencing of host defence pathways. This leads to septicaemic plague[116], which is a rare but potentially lethal form of the disease. If injected into the extravascular part of the dermis, the bacterial cells travel to the lymph nodes, where an acute infection is established, leading to the formation of painful swellings at the lymph nodes, known as buboes[116]. This is known as bubonic plague (see the figure). If untreated, the infection will re-enter the bloodstream, enabling bacteria to travel to the lungs, where pneumonic plague is established (see the figure). *Y. pestis* has acquired additional plasmids, pPla and pFra, which have pivotal roles in the insect and systemic phases of infection, respectively[52,53,117].



M cell, microfold cell.

mammalian disease, but not in *Yersinia* species that are not mammalian pathogens. The second locus was the chromosomally encoded attachment and invasion locus (*ail*), which has roles in the attachment to, and invasion of, mammalian epithelial cells and in resistance to killing by the host complement system[43–45]. The pivotal role of pYV in the emergence of mammalian pathogenesis is highlighted by the revelation from genomics studies that closely related versions of pYV were acquired independently on at least three occasions in the *Yersinia* genus: twice in the *Y. enterocolitica* lineage and once in the *Y. pseudotuberculosis*–*Y. pestis* lineage[8]. This is contrary to the previously favoured hypothesis that all three pathogenic species were descended from a common ancestor that had already acquired this plasmid[9,27,46]. Genomics studies also showed that *Y. enterocolitica* does not share a common pathogenic ancestor with *Y. pseudotuberculosis* as previously thought[8]. Instead, several independent *ail* and pYV acquisitions were key gene gain events that marked the parallel independent evolution of mammalian pathogenesis in the *Yersinia* genus. This repeated gain of just two genetic loci is surprising if one considers the numerous virulence-associated genes that have been reported in *Yersinia* spp. using conventional bacterial genetics experiments. Some of these genes encode key virulence factors, such as invasin[47], the Fes/Fep iron acquisition system[48], the type IVb pilus encoded by the *tad* locus[49], Myf fimbriae[50] and the heat stable toxin Yst[51], that are widely distributed across pathogenic and non-pathogenic members of the genus. From the eco–evo perspective of pathogen evolution, this discrepancy could be explained by environmental *Yersinia* species acquiring and maintaining virulence factors to combat predatory protozoa, nematodes, bacteria and phages, or even to colonize plants. As such, the role of the accessory gene pool in pathogenic bacteria and pathogenesis may be better understood when studied in the wider context of non-pathogenic members of the genus, a strategy that has yet to be widely applied to other enteric pathogens. One benefit of studying the role of virulence genes in non-pathogenic isolates may be that, through an improved understanding of the ecology of these bacterial species, we will be better able to evaluate the hypothesis that pathogen evolution should be viewed from an eco–evo perspective.

In addition to the crucial gains of the virulence plasmid pYV and the *ail* locus in the lineages of all three human pathogenic *Yersinia* species, several other loci have been acquired and maintained by clonal expansion in individual *Yersinia* lineages. Perhaps the most formative of these is the acquisition by *Y. pestis* of the pFra (also known as pMT1) plasmid, which encodes Yersinia murine toxin (Ymt) and the F1 capsular protein, and the pPla (also known as pPCP1) plasmid, which encodes the plasminogen activator protein Pla. The acquisition of these two plasmids is a significant event as Ymt is required for the flea-borne transmission of *Y. pestis*[52] and Pla has recently been shown to be a key virulence factor responsible for the fulminant lung infection associated with *Y. pestis* but not with other *Yersinia* species[53]. Other acquired elements specific to individual lineages of the *Yersinia* genus include the high pathogenicity island (HPI), which is a large integrative and conjugative element (ICE)[54] that is found in several bacterial genera, including extraintestinal pathogenic *Escherichia coli* (ExPEC)[55]. The HPI encodes the siderophore yersiniabactin, which sequesters iron from the host[55,56], imports zinc into bacterial cells[57] and limits the respiratory oxidative burst response of immune cells[58]. In the *Yersinia* genus, HPI elements that contain a fully functioning yersiniabactin are only found in strains belonging to the *Y. pseudotuberculosis*–*Y. pestis* species complex, in which the presence of yersiniabactin is isolate-dependent, and in the highly virulent phylogroup 2 lineage of *Y. enterocolitica*, in which yersiniabactin is uniformly present. The key remaining gene gain events in the evolution of pathogenesis in the *Yersinia* genus are the acquisition of the locus encoding the Ysa T3SS by phylogroup 2 of *Y. enterocolitica* and the acquisition of the haemin storage (*hms*) locus by the *Y. pseudotuberculosis*–*Y. pestis* species complex. The locus encoding the Ysa T3SS is located in a region of the genome known as the plasticity zone[59]; first identified in the archetypal phylogroup 2 strain 8081 (REF. 9), the plasticity zone seems to have been the site of several independent and phylogroup-specific acquisitions and deletions[9]. The Ysa T3SS has been shown to secrete numerous chromosomally encoded T3SS effector proteins[60], as well as the Yop effector proteins of the Ysc T3SS that are encoded by the virulence plasmid pYV[61]. Although mutagenesis studies have suggested that the Ysa T3SS has only a minor role in mammalian pathogenesis[60], *ysa* deletion mutants were shown to be attenuated in their ability to survive inside cultured *Drosophila melanogaster* cells, which suggests that the Ysa T3SS acquired by phylogroup 2 of *Y. enterocolitica* is more important for survival in insect vectors than for mammalian pathogenesis[62]. This is consistent with the eco–evo perspective, as the virulence factor is present in a mammalian pathogen but acquired for survival in an environment outside of the mammalian host. The *hms* locus is located on the chromosome and encodes a haemin storage system that is unique to the *Y. pseudotuberculosis*–*Y. pestis* species complex and is responsible for the formation of pigmented colonies on Congo-red agar[63]. This locus is an important factor in *Y. pestis* pathogenesis as it is required for the formation of biofilms that block the flea foregut, which results in reflux of blood from the infected feeding flea back onto the bite site, thereby facilitating transmission to a mammalian host[64].

Regulatory control

In addition to their acquisition, the stable maintenance of mobile genetic elements has been a key process in the emergence of pathogenesis in the *Yersinia* genus. A newly acquired locus is more likely to be stably maintained in a population if only a minimal fitness cost is associated with integration into the genome of the host bacterium and/or the genetic determinant encoded by the locus is under a high selective pressure. To achieve this, many acquired elements fall under strict transcriptional and/or translational control to minimize the

Accessory gene pool
The set of genes that have been shown to be differentially present in individual genomes within a species or genus.

Yersinia murine toxin
(Ymt). First characterized as a determinant of lethality in mice but now known to have a crucial role in the ability of *Yersinia pestis* to survive in fleas.

F1 capsular protein
Protein antigen found on the surface of pathogenic *Yersinia* spp. that is thought to modulate the targeting of bacteria to sites of infection.

Pla
A protease found in *Yersinia pestis* that is encoded on a *Y. pestis*-specific plasmid. Pla is required for pneumonic infection.

Integrative and conjugative element
(ICE). A class of bacterial mobile elements that uses self-encoded integrase and excision factors for excision and transfer to a donor cell.

Extraintestinal pathogenic *Escherichia coli*
(ExPEC). Strains of *Escherichia coli* that can asymptomatically colonize the mammalian intestinal tract but cause disease in sites such as the urinary tract, blood stream or meninges.

Regulons
Sets of genes in which each set is under the control of a common regulatory system.

MarR/SlyA
A family of transcriptional regulators found in bacteria. Most of these regulators activate transcription by alleviating histone-like nucleoid structuring protein (H-NS)-mediated repression.

Promiscuous regulatory system
A set of promiscuous regulators and cognate regulon; these promiscuous regulators assume transcriptional control of numerous genes that do not come under fine-scale environmental control.

fitness costs associated with their acquisition. For example, many acquired genetic elements in the genomes of Enterobacteriaceae are regulated by the histone-like nucleoid structuring protein (H-NS), which binds to AT-rich regions of DNA to silence transcriptional activity[65]. Once under the strict negative control of H-NS, the acquired elements are integrated into existing regulons in which transcription is activated by a positive regulator that functions antagonistically with H-NS to tightly control the expression of the new genes. In the *Yersinia* genus, the regulator of virulence protein A (RovA), which is a member of the MarR/SlyA family of global transcriptional regulators[66], is present in all species[8] and has been suggested to fulfil the role of a positive regulator that antagonizes H-NS[8,66]. Indeed, RovA antagonizes H-NS silencing of the gene encoding invasin, which is involved in the early stages of *Yersinia* spp. attachment to cells and internalization during infection, by binding to the promoter region to mask the H-NS binding site, relieving transcriptional suppression[67]. Genes in the RovA regulon are silenced by H-NS and a co-repressor, YmoA; together, *rovA* and *ymoA* are thought to encode an ancestral promiscuous regulatory system with a regulon that has acquired distinct sets of genes in different lineages (FIG. 1). This promiscuity probably explains why RovA and YmoA regulate different genes in *Y. enterocolitica* and the *Y. pseudotuberculosis*–*Y. pestis* species complex[66,68]; indeed, almost all genes controlled by RovA and YmoA in *Yersinia* spp. are lineage-specific. These lineage specific genes include genes that encode virulence factors, as well as prophage genes and metabolic genes[66]. The RovA–YmoA regulon also includes *lcrF*, which encodes a transcriptional activator of virulence genes; this further highlights the importance of the RovA–YmoA regulon as a global regulator of virulence.

The regulation of gene expression by YmoA and LcrF responds to the mammalian environment, as the YmoA protein and the *lcrF* transcript are thermosensitive switches. The virulence factors regulated by LcrF, such as Yop proteins, are encoded on the virulence plasmid pYV.

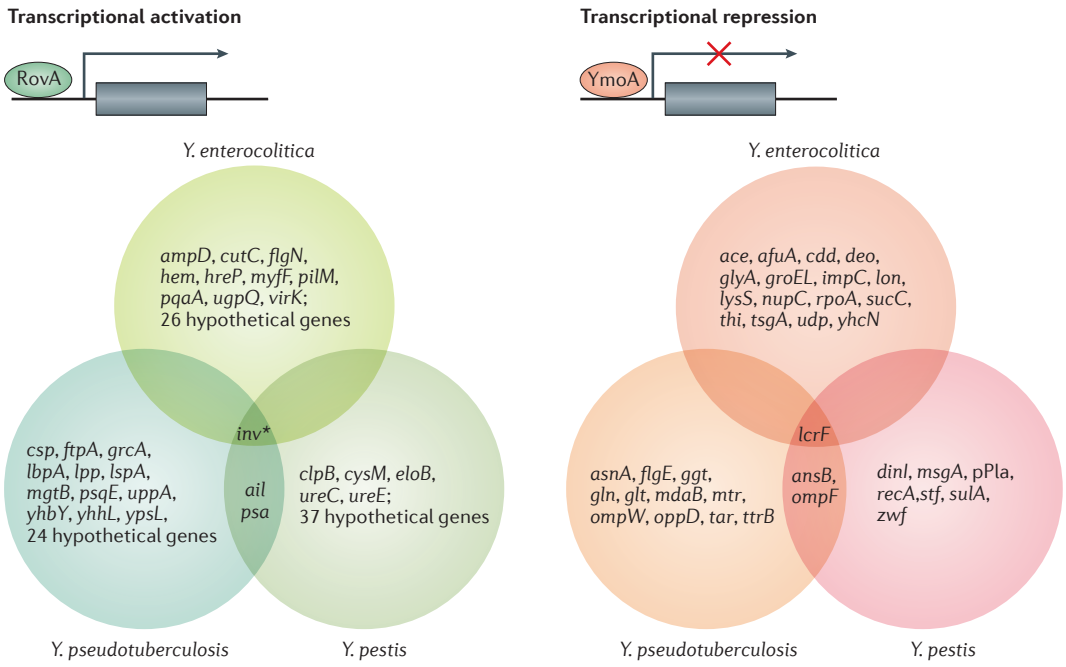


Figure 1 | **Comparison of the RovA–YmoA regulons of *Yersinia enterocolitica*, *Yersinia pseudotuberculosis* and *Yersinia pestis*.** An analysis of gene regulation by the regulator of virulence protein A (RovA) in *Yersinia enterocolitica*, *Yersinia pseudotuberculosis* and *Yersinia pestis* shows that the key acquired virulence plasmid pYV (represented by *lcrF*) is the only locus for which repression by RovA is conserved in all three species. The activation of *inv* (known as *invA* in *Y. enterocolitica*; the gene encodes an invasin protein that mediates host cell invasion) by RovA is also shown as conserved in all three species; however, the asterix indicates that *inv* is a pseudogene in *Y. pestis*. Aside from the conserved activation of the *psa* locus in *Y. pseudotuberculosis* and *Y. pestis* and *myfF* (the functional orthologue of *psaF*) in *Y. enterocolitica*, the RovA–YmoA regulons of the three species otherwise lack common activated and repressed genes, which highlights the key roles RovA and YmoA have in controlling the expression of lineage-specific loci. *ail*, attachment and invasion locus; *ampD*, *N*-acetyl-anhydromuramyl-L-alanine amidase; *ansA*, asparaginase; *cdd*, cytidine deaminase; *csp*, cold shock protein; *cysM*, cysteine synthase; *ggt*, γ-glutamyltranspeptidase; *gln*, glutamine ABC transporter locus; *glt*, sodium/glutamate symport carrier protein; *glyA*, serine hydroxymethyltransferase; *hreP*, subtilisin/kexin-like protease; *lpp*, major outer membrane lipoprotein; *lysS*, lysine-tRNA ligase; *mdaB*, modulator of drug activity B; *mgtB*, Mg^{2+} transport ATPase protein B; *msgA*, macrophage survival gene A; *mtr*, tryptophan-specific transport protein; *myfF*, mucoid *Yersinia* factor F; *ompF*, outer membrane protein F; *oppD*, oligopeptide ABC transporter ATP-binding protein; *psa*, pH 6 antigen; *rpoA*, DNA-directed RNA polymerase subunit α; *stf*, bacteriophage tail fibre protein; *sucC*, succinyl-CoA synthetase subunit β; *tar*, methyl-accepting chemotaxis protein; *ttrB*, tetrathionate reductase subunit B; *udp*, uridine phosphorylase; *ugpQ*, glycerophosphodiester phosphodiesterase; *ureC*, urease subunit α; *zwf*, glucose-6-phosphate 1-dehydrogenase.

Repression of these genes occurs through the silencing of *lcrF*, which is itself also located on the pYV plasmid, by YmoA and H-NS. The mechanism of the thermosensing switch at low temperatures is a tandem process that relies on the binding of YmoA directly to the promoter region of the *lcrF* locus and the formation of a double stem loop in an intergenic region in the *lcrF* transcript. The stem loop conceals the ribosome binding site in the transcript from the ribosome[69]. As a result of this combined transcriptional and post-transcriptional regulatory mechanism, *lcrF* transcription and translation are repressed at low temperatures[69]. RovA is also a thermosensitive switch that binds to promoters at low temperatures[70]. As a result of YmoA and LcrF thermosensitivity, the virulence factors that are encoded by pYV are not expressed at temperatures below 37 °C but are instead only expressed after a shift to mammalian body temperatures[69]. By contrast, the RovA regulon, which includes genes encoding virulence factors such as Ail and invasin, is activated by RovA at temperatures below 37 °C but not at mammalian body temperatures[70].

Overall, the importance of thermosensitive gene regulation by RovA and YmoA for the evolution of pathogenesis in the *Yersinia* genus is evident from studies describing their regulons. However, independent of temperature, the genes involved in the biosynthesis of flagella are expressed at high levels in *Y. enterocolitica* and *Y. pseudotuberculosis*, and insect transmission loci, such as *hms*, are expressed at high levels in *Y. pestis*[71,72], which suggests the involvement of additional regulatory factors. Aside from temperature, other environmental triggers known to influence the regulation of virulence factors in pathogenic *Yersinia* species include low levels of calcium[73] or magnesium[74] divalent cations. Depletion of Mg^{2+} is sensed by the PhoPQ two-component regulatory system[74], which regulates the level of RovA and, as such, is fundamental to virulence phenotypes in all three human pathogenic species of *Yersinia* through the control of the RovA regulon[75–77].

The presence of an ancestral promiscuous regulatory system that controls the expression of acquired virulence factors shows remarkable similarity to the regulation of the acquired *Salmonella* pathogenicity island 1 (SPI-1) by *Salmonella* spp. This pathogenicity island encodes a T3SS and its cognate secreted effector proteins, which remodel the actin cytoskeleton of infected cells to promote the uptake of attached bacteria into endocytic vacuoles[78]. As with the virulence plasmid pYV, the acquisition of SPI-1 was a landmark event in the evolution of the *Salmonella* genus, and it is present in both species that have been described in the genus, *S. enterica* and *Salmonella bongori*[78]. Furthermore, as with the virulence factors that are regulated by RovA and YmoA in *Yersnia* spp., SPI-1 is under negative transcriptional control by H-NS. The H-NS-mediated repression of SPI-1 is alleviated by PhoPQ and several transcriptional regulators, including the HilD and HilA regulatory cascade and the RovA orthologue SlyA[79]. Recent work in *S.* Typhimurium showed that SPI-1 was readily lost from the genomes of H-NS null mutants, suggesting that H-NS silencing of SPI-1 has been an essential step in the fixation of SPI-1 in all *Salmonella* species, and therefore for the evolution of pathogenesis in this genus[80].

The shared features of acquired virulence factor regulation in *Yersinia* spp. and *Salmonella* spp. suggests that an emerging theme in the evolution of pathogenesis in enteric bacteria is not only the acquisition of one or two key genetic loci, but also the coordinated transcriptional control of these acquired loci by H-NS and other promiscuous species-specific global regulators.

The role of gene loss

Although the importance of gene gain in the evolution of bacterial pathogenicity is perhaps more intuitive and has been better documented, gene loss has also been shown to have a central role in the emergence of bacterial pathogens, notably in host adaptation[81]. Most studies examining the evolution of pathogenesis in *Yersinia* spp. have focused on the evolutionary events that led to the emergence of *Y. pestis* from *Y. pseudotuberculosis*[28]. Although there were important gene gain events in this evolutionary pathway, the most remarkable observation of early comparative genomic analyses between the two species was the functional loss of approximately 10% of the genome in *Y. pestis*, including numerous genes associated with virulence and metabolism in *Y. psuedotuberculosis*[10]. This is a considerably larger proportion of the genome than has been acquired by gene gain events in the *Y. pestis* lineage. The high level of gene attrition was largely attributable to the expansion of four major classes of insertion sequence element (IS*1541*, IS*100*, IS*285* and IS*1661*), leading to gene inactivation at the sites of insertion. Other processes of functional gene loss include the deletion of genes or larger genetic loci and SNP-based pseudogene formation[32,33,82,83]. Many of the genes that were functionally lost encode virulence factors, such as invasin and YadA, metabolic loci, such as those responsible for dicarboxylic amino acid metabolism and uracil transport, or loci involved in motility[10,27] (FIG. 2). *Y. pseudotuberculosis* requires these functions for the successful colonization of the mammalian gastrointestinal tract, which is a niche that was no longer frequently inhabited by *Y. pestis* as it adapted to colonization of the midgut of fleas and, in mammals, to become a systemic pathogen that invades the lymphatic system[27].

The importance of gene loss in the evolution of *Y. pestis* has recently been further demonstrated in the identification of mutations associated with adaptation to flea-borne transmission. Using a combination of comparative genomics and conventional bacterial genetics, three key loss-of-function mutations were identified that had become fixed in the *Y. pestis* population, all of which increased cyclic di-GMP-dependent biofilm formation in the flea foregut[6] (FIG. 2). In each case, loss of function was caused by a SNP that replaced a functional gene with a pseudogene. Of the three mutated genes, two encoded EAL-domain phosphodiesterases (PDEs) and one encoded the RcsA component of the Rcs signal transduction system. PDEs repress the transcription of the *hms* operon by degrading the bacterial signalling molecule cyclic di-GMP, whereas RcsA forms a dimer

Two-component regulatory system
A bacterial sensor–kinase system (composed of an outer membrane sensor and a response regulator) that regulates gene expression in response to a specific environmental stimulus.

Insertion sequence element
An insertion sequence element is the simplest type of bacterial transposable element and encodes only the gene required for its own transposition, flanked by insertion sequence repeats.

Cyclic di-GMP
A secondary messenger molecule used in bacterial signal transduction to modulate gene expression in response to environmental perturbations.

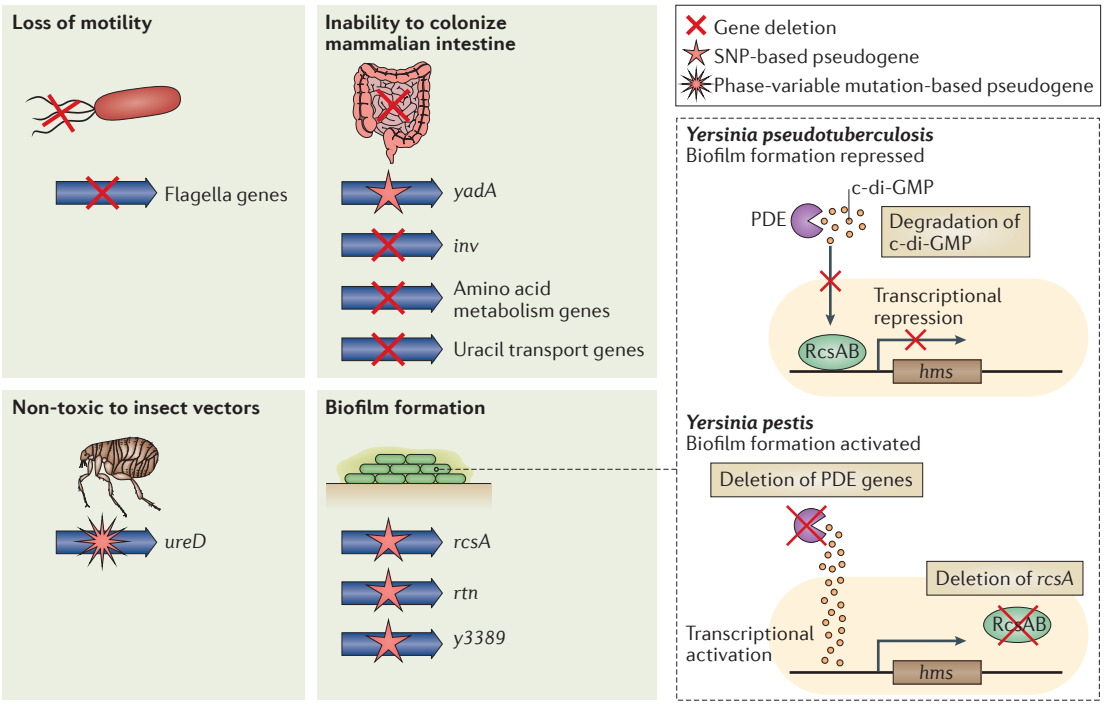


Figure 2 | **Gene loss in the emergence of *Yersinia pestis* from *Yersinia pseudotuberculosis*.** A comparison of the genome of *Yersinia pestis* with that of *Yersinia pseudotuberculosis* reveals deletion events in genes — encoding motility factors and components of metabolic pathways — that enable *Yersinia pseudotuberculosis* to colonize the mammalian intestinal tract. These deletions in the ancestral *Y. pseudotuberculosis* genome would have occurred as a result of a shift away from a niche in the mammalian gut environment towards an insect-vector phase of the life cycle during the emergence of *Y. pestis*. As a result, the functional loss of insect toxins was selected for; in *Y. pseudotuberculosis*, the function of urease, which has high toxicity in fleas, requires urease accessory protein (encoded by *ureD*) but in *Y. pestis* a phase-variable mutation has reversibly inactivated *ureD*. Furthermore, for transmission from the flea to the mammalian host to occur, *Y. pestis* must form biofilms in the flea foregut. Accordingly, three single-nucleotide polymorphisms (SNPs) have become fixed in *Y. pestis* that have inactivated three genes to pseudogenes to promote biofilm formation through the loss of gene function. First, the loss of function of *rcsA* alleviates transcriptional repression of the haemin storage (*hms*) operon, which encodes factors required for biofilm formation, by the RcsA–RcsB dimer. Second, SNPs have inactivated two genes (*rtn* and *y3389*) that encode phosphodiesterases (PDEs), which are enzymes that degrade cyclic di-GMP (c-di-GMP), a molecule that acts as a transcriptional activator of *hms* and thus biofilm formation. *inv*, invasin encoding gene.

with the response regulator RcsB to transcriptionally represses the *hms* operon through DNA binding[6]. As the *hms* operon promotes biofilm formation, the functional loss of the PDE enzymes and RcsA in the *Y. pestis* lineage led to an enhanced ability to form biofilms in the foregut of fleas, which enabled a shift in pathogenic lifestyle[6]. The observation that these SNP-based pseudogenes have become fixed in the *Y. pestis* population but that the genes have not yet been fully degraded[30,84] is consistent with the relatively recent emergence of *Y. pestis*[28] from a *Y. pseudotuberculosis* ancestor. A fourth loss-of-function mutation that has been identified in the *Y. pestis* lineage is a phase-variable mutation in the gene encoding the urease accessory protein (*ureD*). As this protein is required for urease function, the mutation also leads to the functional loss of urease. In *Y. pseudotuberculosis*, it has been shown that functional *ureD* is associated with significant oral toxicity in the flea[85]; therefore, the mutation in *ureD* may have contributed to the emergence of *Y. pestis*[6,85] (FIG. 2). However, as a phase-variable mutation, loss of function is reversible by spontaneous reversion. Therefore, *ureD* may be a contingency gene that provides *Y. pestis* with the flexibility of retaining urease expression as and when it is required, perhaps in the soil-dwelling stage of its life cycle[86].

Genomic analysis of *Y. enterocolitica* showed that, as with *Y. pestis*, gene loss and metabolic restriction also mark the evolution of certain subtypes of this species[8] (FIG. 3). In addition to the highly pathogenic phylogroup 2 described above, *Y. enterocolitica* can be subdivided into the non-pathogenic phylogroup 1, which is an ancestral lineage, and phylogroups 3–5, which show limited pathogenicity in mouse models but are the most successful lineages in terms of the prevalence of disease and are the most common isolates in livestock and human clinical samples[8]. Phylogroup 6 is a rare lineage that has only ever been isolated from wild hares[87,88]. Similarly to the emergence of *Y. pestis*, the divergence of the host-restricted phylogroup 6 from the broad host range lineage of phylogroups 3–5 was accompanied by the expansion of a single insertion sequence element, IS*1667*, which drove the functional loss of numerous metabolic genes. The functional loss of metabolic loci is particularly pronounced in phylogroup 6 and

Phase-variable mutation
A reversible mutation that can switch from wild-type to mutant, or from mutant to wild-type, in cell progeny.

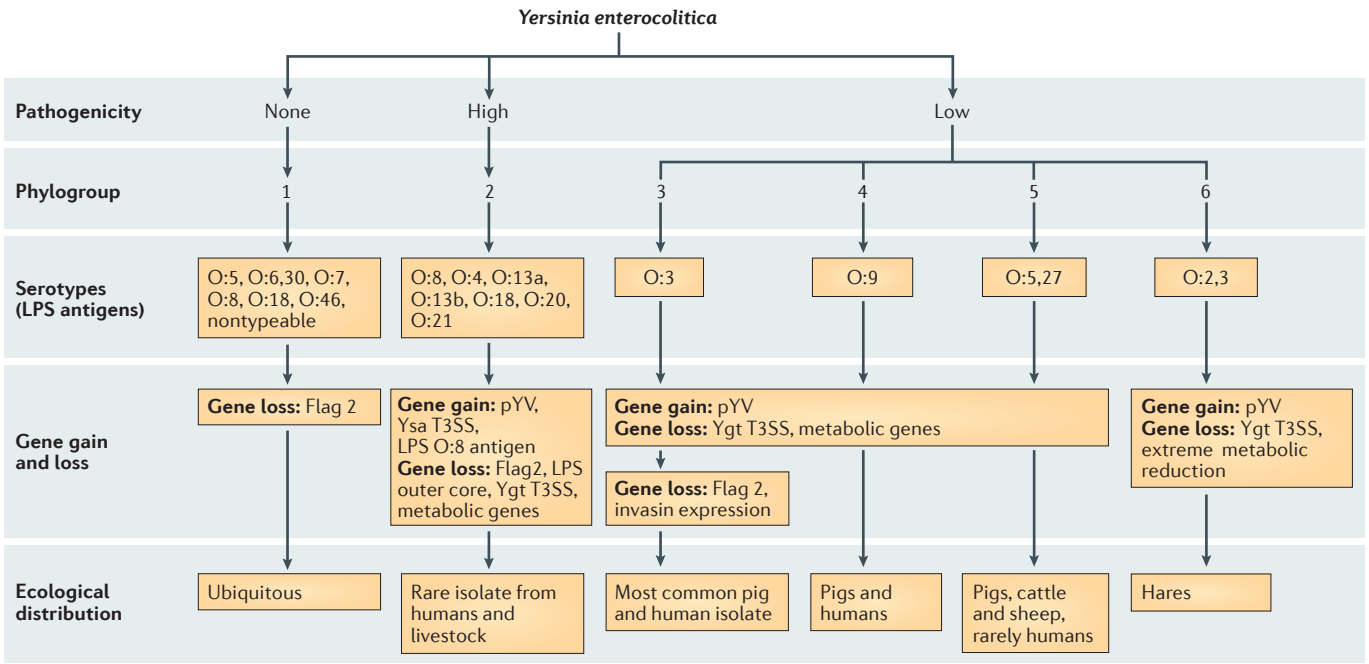


Figure 3 | **Gene gain and gene loss in *Yersinia enterocolitica*.** Isolates of *Yersinia enterocolitica* can be organized into phylogenetically distinct phylogroups. Although these groupings are classified according to their genomic relatedness they also reflect distinct serotypes (determined by lipopolysaccharide (LPS) structures), owing to the acquisition of specific LPS operons in each of the pathogenic phylogroups (such as the LPS O:8 antigen acquired by phylogroup 2, which is also present in *Yersinia pestis*). Phylogroup 1 is most similar to ancestral *Y. enterocolitica* and is non-pathogenic. Comparative genomic analyses suggest that the emergence of pathogenicity in this species was marked by the independent acquisition of the virulence plasmid pYV, which encodes the Ysc type III secretion system (T3SS), by the lineages of phylogroup 2 and phylogroups 3–5, and that the emergence of the enhanced pathogenicity of phylogroup 2 was the result of the acquisition of the high pathogenicity island (HPI). As opposed to the acquired Ysc and Ysa T3SSs, the Ygt T3SS is ancestral to the *Yersinia* genus; however, Ygt is lost from all the pathogenic phylogroups of *Y. enterocolitica*. Phylogroup 3 has become the dominant isolate found in pig reservoirs and cases of human disease; this phylogroup has a deletion in the Flag 2 locus, which encodes secondary flagellum genes, as well as an insertion in the promoter of the gene encoding invasin, which increases the transcription of this key virulence factor.

is thought to reflect its extreme niche restriction to hares. Interestingly, the metabolic pathways that are inactivated in phylogroup 6 overlap, in part, with the metabolic pathways lost in *Y. pestis*, as in both cases loss-of-function mutations have arisen in the operons for cobalamin biosynthesis, tetrathionate respiration and the pyruvate dehydrogenase complex, which is responsible for the use of molecular hydrogen in anaerobic respiration (see below). The metabolic predictions made from these genetic observations are directly supported by high-throughput metabolic phenotyping[8].

Of the functions that were lost as different lineages of *Y. enterocolitica* diverged and adapted to new niches, one of the earliest events is likely to have been the decay of the Ygt T3SS, which is a T3SS ancestral to the *Yersinia* genus and distinct from the Ysc and Ysa T3SSs acquired in pathogenic lineages. Ygt has not yet been fully characterized but it shares a high level of sequence identity with the T3SS encoded on the SPI-2 genomic island, which is required for intracellular survival of *S. enterica* cells[89]. In *Y. enterocolitica* phylogroup 1, which is non-pathogenic, the Ygt T3SS seems to include the full complement of effector proteins encoded by the SPI-2 locus[8], whereas in *Y. enterocolitica* phylogroups 3–6, which have low pathogenic potential, the Ygt T3SS seems to have been inactivated either by SNPs or gene deletions. Almost all genomic traces of the Ygt T3SS have been lost from isolates of *Y. enterocolitica* phylogroup 2, which are highly pathogenic[8], but these isolates have acquired the Ysa T3SS, which may offset the loss of Ygt, providing an example of how dynamic the evolution of pathogenesis can be through gene gain and gene loss[59]. A comprehensive functional characterization of the Ygt T3SS is required to determine whether the selective maintenance of the Ygt T3SS in *Y. enterocolitica* phylogroup 1 is a classic example of the eco–evo principle, in which there is selection for a protective benefit in a non-mammalian host environment.

Gene inactivation mediated by insertion sequence elements is evident in *Y. enterocolitica* phylogroup 3, which comprises strains of low pathogenicity traditionally classified as serotype O:3 strains[90,91]. However, the insertion of these elements does not always result in gene inactivation, as molecular analysis showed that insertion of an IS*1667* element into the promoter of *invA* created an additional RovA binding site that increased expression of the gene[92]. Comparative analysis between the other low-pathogenicity phylogroups of *Y. enterocolitica* suggests that this insertion is the primary event that led to the enhanced virulence observed in isolates from

Galliforme birds
An order of approximately 290 bird species that are best described as fowl-like or poultry-like.

phylogroup 3 compared with isolates from phylogroup 4 and phylogroup 5 (REF. 93). The upregulation of invasin enables isolates from phylogroup 3 to colonize porcine tissues more effectively than other phylogroups and, as a result, isolates from this phylogroup are the predominant isolates identified in human clinical samples in Europe during the past 10 years and in contemporary veterinary epidemiological surveys[93], reflecting the origin of most human infections in the consumption of pork products.

Gene loss in other Enterobacteriaceae

Functional gene loss in metabolic loci has been a key event in the transition of *Yersinia* species from a broad host range to a niche-restricted and host-adapted pathogen on at least two occasions: in *Y. enterocolitica* and in the *Y. pseudotuberculosis*–*Y. pestis* species complex. As with the acquisition of virulence genes and their incorporation into regulatory systems, the loss of metabolic loci has also been observed in the emergence of pathogenesis in *Salmonella* serovars. In particular, *S.* Typhi[94] has many similarities with *Y. pestis,* as it is niche-restricted (to the human host) and causes an acute systemic disease known as typhoid, as opposed to *S.* Typhimurium, which has a broad host range and causes an inflammatory infection of the intestine in mammals[95] (and is therefore more akin to *Y. pseudotuberculosis* and *Y. enterocolitica* than to *Y. pestis*). As with *S.* Typhi, isolates from the serovars *Salmonella enterica* subsp. *enterica* serovar Gallinarum and *Salmonella enterica* subsp. *enterica* serovar Pullorum cause a typhoid-like disease, but these infections are restricted to galliforme birds[96]. Comparative genomic analysis of these four serovars has shown that functional gene loss in loci involved in anaerobic metabolism — more specifically, in the *cob*, *pdu* and *ttr* operons — is a marker of the transition from intestinal pathogens (such as *S.* Typhimurium) to invasive pathogens (such as *S.* Typhi, *S.* Gallinarum and *S.* Pullorum)[97].

In *S.* Typhimurium, which maintains the *cob*, *pdu* and *ttr* operons, the *cob* operon encodes factors that are required for the synthesis of cobalamin (also known as vitamin B_{12}), which is activated under anaerobic conditions. Cobalamin is required as a cofactor for several enzymes in the 1,2-propanediol utilization pathway, such as propanediol dehydratase, which is encoded by *pduCDE*[98]. 1,2-Propanediol is a byproduct of fermentative growth on rhamnose and fucose[98], which are common constituents of plant cell walls and intestinal epithelial cells that line the gut[99]. In addition to cobalamin and *pdu*-encoded enzymes, the anaerobic degradation of 1,2-propanediol requires tetrathionate as a terminal electron acceptor, and the products of the *ttr* genes reduce tetrathionate for this purpose[100]. Unlike 1,2-propanediol, tetrathionate had not historically been described as present in the mammalian gut, but was instead known to be present in soils. However, in recent years it has been shown that tetrathionate is produced naturally during the inflammatory response to infection with *S.* Typhimurium. Under such conditions, the functions encoded by the *cob*, *pdu* and *ttr* operons combine to provide a competitive growth advantage for *S.* Typhimurium, as they enable it to outgrow the natural and largely fermentative gut microbiota by metabolizing 1,2-propanediol, which is not readily fermented[101,102] (FIG. 4). As the invasive typhoidal *Salmonella* serovars have evolved away from an intestinal lifestyle, functions in the *cob*, *pdu* and *ttr* operons have been sequentially lost. The loss of these operons represents a signature of the change in niche and a movement away from causing intestinal disease to a more systemic infection cycle[103].

The distribution of the *cob*, *pdu* and *ttr* loci in the *Yersinia* genus is similar to the distribution of these loci in the *Salmonella* genus. The operons are present and intact in *Y. enterocolitica* and in the majority of environmental *Yersinia* species[104], but absent from the *Y. pseudotuberculosis*–*Y. pestis* species complex. As mentioned above, tetrathionate is produced in the vertebrate gut during the inflammatory response to infection with *S.* Typhimurium. Similarly, studies on *Y. enterocolitica* pathogenesis have shown that infections with this pathogen are characterized by inflammation of the gut[105]. Moreover, the acquisition of virulence factors seems to be essential for the stimulation of inflammation by both *S.* Typhimurium and *Y. enterocolitica*, with a requirement for the SPI-1- and SPI-2-encoded T3SSs in *S.* Typhimurium infection and pYV in *Y. enterocolitica* infection[106]. It is possible that, similarly to *S.* Typhimurium, the induction of a pYV-associated inflammatory response during intestinal infection[107], combined with the expression of the *cob*, *pdu* and *ttr* operons, enables *Y. enterocolitica* to gain a metabolic advantage

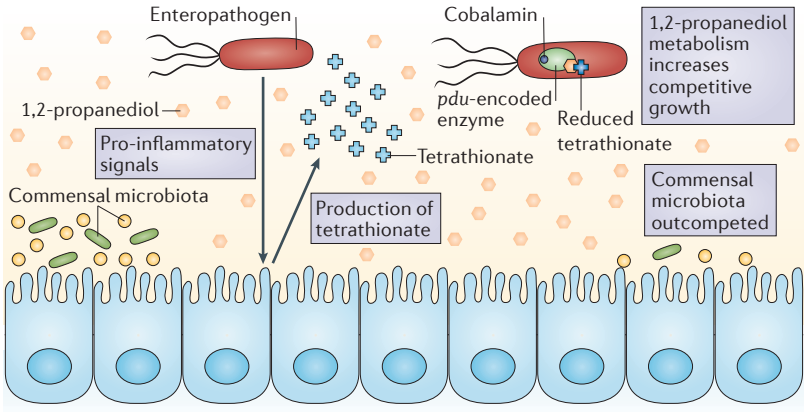


Figure 4 | **Enteropathogenic Enterobacteriaceae use the *cob*, *pdu* and *ttr* operons to outcompete the normal intestinal microbiota during mammalian infection.** The *cob*, *pdu* and *ttr* operons are maintained specifically in the enteropathogenic *Salmonella* serovars (such as *Salmonella enterica* subsp. *enterica* serovar Typhimurium) and are also present in *Yersinia enterocolitica*. These enteropathogens cause inflammation during infection, which leads to the overproduction of tetrathionate by the intestinal epithelium. In this environment, the *cob*, *pdu* and *ttr* operons together enable *Y. enterocolitica* and *S.* Typhimurium to metabolize 1,2-propanediol (a carbon source present in the gut) and thus provide a metabolic advantage to these species when competing with the normal intestinal microbiota in the inflamed gut. The *cob* operon encodes factors that are required for cobalamin (also known as vitamin B_{12}) biosynthesis. Cobalamin activates enzymes that degrade 1,2-propanediol, and these enzymes are encoded by the *pdu* operon. The degradation of 1,2-propanediol also requires the use of reduced tetrathionate as a terminal electron acceptor during anaerobic respiration, and the reduction of tetrathionate requires factors encoded by the *ttr* operon. Metabolism of 1,2-propanediol by enteropathogenic Enterobacteriaceae enables them to outcompete the normal gut microbiota and leads to overgrowth and colonization.

Phenotype microarrays
High-throughput, automated assays that determine the ability of living cells to metabolize metabolic substrates.

over the resident gut microbiota by using tetrathionate and naturally occurring carbon sources to respire under these conditions. The complete loss of the *cob*, *pdu* and *ttr* loci in *Y. pseudotuberculosis* and *Y. pestis* may explain why these pathogens are incapable of causing the inflammatory infection of the mammalian intestine observed in *Y. enterocolitica,* despite their retention of pYV.

Summary and concluding remarks

The study of pathogenic *Yersinia* species has historically been fundamental to our understanding of the mechanisms of bacterial pathogenesis, such as the role of T3SSs[40,108], the molecular mechanisms by which pathogens invade the intestinal epithelium[43] and the importance of plasmids in conferring virulence[12,39]. As we have described in this Review, the study of the *Yersinia* genus has also now — in the genomics era — become fundamental to our understanding of the evolution of mammalian pathogenesis, as it is the first multispecies bacterial genus for which all species have been sequenced and placed into phylogenomic context[8]. By combining a genomics approach with high-throughput metabolic analyses using phenotype microarrays, it has been possible to trace the evolutionary origins of each species, compare the pathogenic and non-pathogenic species and reveal the parallel evolutionary paths of the emergence of virulence in humans[8] (FIG. 5). This combined approach has established that the *Yersinia* genus has evolved from environmental ancestors and that the human pathogens *Y. enterocolitica* and the *Y. pseudotuberculosis*–*Y. pestis*

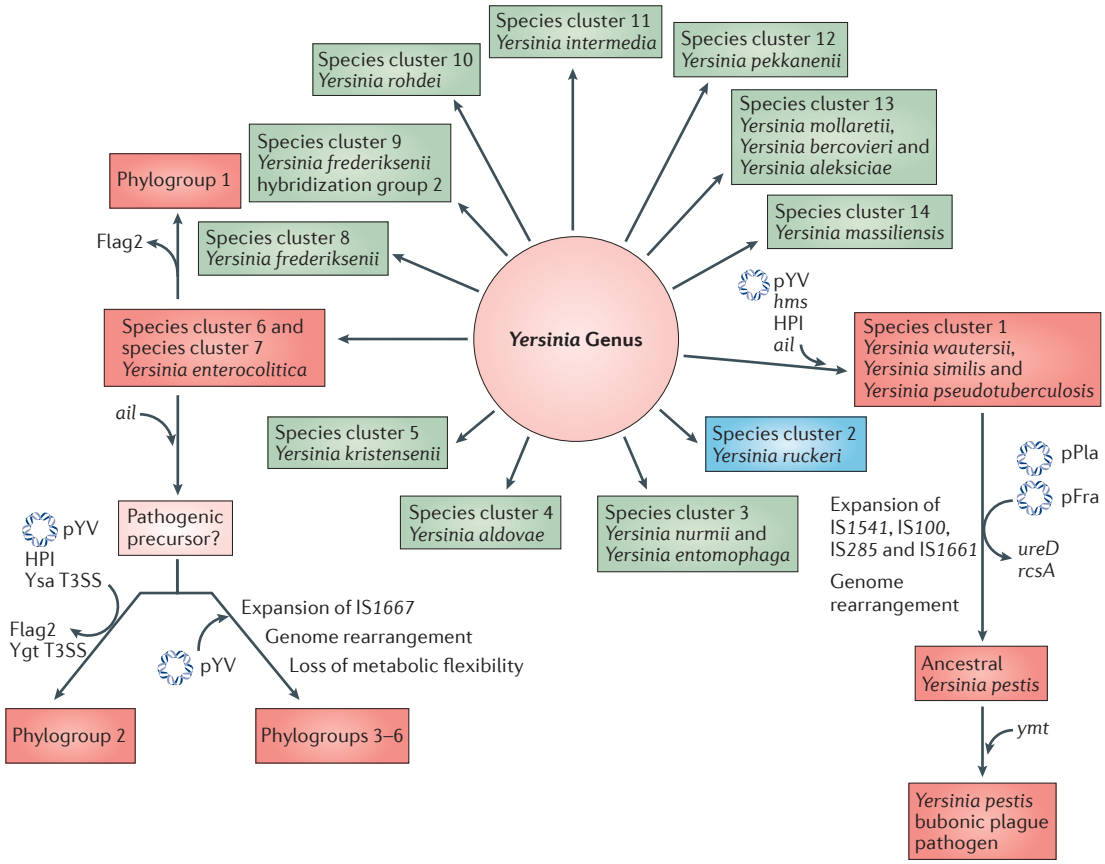


Figure 5 | **Key gene gain and gene loss events in the evolution of mammalian pathogenesis in the *Yersinia* genus.** The formation of each pathogenic lineage in the *Yersinia* genus is marked by key gene gain and gene loss events. The three mammalian pathogens of the genus (*Yersinia enterocolitica*, *Yersinia pseudotuberculosis* and *Yersinia pestis*) are present in two distinct lineages. In the first, *Y. enterocolitica* diverged into a non-pathogenic phylogroup (phylogroup 1) and a precursor of the pathogenic phylogroups 2–6. In the second, an ancestor of species cluster 1 diverged into the non-pathogenic species *Yersinia similis* and *Yersinia wautersii* and the pathogen *Y. pseudotuberculosis*. In turn, the plague pathogen *Y. pestis* emerged from *Y. pseudotuberculosis* — following the acquisition of *Yersinia* murine toxin (*ymt*) by the pFra plasmid — which ultimately gave rise to flea-borne transmission and thereby the ability of modern day *Y. pestis* to cause bubonic plague. The most prominent of the gene gain events during the emergence of pathogenesis in the two pathogenic lineages is the acquisition of the pYV virulence plasmid, which encodes the Ysc type III secretion system (T3SS). pYV was acquired independently in each lineage, including twice in the *Y. enterocolitica* lineage, which shows that human pathogenicity has emerged on at least three occasions from environmental generalists; in each case, a similar theme of gene gain, gene loss, metabolic reduction and genome rearrangements has driven this process. For example, the acquired chromosomally encoded attachment and invasion locus (*ail*) is present in all three human pathogens, whereas the gene encoding the urease accessory protein (*ureD*) was lost in *Y. pestis* to enable its survival in an insect vector. Flag 2, secondary flagellum gene cluster; HPI, high pathogenicity island; *hms*, haemin storage locus; T3SS, type III secretion system.

Palaeomicrobiology
The study of historical infectious disease or microbiota by recovering DNA from ancient human (or other host) remains.

species complex emerged independently following the common themes of 'add, stir and reduce'. In *Y. pestis*, gene gain was driven by the acquisition of the lineage-specific virulence plasmids pFra and pPla, whereas considerable loss of gene function occurred as a result of the expansion of several insertion sequence elements, which led to widespread genome rearrangements, and other mutational processes. However, in the *Y. enterocolitica* lineages, these themes of 'add, stir and reduce' are weighted differently in the various phylogroups. Phylogroup 2 acquired the HPI and the Ysa T3SS, and lost the secondary flagellum and the Ygt T3SS. In phylogroups 3–6, a general loss of metabolic flexibility arose from genome rearrangements driven by the expansion of IS*1667*; in the case of phylogroup 6, the loss of metabolic capability was extreme.

Genomic studies of *Yersinia* spp. have also been notable in the field of palaeomicrobiology, with the near complete reconstruction of the genomes of the *Y. pestis* strains responsible for the Black Death and Justinian plague from ancient DNA samples, the interpretation of which has shed light on historical episodes of disease[84,109]. These palaeomicrobiology studies can also provide new data that better resolve the microevolution of pathogenesis within the *Y. pestis* lineage. Indeed, this microevolution is beginning to be unravelled, with studies elucidating the relationship between ancestral and modern *Y. pestis* strains[110,111]. One very recent study has shown that *Y. pestis* was a pathogen of Bronze Age humans but had not yet acquired the Ymt toxin, which is required for the survival of the bacterium in the flea vector[111]. This interesting finding means that *Y. pestis* was already a human pathogen when it evolved the ability to infect fleas.

Genomics studies have also helped to evaluate the validity of the eco–evo perspective in the emergence of bacterial pathogenesis. When applied to comparative analysis of virulence factor genes in complete genome sequence datasets of all *Yersinia* species, the eco–evo perspective — in which organisms are evaluated broadly in the light of evolution and ecology, rather than the narrow constraints of their behaviour in human infection — has informed our understanding of the key evolutionary steps that led to the emergence of mammalian pathogenesis. This suggests that it will be profitable to move away from a human-centric view of bacterial pathogenesis in favour of a greater consideration of the environmental context, habitat and niches of these organisms.

Finally, genomics data from the *Yersinia* genus shows that a genus-wide sequencing approach to the study of bacterial pathogenesis leads to new information on the distribution of virulence-associated genes and on which of these genes are the most appropriate markers for novel diagnostics or the most promising targets for therapeutics. However, even with the numerous *Yersinia* spp. isolates sequenced to date, the origin and reservoir of mobile genetic elements, such as the virulence plasmid pYV, remain obscure. It may be that we have only sequenced the tip of the bacterial iceberg and that future large-scale *Yersinia* spp. genomics and microbiome studies, especially when coupled with high-throughput phenotyping methods, will provide more insights into the evolutionary histories of *Yersinia* spp. genomes and mobile elements. The web of bacterial life revealed by these studies will probably show an interrelatedness of bacteria that blurs species boundaries. Furthermore, when viewed from the eco–evo perspective, new genomic data promises to provide greater insights into how and why bacteria evolve to be more pathogenic, and thus has the potential to avert future threats posed by the emergence of new pathogenic strains.

1. Fuchs, T., Bresolin, G., Marcinowski, L., Schachtner, J. & Scherer, S. Insecticidal genes of *Yersinia* spp.: taxonomical distribution, contribution to toxicity towards *Manduca sexta* and *Galleria mellonella*, and evolution. *BMC Microbiol.* **8**, 214 (2008).
2. Lithgow, K. V. *et al.* A general protein *O*-glycosylation system within the *Burkholderia cepacia* complex is involved in motility and virulence. *Mol. Microbiol.* **92**, 116–137 (2014).
3. Pukklay, P. *et al.* Involvement of EnvZ–OmpR two-component system in virulence control of *Escherichia coli* in *Drosophila melanogaster*. *Biochem. Biophys. Res. Commun.* **438**, 306–311 (2013).
4. Kumar, P. & Virdi, J. S. Identification and distribution of putative virulence genes in clinical strains of *Yersinia enterocolitica* biovar 1A by suppression subtractive hybridization. *J. Appl. Microbiol.* **113**, 1263–1272 (2012).
5. Korea, C. G., Badouraly, R., Prevost, M.-C., Ghigo, J.-M. & Beloin, C. *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. *Environ. Microbiol.* **12**, 1957–1977 (2010).
6. Sun, Y. C., Jarrett, C. O., Bosio, C. F. & Hinnebusch, B. J. Retracing the evolutionary path that led to flea-borne transmission of *Yersinia pestis*. *Cell Host Microbe* **15**, 578–586 (2014).
The identification of key gene gain and loss events in the emergence of *Y. pestis*.
7. Viana, D. *et al.* A single natural nucleotide mutation alters bacterial pathogen host tropism. *Nat. Genet.* **47**, 361–366 (2015).
8. Reuter, S. *et al.* Parallel independent evolution of pathogenicity within the genus *Yersinia*. *Proc. Natl Acad. Sci. USA* **111**, 6768–6773 (2014).
The first population genomic analysis of an entire bacterial genus, and a key advance in the characterization of *Yersinia* spp. evolution.
9. Thomson, N. R. *et al.* The complete genome sequence and comparative genome analysis of the high pathogenicity *Yersinia enterocolitica* strain 8081. *PLoS Genet.* **2**, e206 (2006).
10. Chain, P. S. G. *et al.* Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc. Natl Acad. Sci. USA* **101**, 13826–13831 (2004).
11. Isberg, R. R., Voorhis, D. L. & Falkow, S. Identification of invasin: a protein that allows enteric bacteria to penetrate cultured mammalian cells. *Cell* **50**, 769–778 (1987).
12. Portnoy, D. A. & Falkow, S. Virulence-associated plasmids from *Yersinia enterocolitica* and *Yersinia pestis*. *J. Bacteriol.* **148**, 877–883 (1981).
13. Portnoy, D. A., Moseley, S. L. & Falkow, S. Characterization of plasmids and plasmid-associated determinants of *Yersinia enterocolitica* pathogenesis. *Infect. Immun.* **31**, 775–782 (1981).
14. Rosqvist, R., Hakansson, S., Forsberg, A. & Wolf-Watz, H. Functional conservation of the secretion and translocation machinery for virulence proteins of yersiniae, salmonellae and shigellae. *EMBO J.* **14**, 4187–4195 (1995).
15. Yao, T., Mecsas, J., Healy, J. I., Falkow, S. & Chien, Y. Suppression of T and B lymphocyte activation by a *Yersinia pseudotuberculosis* virulence factor, YopH. *J. Exp. Med.* **190**, 1343–1350 (1999).
16. Monack, D. M., Mecsas, J., Bouley, D. & Falkow, S. *Yersinia*-induced apoptosis *in vivo* aids in the establishment of a systemic infection of mice. *J. Exp. Med.* **188**, 2127–2137 (1998).
17. Boland, A. & Cornelis, G. R. Role of YopP in suppression of tumor necrosis factor alpha release by macrophages during *Yersinia* infection. *Infect. Immun.* **66**, 1878–1884 (1998).
18. Carniel, E. Evolution of pathogenic *Yersinia*, some lights in the dark. *Adv. Exp. Med. Biol.* **529**, 3–12 (2003).
19. Sprague, L. D. & Neubauer, H. *Yersinia aleksiciae* sp. nov. *Int. J. Syst. Evol. Microbiol.* **55**, 831–835 (2005).
20. Sprague, L. D., Scholz, H. C., Amann, S., Busse, H. J. & Neubauer, H. *Yersinia similis* sp. nov. *Int. J. Syst. Evol. Microbiol.* **58**, 952–958 (2008).
21. Sulakvelidze, A. Yersiniae other than *Y. enterocolitica*, *Y. pseudotuberculosis*, and *Y. pestis*: the ignored species. *Microbes Infect.* **2**, 497–513 (2000).
22. Murros-Kontiainen, A. *et al. Yersinia nurmii* sp. nov. *Int. J. Syst. Evol. Microbiol.* **61**, 2368–2372 (2011).
23. Murros-Kontiainen, A. *et al. Yersinia pekkanenii* sp. nov. *Int. J. Syst. Evol. Microbiol.* **61**, 2363–2367 (2011).
24. Hurst, M. R. H., Becher, S. A., Young, S. D., Nelson, T. L. & Glare, T. R. *Yersinia entomophaga* sp. nov., isolated from the New Zealand grass grub *Costelytra zealandica*. *Int. J. Syst. Evol. Microbiol.* **61**, 844–849 (2011).
25. Merhej, V., Adekambi, T., Pagnier, I., Raoult, D. & Drancourt, M. *Yersinia massiliensis* sp. nov., isolated from fresh water. *Int. J. Syst. Evol. Microbiol.* **58**, 779–784 (2008).
26. Savin, C. *et al.* The *Yersinia pseudotuberculosis* complex: characterization and delineation of a new species, *Yersinia wautersii*. *Int. J. Med. Microbiol.* **304**, 452–463 (2014).
27. Wren, B. W. The Yersiniae — a model genus to study the rapid evolution of bacterial pathogens. *Nat. Rev. Microbiol.* **1**, 55–64 (2003).

28. Achtman, M. *et al. Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc. Natl Acad. Sci. USA* **96**, 14043–14048 (1999).
The first paper to show that *Y. pestis* is a recently emerged clone of *Y. pseudotuberculosis*.
29. Achtman, M. *et al.* Microevolution and history of the plague bacillus, *Yersinia pestis*. *Proc. Natl Acad. Sci. USA* **101**, 17837–17842 (2004).
30. Bos, K. I. *et al. Yersinia pestis*: new evidence for an old infection. *PLoS ONE* **7**, e49803 (2012).
31. Chain, P. S. *et al.* Complete genome sequence of *Yersinia pestis* strains Antiqua and Nepal516: evidence of gene reduction in an emerging pathogen. *J. Bacteriol.* **188**, 4453–4463 (2006).
32. Parkhill, J. *et al.* Genome sequence of *Yersinia pestis,* the causative agent of plague. *Nature* **413**, 523–527 (2001).
33. Deng, W. *et al.* Genome sequence of *Yersinia pestis* KIM. *J. Bacteriol.* **184**, 4601–4611 (2002).
34. Savin, C. *et al.* Draft genome sequence of a clinical strain of *Yersinia enterocolitica* (IP10393) of bioserotype 4/O:3 from France. *Genome Announc.* **1**, e00150-12 (2013).
35. Batzilla, J., Hoper, D., Antonenka, U., Heesemann, J. & Rakin, A. Complete genome sequence of *Yersinia enterocolitica* subsp. *palearctica* serogroup O:3. *J. Bacteriol.* **193**, 2067 (2011).
36. Fuchs, T. M., Brandt, K., Starke, M. & Rattei, T. Shotgun sequencing of *Yersinia enterocolitica* strain W22703 (biotype 2, serotype O:9): genomic evidence for oscillation between invertebrates and mammals. *BMC Genomics* **12**, 168 (2011).
37. Wang, X. *et al.* Complete genome sequence of a *Yersinia enterocolitica* 'Old World' (3/O:9) strain and comparison with the 'New World' (1B/O:8) strain. *J. Clin. Microbiol.* **49**, 1251 (2011).
38. Chen, P. *et al.* Genomic characterization of the *Yersinia* genus. *Genome Biol.* **11**, R1 (2010).
39. Portnoy, D. A., Wolf-Watz, H., Bolin, I., Beeder, A. B. & Falkow, S. Characterization of common virulence plasmids in *Yersinia* species and their role in the expression of outer membrane proteins. *Infect. Immun.* **43**, 108–114 (1984).
40. Cornelis, G. R. & Wolf-Watz, H. The *Yersinia* Yop virulon: a bacterial system for subverting eukaryotic cells. *Mol. Microbiol.* **23**, 861–867 (1997).
41. Grosdent, N., Maridonneau-Parini, I., Sory, M.-P. & Cornelis, G. R. Role of Yops and adhesins in resistance of *Yersinia enterocolitica* to phagocytosis. *Infect. Immun.* **70**, 4165–4176 (2002).
42. Bohn, E. *et al.* Gene expression patterns of epithelial cells modulated by pathogenicity factors of *Yersinia enterocolitica*. *Cell. Microbiol.* **6**, 129–141 (2004).
43. Miller, V. L. & Falkow, S. Evidence for two genetic loci in *Yersinia enterocolitica* that can promote invasion of epithelial cells. *Infect. Immun.* **56**, 1242–1248 (1988).
44. Pierson, D. E. & Falkow, S. The *ail* gene of *Yersinia enterocolitica* has a role in the ability of the organism to survive serum killing. *Infect. Immun.* **61**, 1846–1852 (1993).
45. Biedzka-Sarek, M., Venho, R. & Skurnik, M. Role of YadA, Ail, and lipopolysaccharide in serum resistance of *Yersinia enterocolitica* serotype O:3. *Infect. Immun.* **73**, 2232–2244 (2005).
46. Carniel, E. Plasmids and pathogenicity islands of *Yersinia*. *Curr. Top. Microbiol. Immunol.* **264**, 89–108 (2002).
47. Pepe, J. C. & Miller, V. L. *Yersinia enterocolitica* invasin: A primary role in the initiation of infection. *Proc. Natl Acad. Sci. USA* **90**, 6473–6477 (1993).
48. Schubert, S., Fischer, D. & Heesemann, J. Ferric enterochelin transport in *Yersinia enterocolitica*: molecular and evolutionary aspects. *J. Bacteriol.* **181**, 6387–6395 (1999).
49. Tomich, M., Planet, P. J. & Figurski, D. H. The *tad* locus: postcards from the widespread colonization island. *Nat. Rev. Microbiol.* **5**, 363–375 (2007).
50. Iriarte, M. *et al.* The Myf fibrillae of *Yersinia enterocolitica*. *Mol. Microbiol.* **9**, 507–520 (1993).
51. Delor, I. & Cornelis, G. R. Role of *Yersinia enterocolitica* Yst toxin in experimental infection of young rabbits. *Infect. Immun.* **60**, 4269–4277 (1992).
52. Hinnebusch, B. J. *et al.* Role of *Yersinia* murine toxin in survival of *Yersinia pestis* in the midgut of the flea vector. *Science* **296**, 733–735 (2002).
53. Zimbler, D. L., Schroeder, J. A., Eddy, J. L. & Lathem, W. W. Early emergence of *Yersinia pestis* as a severe respiratory pathogen. *Nat. Commun.* **6**, 7487 (2015).
A recent publication showing the key role of Pla in the evolution of the unique virulence characteristics of *Y. pestis*.
54. Lesic, B. *et al.* Excision of the high-pathogenicity island of *Yersinia pseudotuberculosis* requires the combined actions of its cognate integrase and Hef, a new recombination directionality factor. *Mol. Microbiol.* **52**, 1337–1348 (2004).
55. Schubert, S., Picard, B., Gouriou, S., Heesemann, J. & Denamur, E. *Yersinia* high-pathogenicity island contributes to virulence in *Escherichia coli* causing extraintestinal infections. *Infect. Immun.* **70**, 5335–5337 (2002).
56. Carniel, E., Guilvout, I. & Prentice, M. Characterization of a large chromosomal 'high-pathogenicity island' in biotype 1B *Yersinia enterocolitica*. *J. Bacteriol.* **178**, 6743–6751 (1996).
57. Bobrov, A. G. *et al.* The *Yersinia pestis* siderophore, yersiniabactin, and the ZnuABC system both contribute to zinc acquisition and the development of lethal septicaemic plague in mice. *Mol. Microbiol.* **93**, 759–775 (2014).
58. Paauw, A., Leverstein-van Hall, M. A., van Kessel, K. P. M., Verhoef, J. & Fluit, A. C. Yersiniabactin reduces the respiratory oxidative stress response of innate immune cells. *PLoS ONE* **4**, e8240 (2009).
59. Haller, J. C., Carlson, S., Pederson, K. J. & Pierson, D. E. A chromosomally encoded type III secretion pathway in *Yersinia enterocolitica* is important in virulence. *Mol. Microbiol.* **36**, 1436–1446 (2000).
60. Matsumoto, H. & Young, G. M. Proteomic and functional analysis of the suite of Ysp proteins exported by the Ysa type III secretion system of *Yersinia enterocolitica* Biovar 1B. *Mol. Microbiol.* **59**, 689–706 (2006).
61. Young, B. M. & Young, G. M. Evidence for targeting of Yop effectors by the chromosomally encoded Ysa type III secretion system of *Yersinia enterocolitica*. *J. Bacteriol.* **184**, 5563–5571 (2002).
62. Walker, K. A., Maltez, V. I., Hall, J. D., Vitko, N. P. & Miller, V. L. A phenotype at last: essential role for the *Yersinia enterocolitica* Ysa type III secretion system in a *Drosophila melanogaster* S2 cell model. *Infect. Immun.* **81**, 2478–2487 (2013).
63. Perry, R. D., Pendrak, M. L. & Schuetze, P. Identification and cloning of a hemin storage locus involved in the pigmentation phenotype of *Yersinia pestis*. *J. Bacteriol.* **172**, 5929–5937 (1990).
64. Hinnebusch, B. J., Perry, R. D. & Schwan, T. G. Role of the *Yersinia pestis* hemin storage (hms) locus in the transmission of plague by fleas. *Science* **273**, 367–370 (1996).
65. Schroder, O. & Wagner, R. The bacterial regulatory protein H-NS — a versatile modulator of nucleic acid structures. *Biol. Chem.* **383**, 945–960 (2002).
66. Cathelyn, J. S., Ellison, D. W., Hinchliffe, S. J., Wren, B. W. & Miller, V. L. The RovA regulons of *Yersinia enterocolitica* and *Yersinia pestis* are distinct: evidence that many RovA-regulated genes were acquired more recently than the core genome. *Mol. Microbiol.* **66**, 189–205 (2007).
A report of the lineage specificity of RovA regulons in pathogenic *Yersinia* species.
67. Ellison, D. W., Lawrenz, M. B. & Miller, V. L. Invasin and beyond: regulation of *Yersinia* virulence by RovA. *Trends Microbiol.* **12**, 296–300 (2004).
68. Bucker, R., Heroven, A. K., Becker, J., Dersch, P. & Wittmann, C. The pyruvate–tricarboxylic acid cycle node: a focal point of virulence control in the enteric pathogen *Yersinia pseudotuberculosis*. *J. Biol. Chem.* **289**, 30114–30132 (2014).
The only published characterization of the *Y. pseudotuberculosis* RovA regulon.
69. Bohme, K. *et al.* Concerted actions of a thermo-labile regulator and a unique intergenic RNA thermosensor control *Yersinia* virulence. *PLoS Pathog.* **8**, e1002518 (2012).
A comprehensive analysis of the mechanism by which YmoA regulates genes on the virulence plasmid pYV.
70. Quade, N. *et al.* Structural basis for intrinsic thermosensing by the master virulence regulator RovA of *Yersinia*. *J. Biol. Chem.* **287**, 35796–35803 (2012).
71. Kapatral, V. *et al.* Gene array analysis of *Yersinia enterocolitica* FlhD and FlhC: regulation of enzymes affecting synthesis and degradation of carbamoylphosphate. *Microbiology* **150**, 2289–2300 (2004).
72. Perry, R. D. *et al.* Temperature regulation of the hemin storage (Hms$^+$) phenotype of *Yersinia pestis* is posttranscriptional. *J. Bacteriol.* **186**, 1638–1647 (2004).
73. Fälker, S., Schmidt, M. A. & Heusipp, G. Altered Ca^{2+} regulation of Yop secretion in *Yersinia enterocolitica* after DNA adenine methyltransferase overproduction is mediated by Clp-dependent degradation of LcrG. *J. Bacteriol.* **188**, 7072–7081 (2006).
74. Grabenstein, J. P., Marceau, M., Pujol, C., Simonet, M. & Bliska, J. B. The response regulator PhoP of *Yersinia pseudotuberculosis* is important for replication in macrophages and for virulence. *Infect. Immun.* **72**, 4973–4984 (2004).
75. Rebeil, R. *et al.* Induction of the *Yersinia pestis* PhoP–PhoQ regulatory system in the flea and its role in producing a transmissible infection. *J. Bacteriol.* **195**, 1920–1930 (2013).
76. Pisano, F. *et al.* Influence of PhoP and intra-species variations on virulence of *Yersinia pseudotuberculosis* during the natural oral infection route. *PLoS ONE* **9**, e103541 (2014).
77. Nuss, A. M. *et al.* A direct link between the global regulator PhoP and the Csr regulon in *Y. pseudotuberculosis* through the small regulatory RNA CsrC. *RNA Biol.* **11**, 580–593 (2014).
78. Fookes, M. *et al. Salmonella bongori* provides insights into the evolution of the *Salmonellae*. *PLoS Pathog.* **7**, e1002191 (2011).
A key publication highlighting the evolution of pathogenesis in the *Salmonella* genus.
79. Navarre, W. W. *et al.* Co-regulation of *Salmonella enterica* genes required for virulence and resistance to antimicrobial peptides by SlyA and PhoP/PhoQ. *Mol. Microbiol.* **56**, 492–508 (2005).
80. Ali, S. S. *et al.* Silencing by H-NS potentiated the evolution of *Salmonella*. *PLoS Pathog.* **10**, e1004500 (2014).
A study showing the importance of H-NS in the evolution of SPI-1-mediated pathogenesis in *Salmonella* spp.
81. Pallen, M. J. & Wren, B. W. Bacterial pathogenomics. *Nature* **449**, 835–842 (2007).
82. Morelli, G. *et al. Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nat. Genet.* **42**, 1140–1143 (2010).
A comprehensive analysis of the pattern of fine-scale and ongoing recent evolution in *Y. pestis*.
83. Song, Y. *et al.* Complete genome sequence of *Yersinia pestis* strain 91001, an isolate avirulent to humans. *DNA Res.* **11**, 179–197 (2004).
84. Bos, K. I. *et al.* A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* **478**, 506–510 (2011).
85. Chouikha, I. & Hinnebusch, B. J. Silencing urease: a key evolutionary step that facilitated the adaptation of *Yersinia pestis* to the flea-borne transmission route. *Proc. Natl Acad. Sci. USA* **111**, 18709–18714 (2014).
86. Sebbane, F., Devalckenaere, A., Foulon, J., Carniel, E. & Simonet, M. Silencing and reactivation of urease in *Yersinia pestis* is determined by one G residue at a specific position in the *ureD* gene. *Infect. Immun.* **69**, 170–176 (2001).
87. Aleksic, S. & Wuthe, H. H. *Yersinia enterocolitica* serovar 2a, wb, 3:b, c biovar 5 in hares and sheep. *Berl. Munch. Tierarztl. Wochenschr.* **110**, 176–177 (1997).
88. Swaminathan, B., Harmon, M. C. & Mehlman, I. J. *Yersinia enterocolitica*. *J. Appl. Bacteriol.* **52**, 151–183 (1982).
89. Shea, J. E., Hensel, M., Gleeson, C. & Holden, D. W. Identification of a virulence locus encoding a second type III secretion system in *Salmonella* Typhimurium. *Proc. Natl Acad. Sci. USA* **93**, 2593–2597 (1996).
90. Schaake, J. *et al.* Essential role of invasin for colonization and persistence of *Yersinia enterocolitica* in its natural reservoir host, the pig. *Infect. Immun.* **82**, 960–969 (2014).
91. Schaake, J. *et al.* Human and animal isolates of *Yersinia enterocolitica* show significant serotype-specific colonization and host-specific immune defense properties. *Infect. Immun.* **81**, 4013–4025 (2013).
92. Uliczka, F. *et al.* Unique cell adhesion and invasion properties of *Yersinia enterocolitica* O:3, the most frequent cause of human Yersiniosis. *PLoS Pathog.* **7**, e1002117 (2011).
93. Valentin-Weigand, P. Heesemann, J. Dersch, P. Unique virulence properties of *Yersinia enterocolitica* O:3 - An emerging zoonotic pathogen using pigs as preferred reservoir host. *Int. J. Med. Microbiol.* **14**, 00090–00093 (2014).

94. Parkhill, J. *et al.* Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* **413**, 848–852 (2001).
95. McClelland, M. *et al.* Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* **413**, 852–856 (2001).
96. Thomson, N. R. *et al.* Comparative genome analysis of *Salmonella* Enteritidis PT4 and *Salmonella* Gallinarum 287/91 provides insights into evolutionary and host adaptation pathways. *Genome Res.* **18**, 1624–1637 (2008).
97. Nuccio, S.-P. & Baumler, A. J. Comparative analysis of *Salmonella* genomes identifies a metabolic network for escalating growth in the inflamed gut. *mBio* **5**, e00929-14 (2014).
98. Badia, J., Ros, J. & Aguilar, J. Fermentation mechanism of fucose and rhamnose in *Salmonella* Typhimurium and *Klebsiella pneumoniae*. *J. Bacteriol.* **161**, 435–437 (1985).
99. Bry, L., Falk, P. G., Midtvedt, T. & Gordon, J. I. A model of host–microbial interactions in an open mammalian ecosystem. *Science* **273**, 1380–1383 (1996).
100. Price-Carter, M., Tingey, J., Bobik, T. A. & Roth, J. R. The alternative electron acceptor tetrathionate supports B12-dependent anaerobic growth of *Salmonella enterica* serovar Typhimurium on ethanolamine or 1,2-propanediol. *J. Bacteriol.* **183**, 2463–2475 (2001).
101. Winter, S. E. *et al.* Gut inflammation provides a respiratory electron acceptor for *Salmonella*. *Nature* **467**, 426–429 (2010).
A key account of the role of *cob*, *pdu* and *ttr* in inflammation and intestinal colonization by *Salmonella*.
102. Thiennimitr, P. *et al.* Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. *Proc. Natl Acad. Sci. USA* **108**, 17480–17485 (2011).
103. Langridge, G. C. *et al.* Patterns of genome evolution that have accompanied host adaptation in *Salmonella*. *Proc. Natl Acad. Sci. USA* **112**, 863–868 (2015).
104. Prentice, M. B. *et al.* Cobalamin synthesis in *Yersinia enterocolitica* 8081. Functional aspects of a putative metabolic island. *Adv. Exp. Med. Biol.* **529**, 43–46 (2003).
105. Rabson, A. R., Hallett, A. F. & Koornhof, H. J. Generalized *Yersinia enterocolitica* infection. *J. Infect. Dis.* **131**, 447–451 (1975).
106. Lian, C. J., Hwang, W. S., Kelly, J. K. & Pai, C. H. Invasiveness of *Yersinia enterocolitica* lacking the virulence plasmid: an in-vivo study. *J. Med. Microbiol.* **24**, 219–226 (1987).
107. Buret, A., O'Loughlin, E. V., Curtis, G. H. & Gall, D. G. Effect of acute *Yersinia enterocolitica* infection on small intestinal ultrastructure. *Gastroenterology* **98**, 1401–1407 (1990).
108. Tardy, F. *et al. Yersinia enterocolitica* type III secretion-translocation system: channel formation by secreted Yops. *EMBO J.* **18**, 6793–6799 (1999).
109. Wagner, D. M. *et al. Yersinia pestis* and the plague of Justinian 541–543 AD: a genomic analysis. *Lancet Infect. Dis.* **14**, 319–326 (2014).
110. Cui, Y. *et al.* Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proc. Natl Acad. Sci. USA* **110**, 577–582 (2013).
111. Rasmussen, S. *et al.* Early divergent strains of *Y. pestis* in Eurasia 5000 years ago. *Cell.* **163**, 1–12 (2015).
112. Hall, M. *et al.* Use of whole-genus genome sequence data to develop a multilocus sequence typing tool that accurately identifies *Yersinia* isolates to the species and subspecies levels. *J. Clin. Microbiol.* **53**, 35–42 (2015).
113. Bottone, E. J. *Yersinia enterocolitica*: overview and epidemiologic correlates. *Microbes Infect.* **1**, 323–333 (1999).
114. Barnes, P. D., Bergman, M. A., Mecsas, J. & Isberg, R. R. *Yersinia pseudotuberculosis* disseminates directly from a replicating bacterial pool in the intestine. *J. Exp. Med.* **203**, 1591–1601 (2006).
115. Bottone, E. J. *Yersinia enterocolitica*: the charisma continues. *Clin. Microbiol. Rev.* **10**, 257–276 (1997).
116. Sebbane, F., Jarrett, C. O., Gardner, D., Long, D. & Hinnebusch, B. J. Role of the *Yersinia pestis* plasminogen activator in the incidence of distinct septicemic and bubonic forms of flea-borne plague. *Proc. Natl Acad. Sci. USA* **103**, 5526–5530 (2006).
117. Sha, J. *et al.* Characterization of an F1 deletion mutant of *Yersinia pestis* CO92, pathogenic role of F1 antigen in bubonic and pneumonic plague, and evaluation of sensitivity and specificity of F1 antigen capture-based dipsticks. *J. Clin. Microbiol.* **49**, 1708–1715 (2011).

Acknowledgements
A.M. acknowledges funding from the UK Department for Environment, Food and Rural Affairs (Defra), B.W.W. acknowledges funding from Defra and the UK Biotechnology and Biological Sciences Research Council (BBSRC) and N.R.T. acknowledges funding from the Wellcome Trust (core award 098051). The authors thank their respective funding bodies for supporting the *Yersinia* spp. research in their laboratories. The authors also thank Nottingham Trent University for their generous funding through the Vice Chancellor's Researcher Development Scheme for the Ph.D. research of S.R. on evolution in *Yersinia* species.

Competing interests statement
The authors declare no competing interests.