

# THE YERSINIAE — A MODEL GENUS TO STUDY THE RAPID EVOLUTION OF BACTERIAL PATHOGENS

*Brendan W. Wren*

*Yersinia pestis*, the causative agent of plague, seems to have evolved from a gastrointestinal pathogen, *Yersinia pseudotuberculosis*, in just 1,500–20,000 years — an ‘eye blink’ in evolutionary time. The third pathogenic *Yersinia*, *Yersinia enterocolitica*, also causes gastroenteritis but is distantly related to *Y. pestis* and *Y. pseudotuberculosis*. Why do the two closely related species cause remarkably different diseases, whereas the distantly related enteropathogens cause similar symptoms? The recent availability of whole-genome sequences and information on the biology of the pathogenic yersiniae have shed light on this paradox, and revealed ways in which new, highly virulent pathogens can evolve.

## ENTEROPATHOGENIC

Describes a pathogen that resides in the gastrointestinal tract.

## PANDEMIC

The worldwide spread of an epidemic.

The yersiniae are Gram-negative bacteria that belong to the family *Enterobacteriaceae*. They consist of 11 species that have traditionally been distinguished by DNA–DNA hybridization and biochemical analyses<sup>1–4</sup>. Three of them are pathogenic to humans: *Yersinia pestis* and the ENTEROPATHOGENIC yersiniae, *Yersinia pseudotuberculosis* and *Yersinia enterocolitica*. All three species target the lymph tissues during infection and carry a 70-kb virulence plasmid (pYV), which is essential for infection in these tissues, as well as to overcome host defence mechanisms<sup>1–4</sup>.

Elegant population genetics studies have indicated that *Y. pestis* recently evolved from *Y. pseudotuberculosis* — just 1,500–20,000 years ago<sup>5</sup>. However, *Y. pestis* has a very different ecology and epidemiology to the enteropathogenic yersiniae, and causes a markedly different disease. From a mammalian enteropathogen that is widely found in the environment, it has rapidly transformed into a blood-borne pathogen that is also able to parasitize insects and cause systemic disease. The varied ecology, pathogenicity and host range of these related species, together with the availability of mouse models for both gastroenteritis and plague that mimic human disease, and the relative ease of constructing defined mutants, make the yersiniae a model genus in which to study the genetics and evolution of bacterial pathogens.

This review covers emerging themes from the genome sequences of the pathogenic yersiniae and discusses how this information is guiding hypotheses on the evolution of this genus.

*Yersinia pestis* — the causative agent of plague — *Y. pestis* has been responsible for three human PANDEMICS — the Justinian plague (fifth to seventh centuries), the Black Death (thirteenth to fifteenth centuries) and modern plague (1870s to the present day)<sup>1,3</sup>. The Black Death alone is estimated to have claimed the lives of one-third of the European population, and has shaped the development of modern civilization. Plague is still with us, circulating in various mammalian species on most continents (FIG. 1). The recent identification of multidrug resistant strains<sup>6</sup> and the possibility that it could be used as an agent of biological warfare, mean that plague still poses a significant threat to human health. An effective vaccine that induces long-lived immunity against bubonic and pneumonic plagues is still not available.

*Y. pestis* has a complex life cycle involving a mammalian reservoir (primarily rodents) and a flea vector<sup>1,3</sup> (FIG. 2). The bacterium ensures transmission by forming a cohesive aggregate that blocks the foregut of infected fleas<sup>7</sup>. This results in futile attempts to feed on a new

Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. e-mail: [brendan.wren@lshtm.ac.uk](mailto:brendan.wren@lshtm.ac.uk)  
doi:10.1038/nrmicro730

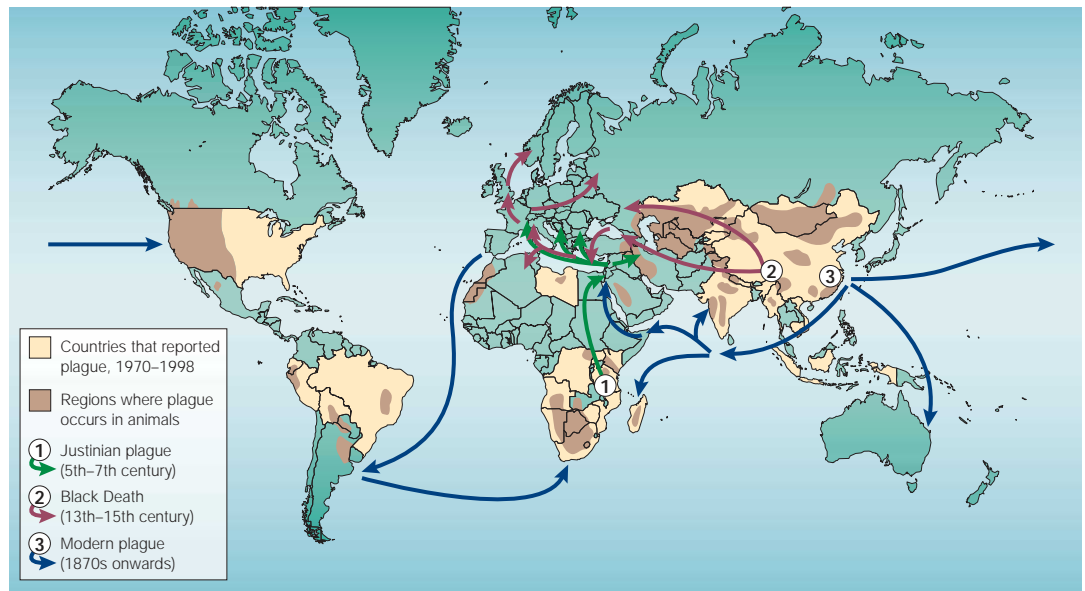


Figure 1 | **World distribution of plague, 1998.** Modified from REF. 5 © (1999) National Academy of Sciences.

host as a blocked flea regurgitates infected blood back into the bite site, effectively injecting the bacteria under the animal's skin<sup>7</sup>. On infection, some of the bacilli are engulfed by macrophages and carried to the regional lymph nodes, draining the infection site. The bulk of the bacteria, protected by plasminogen activator, penetrate tissue, eventually accessing the lymphatic vessels that drain regional lymph nodes<sup>8</sup>, where they multiply, giving rise to the classical symptoms of infection and then BACTERAEMIA<sup>1,3</sup>. This is known as bubonic plague. If the infection then progresses to the lungs, pneumonic plague develops, which is highly infectious and rapidly fatal.

There are three recognized subgroups, or BIOVARs, of *Y. pestis*: Antiqua, Mediaevalis and Orientalis. They are differentiated by their abilities to ferment glycerol and to reduce nitrate, but these differences do not seem to correlate with virulence. Based on epidemiological observations and historical records, each biovar has been associated with one of the three pandemics. Biovar Antiqua is resident in Africa and is descended from bacteria that caused the Justinian plague; biovar Mediaevalis is resident in central Asia and is descended from bacteria that caused the Black Death; and biovar Orientalis, which at present is widespread, is associated with modern plague<sup>9</sup>. The factors influencing the rise and fall of plague EPIDEMICS remain obscure, but it is possible that severe epidemics could be preceded by subtle genetic changes in *Y. pestis* that result in a highly virulent strain. One of the aims of the *Y. pestis* CO92 biovar Orientalis and *Y. pestis* KIM10 biovar Mediaevalis genome projects<sup>10,11</sup> was to identify these key genetic changes.

**Enteropathogenic yersiniae**  
The enteropathogenic yersiniae are found widely in the environment — for example, in soil — and are a common cause of animal infections, affecting several mammalian and avian species<sup>2,4</sup>. In humans, the infection causes gastroenteritis after the consumption of

contaminated food or water. After ingestion, the bacteria pass into the small intestine, where they translocate across the intestinal epithelium through Peyer's patches (FIG. 2). They then migrate to the mesenteric lymph nodes and are subsequently found in the liver and spleen, where they replicate externally to the host cells<sup>12</sup>. After multiplication, rapid inflammation ensues, which gives rise to the symptoms that are associated with gastroenteritis, such as mesenteric lymphadenitis and terminal ileitis.

*Y. enterocolitica* comprises a biochemically and genetically heterogeneous collection of organisms that has been divided into six biogroups — known as 1A, 1B, 2, 3, 4 and 5 — that can be differentiated by biochemical tests. These can be placed into three lineages: a non-pathogenic group (biogroup 1A); a weakly pathogenic group that is unable to kill mice (biogroups 2 to 5); and a highly pathogenic, mouse-lethal group (biogroup 1B). Biogroup 1A lacks the *Yersinia* virulence plasmid pYV and seems to be distantly related to the other biogroups, whereas biogroup 1B forms a geographically distinct group of strains<sup>13</sup> that are frequently isolated in North America (the so-called 'New-World' strains) and biogroups 2 to 5 are predominantly isolated in Europe and Japan ('Old-World' strains). *Y. pseudotuberculosis* is subgrouped into 21 different serological groups based on variations in the O-ANTIGEN of its lipopolysaccharide (LPS). All *Y. pestis* strains fail to express an O-antigen.

**The *Yersinia pestis* genome recipe**  
The first *Yersinia* strain to be sequenced was *Y. pestis* CO92 biovar Orientalis<sup>10</sup>. The CO92 strain was originally isolated in 1992 from a veterinarian in Colorado, USA, who was infected by a sneezing cat. The genome consists of a 4.65-Mb chromosome and three plasmids — pMT1 or pFra (96.2 kb), pYV or pCD (70.3 kb), and pPla or pPCP1 (9.6 kb) (TABLE 1). Close inspection of the chromosome sequence indicates that it has undergone

**BACTERAEMIA**

A bacterial infection of the blood and/or lymphatic system.

**BIOVAR**

The phenotypical distinction of bacteria in the same species based on biological tests, such as simple biochemical and/or enzymatic differences.

**EPIDEMIC**

The rapid spread of an infectious disease in a defined geographical area.

**O-ANTIGEN**

The lipopolysaccharide surface antigen that is present in most Gram-negative bacteria. It is often used as a convenient immunological test to distinguish between bacteria of the same species.

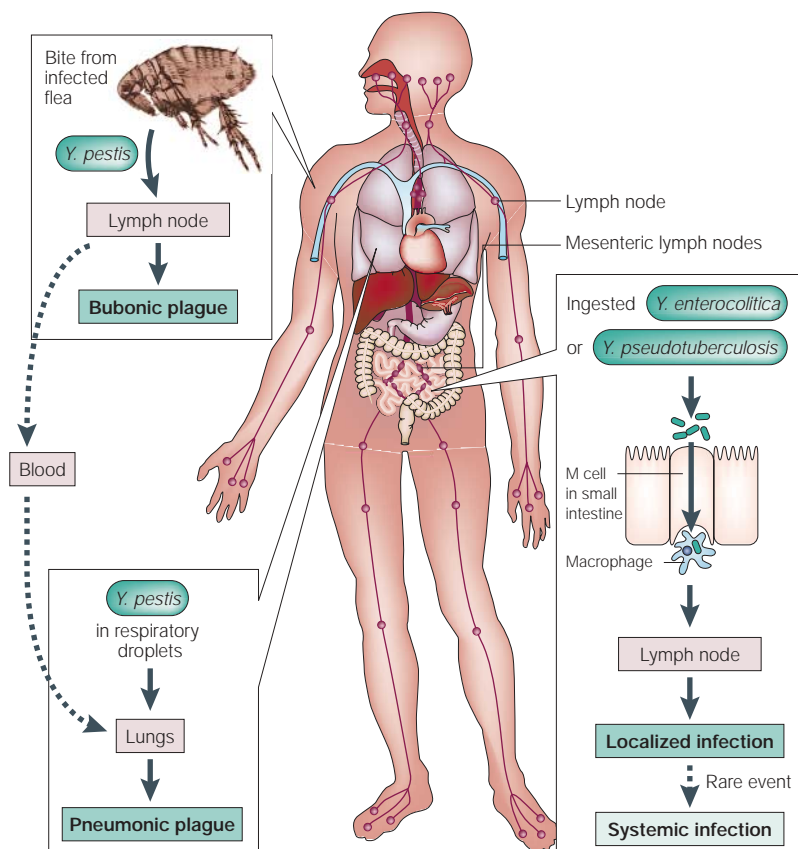


Figure 2 | Steps in the transmission of the pathogenic yersiniae in humans. *Y. pestis* has a rodent reservoir. The rodent's fleas acquire *Y. pestis* from a meal of infected blood, and transmit the bacterium primarily to other rodents, or occasionally to humans — causing bubonic plague in humans. Human-to-human transmission can occur through human fleas. Pneumonic plague is transmitted from human to human through respiratory droplets, or possibly by artificially generated *Y. pestis* aerosols. *Y. enterocolitica* and *Y. pseudotuberculosis* are ingested, and in contrast to *Y. pestis*, enter the lymphatic system through the M cells of the small intestine.

genetic material that is acquired often contributes to an organism's virulence — for example, by broadening its host range, improving its ability to overcome host defences or to cause tissue damage. The process of LATERAL GENE TRANSFER allows microbial pathogens to evolve extremely rapidly, in 'quantum leaps'.

Plasmid acquisition seems to be a key element in the evolutionary jump of *Y. pestis* from an enteric pathogen to a flea-transmitted systemic pathogen. In addition to the virulence plasmid pYV that is common to all pathogenic yersiniae, most *Y. pestis* strains have two further plasmids — pPla that encodes the plasminogen activator Pla<sup>14</sup>, and pMT1 that encodes the putative murine toxin Ymt and the F1 capsule. The precise role of these determinants in host adaptation and virulence is unclear, but there are several indications that they are involved in transmission. Pla, for example, is important for dissemination of *Y. pestis* after subcutaneous injection into a mammalian host<sup>15</sup>, and although capsule-deficient mutants can still cause disease in humans<sup>16</sup>, *Y. pestis* strains that lack the entire pMT1 plasmid are unable to colonize fleas<sup>17</sup>. Recently, Hinnebusch *et al.* have shown that the murine toxin Ymt acts as an intracellular phospholipase D and is required for the survival of *Y. pestis* in the flea midgut compartment, but not in the proventriculus compartment<sup>18</sup>. However, the chromosomally unstable haemin storage (*hms*) locus that encodes outer-surface proteins is also required for flea-borne transmission. Conversely, the *hms* locus is required to infect the proventriculus and not the midgut of the flea<sup>19</sup>. Deletion of the *hms* locus in *Y. pestis* results in changes in blood-feeding behaviour and less efficient transmission of plague. So overall, the acquisition of two plasmids (pPla and pMT1) by horizontal gene transfer along with the pre-existing chromosomal *hms* locus, helps to explain the rapid evolutionary transition of *Y. pestis* to flea-borne transmission.

Although the existence of the pPla and pMT1 plasmids and the *hms* locus was known before the CO92 genome sequence became available, little was known about the rest of the 4.65-Mb chromosome. One characteristic of loci that have been acquired by lateral gene transfer is an atypical G+C content relative to the rest of the genome. This means that the complete genome sequence can be screened for recently acquired genes or cassettes of genes — often referred to as pathogenicity islands — by looking for 'spikes' of G+C variation. G+C analysis of the *Y. pestis* CO92 genome identified at least 21 such regions, including the 102-kb unstable element that contains the *hms* locus<sup>10</sup> (TABLE 2). Among these

fierce genetic flux. It is scarred by genes that have been acquired from other organisms, large sections are jumbled and it seems to be in the early stages of decay. These main features of this highly dynamic genome can be seen as representing three key steps towards the evolution of *Y. pestis*: add DNA, stir and reduce.

**Add DNA.** As the nucleotide sequences of more bacterial genomes become available, it is evident that they are mosaics of DNA sequences from different origins, owing to the lateral exchange of large mobile genetic elements, such as plasmids, phage or transposons. The

Table 1 | Plasmids important in the virulence of pathogenic yersiniae

| Plasmid name                           | Other designations | Size (kb) | Virulence determinants                                | Role in disease                                   |
|--|--------------------|-----------|---|---|
| <i>Yersinia</i> virulence plasmid, pYV | pCD                | 70.3      | Several YOPs, type III secretion system               | Avoidance of the immune system, toxicity          |
| Plasmid-encoding murine toxin, pMT1    | pFra               | 96.2      | Murine toxin (phospholipase), F1 capsule-like antigen | Bacterial transmission by fleas                   |
| Plasminogen-activating plasmid, pPla   | pPCP1, pPst        | 9.6       | Plasminogen activator                                 | Dissemination from intra-dermal site of infection |

LATERAL GENE TRANSFER  
The transfer of DNA — frequently cassettes of genes — between organisms.

Table 2 | Selected putative pathogenicity islands from *Y. pestis* (*Y. pseudotuberculosis*)

| Range           | Overall G+C content | Insertion site of region | Putative function                        |
|-----------------|---------------------|--------------------------|--|
| YPO0255–YPO0273 | 49.1%               | –                        | Type III secretion system                |
| YPO0335–YPO0340 | 36.6%               | tRNA-Phe                 | Insect viral-enhancing factor            |
| YPO0590–YPO0642 | 50.2%               | tRNA-Met                 | Adhesin, autotransporter, protein kinase |
| YPO0684–YPO0697 | 36.6%               | IS 1541                  | Adhesin                                  |
| YPO0770–YPO0778 | 49.2%               | IS 100                   | HPI2 — siderophore biosynthesis          |
| YPO0803–YPO0818 | 32.8%               | –                        | Type-II-related secretion system         |
| YPO0961–YPO0995 | 48.3%               | tRNA                     | Quorum sensing, siderophore biosynthesis |
| YPO1448–YPO1480 | 45.8%               | tRNA-Ser                 | CNF, fatty-acid metabolism               |
| YPO1900–YPO1917 | 56.4%               | tRNA-Asn                 | HPI1 (Yersiniabactin biosynthesis)       |
| YPO1951–YPO2004 | 46.8%               | –                        | Hms (pigmentation locus)                 |
| YPO2311–YPO2321 | 46.9%               | IS 1541                  | Insecticidal toxin complex (TccC)        |
| YPO2434–YPO2443 | 44.0%               | –                        | Iron transport, antibiotic resistance    |
| YPO2934–YPO2948 | 45.4%               | IS 100                   | Chaperone/usher fimbrial system          |
| YPO3673–YPO3682 | 45.1%               | –                        | Insecticidal toxin complex (TcaABC)      |
| YPO4014–YPO4033 | 44.4%               | tRNA-Ser                 | Iron transport                           |

The table shows putative pathogenicity islands that have been identified by notable G+C variation. Overall G+C content of *Y. pestis* chromosome is 47.6%. CNF, cytotoxic necrotizing factor; HPI, high-pathogenicity island.

regions were several genes that seem to have come from other insect pathogens. Sequences related to the parasitism of insects include homologues of insecticidal toxin complexes (Tcs) from *Photorhabdus luminescens*, *Serratia entomophila* and *Xenorhabdus nematophilus*<sup>20</sup>. The toxins are complexes of the products of three different gene families — namely, *tcaA/tcaB/tcdA*, *tcaC/tcdB* and *tccC*. In addition, a predicted coding sequence showing similarity to an insect virus-like enhancer protein — a proteolytic enzyme that can damage insect gut membranes — was also identified in a region of low G+C content<sup>10</sup>. The sequence was flanked by transposase fragments, which indicates horizontal acquisition.

Other apparent acquisitions include a chromosomally encoded type-III secretion system that is similar in gene content and order to the Spi2 type III system of *Salmonella enterica* serovar Typhimurium<sup>21</sup>, and several adhesins and iron-scavenging systems (TABLE 2). The *Y. pestis* CO92 genome sequence also contains several genes that are predicted to encode novel surface antigens that could have a role in virulence. Ten fimbrial-type surface structures that are often important in bacterial attachment were identified, five of which were flanked by genes encoding transposases or integrases, which also indicates horizontal acquisition<sup>10</sup>. A large arsenal of independent gene clusters encoding different fimbriae and adhesins could help *Y. pestis* evade the host immune response, or allow multiple interactions with several different host tissues during the pathogen's complex life cycle.

**Stir.** A striking feature of the *Y. pestis* CO92 genome sequence was the large number of insertion sequence (IS) elements. The total of 140 IS elements exceeds that described in most other bacterial genomes and comprises 3.7% of the genome. IS elements are perfectly repeated sequences, and are likely sites for homologous recombination events that can rearrange the genome.

All bacterial genomes sequenced so far have a small, but detectable bias towards G on the leading strand of the bi-directional replication fork<sup>22</sup>. So, the G/C skew in different parts of the genome highlights any irregularities in the composition of the genome (FIG. 3).

The G/C skew plot of *Y. pestis* CO92 shows three anomalies — two inversions and one translocation<sup>10</sup> (FIG. 3). Each is bounded by IS elements, which indicates that they could be the result of recent recombination events. Polymerase chain reaction analysis indicates that although the chromosome contains both possible orientations of the two inverted sequences, they are present in unequal proportions<sup>10</sup>. This has also been shown to be the case for strain GB (biovar Orientalis) and strain A16 (biovar Antiqua)<sup>10</sup>. It seems that several different chromosomal configurations can exist in the same population, which indicates that genomic rearrangements can occur during the growth of the organism — a feature that has not previously been reported for a bacterium. It is not known how these events affect the biology of the organism, but as the expression of bacterial genes is influenced by their orientation with respect to the direction of DNA replication, it seems reasonable to conclude that such rearrangements could alter virulence.

**Reduce.** Loss of gene function, or genome decay, occurs as a bacterium adapts to its host. For example, many PSEUDOGENES — which are often ignored as sequencing artefacts — could, in fact, be remnants of functional genes from a pathogen that is in the process of 'downsizing' its genome content as it adapts. This is illustrated by the genome sequence of the obligate intracellular pathogen *Rickettsia prowazekii*<sup>23</sup>. The 1.11-Mb *R. prowazekii* genome has a large number of pseudogenes and the highest proportion of non-coding DNA in any prokaryote — more than 24%. The intergenic DNA might represent the scattered remnants of genes that are

#### PSEUDOGENES

A gene that is no longer functional — usually owing to a point mutation — that results in a frame shift in the coding sequence.

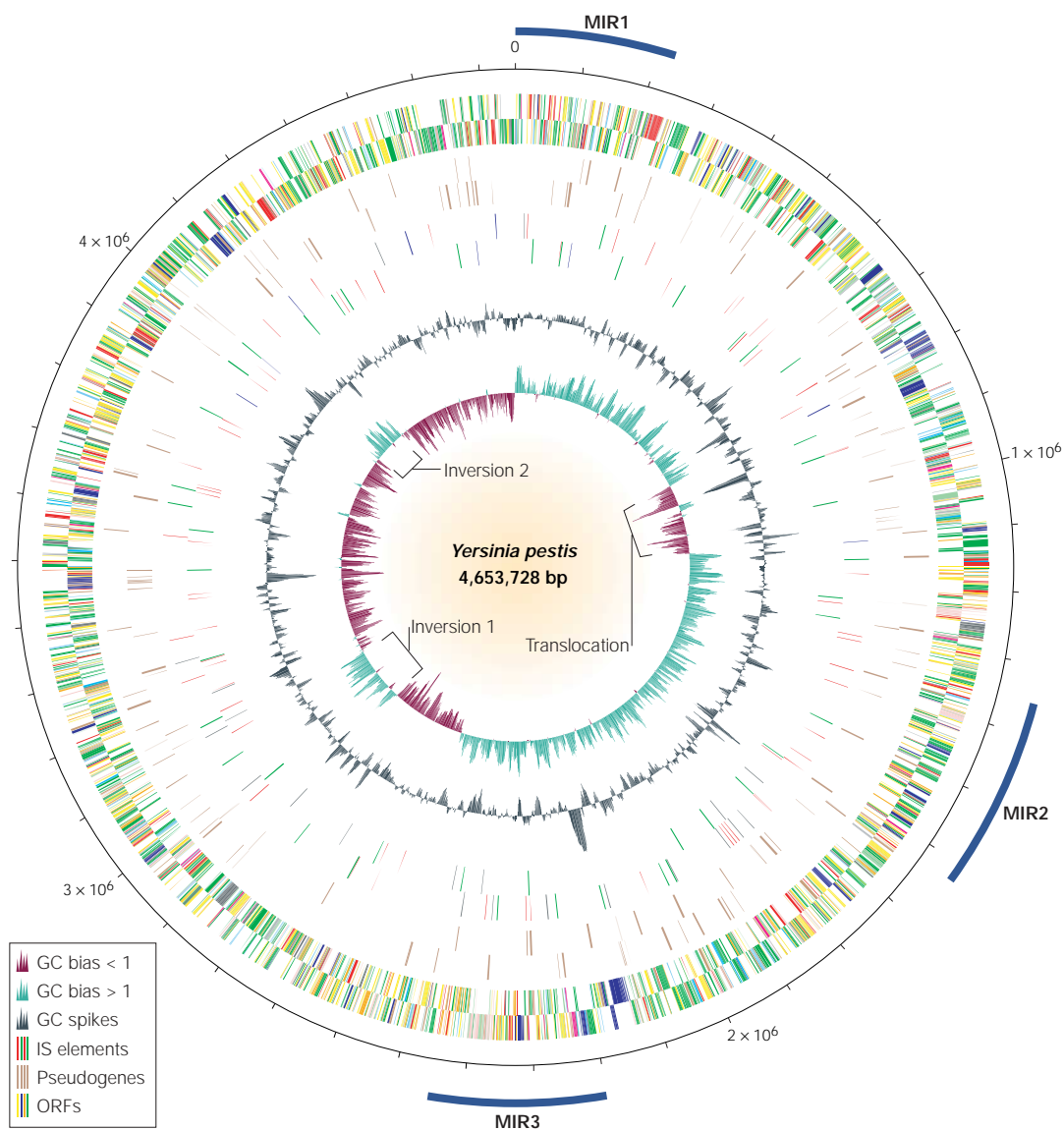


Figure 3 | **Circular representation of the *Y. pestis* CO92 genome.** The outer circle is marked in bases. Circles 1 and 2 (from the outside in), all genes colour-coded by function, forward and reverse strand; circles 3 and 4, pseudogenes; circles 5 and 6, insertion sequence elements; circle 7, G+C content (higher values outward); circle 8, GC bias  $((G-C)/(G+C))$ , turquoise indicates values  $>1$ , burgundy indicates values  $<1$ . Colour-coding for genes: dark blue, pathogenicity or adaptation; black, energy metabolism; red, information transfer; dark green, surface-associated; cyan, degradation of large molecules; magenta, degradation of small molecules; yellow, central or intermediary metabolism; pale blue, regulators; orange, conserved hypothetical; brown, pseudogenes; pink, phage and insertion sequence elements; pale green, unknown; grey, miscellaneous. The three multiple inversion regions (MIR) are highlighted in dark blue. IS, insertion sequence; ORFs, open reading frames. Reproduced from REF. 10 © (2001) Nature Macmillan Magazines Ltd.

no longer required (or are harmful to the existence of the organism), and that have been lost in a step-wise process as the organism acquires an obligate intracellular lifecycle. Analysis of the genome of the leprosy bacillus *Mycobacterium leprae* indicates a similar situation as the genome contains numerous pseudogenes and extensive genetic downsizing that is not found in other mycobacterial species that have a more free-living existence<sup>24</sup>. It has been proposed that the genome of *M. leprae* has evolved to the natural minimal gene-set for mycobacteria, and is now a pathogen that is on the brink of survival<sup>24</sup>.

Close examination of the CO92 nucleotide sequence has identified at least 149 pseudogenes, which represent 4% of the genome. Several mechanisms account for the accumulation of pseudogenes in *Y. pestis*, including IS-element expansion, deletion, point mutation and slippage in tracts of single-nucleotide repeats. The total of 149 is likely to be an underestimate, as certain features such as point mutations are difficult to detect through direct sequencing. Although the number is modest compared to those of *R. prowazekii* and *M. leprae*, it indicates that reductive evolution has begun, and reflects adaptation to new hosts and a gradual change in life cycle.

Table 3 | *Y. pestis* pseudogenes that might be important in the pathogenesis of other *Yersinia*

| Pathogenicity gene | Name        | Predicted function if gene is intact                   | Mutation   |
|--------------------|-------------|--|------------|
| YPO0603            |             | Putative haemagglutinin/haemolysin-related protein     | Partial    |
| YPO0704            |             | Flagellar assembly protein                             | F/S        |
| YPO0727            | <i>flgF</i> | Putative flagellar basal-body rod protein              | IS100      |
| YPO0765            |             | Putative adhesion                                      | F/S        |
| YPO0989            | <i>iucA</i> | Putative siderophore biosynthesis protein, <i>lucA</i> | F/S        |
| YPO1449            | <i>cnf</i>  | Cytotoxic necrotizing factor 1                         | Partial    |
| YPO1562            |             | Putative intimin                                       | IS285      |
| YPO1657            |             | Methyl-accepting chemotaxis protein                    | F/S        |
| YPO1676            |             | Methyl-accepting chemotaxis protein                    | Stop codon |
| YPO1793            | <i>inv</i>  | Invasion   | IS200-like |
| YPO1835            |             | Putative chemotaxis protein                            | F/S        |
| YPO1967            | <i>nmpC</i> | Porin  | IS100      |
| YPO2249            |             | Putative toxin transport protein                       | F/S        |
| YPO2250            |             | Putative toxin secretion protein                       | F/S        |
| YPO2671            | <i>ureD</i> | Urease accessory protein                               | F/S (7G)   |
| YPO2851            | <i>baeS</i> | Two-component system, sensor kinase                    | F/S (7G)   |
| YPO2887            | <i>yapB</i> | Putative autotransporter protein                       | F/S        |
| YPO3100            | <i>fcl</i>  | GDP-L-fucose synthetase                                | F/S (6G)   |
| YPO3102            | <i>gmd</i>  | Probable GDP-mannose 4,6-dehydratase                   | F/S        |
| YPO3105            | <i>wzy</i>  | Putative O-unit polymerase protein                     | F/S (9T)   |
| YPO3108            | <i>wbyl</i> | Putative glycosyltransferase                           | F/S        |
| YPO3114            | <i>ddhB</i> | CDP-D-glucose-4,6-dehydratase                          | F/S (11C)  |
| YPO3340            |             | Putative exogenous ferric siderophore receptor         | Stop codon |
| YPO3679            | <i>tcaB</i> | Insecticidal toxin complex protein, TcaB               | F/S (6A)   |
| YPO4008            | <i>uhpB</i> | Two-component system, sensor kinase                    | IS100      |
| YPO4042            |             | Fimbrial usher protein                                 | F/S        |

The table shows selected *Y. pestis* pseudogenes that might be important in *Y. enterocolitica* and *Y. pseudotuberculosis* pathogenesis. F/S, frameshift (parentheses denote runs of nucleotides that indicate the possibility of reversible switching of the expressed phenotype); IS, insertion sequence; stop codon, truncation of the encoded protein due to a premature stop codon.

A selected list of *Y. pestis* pseudogenes that might once have been involved in host adaptation or virulence is shown in TABLE 3. These include putative insect toxin genes, such as *tcaB* and *tcaA* — which contain a frameshift mutation and an internal deletion, respectively — and the general toxin cytotoxic necrotizing factor 1. The disruption of these genes might be necessary for the life cycle of *Y. pestis*, which persists in the flea gut for relatively long periods of time and would not want to kill its insect host.

As *Y. pestis* has changed its life cycle from that of the ancestral *Y. pseudotuberculosis* strain, it would not be expected to use genes that are required for enteropathogenicity as the newly evolved *Y. pestis* would no longer be transmitted by the faecal–oral route. Enteropathogens specifically adhere to surfaces of the gut and invade the cells lining it. Proteins that are important for this process in *Y. pseudotuberculosis* include *YadA* and *Inv*, both of which are represented by pseudogenes in *Y. pestis*<sup>25,26</sup>. Many of the other pseudogenes reported — for example, a putative intimin adhesion protein — might have encoded adhesin molecules that potentially had a role in enteropathogenesis.

Some pseudogenes might be able to regain their function. Several pathogens have been shown to be capable of switching surface-expressed antigens on or off by slipped-strand mispairing of repeat sequences during replication<sup>27</sup>, and a similar process has been shown in *Y. pestis* in the *ureD* gene. This organism is characteristically UREASE-negative; however, activity can be restored *in vitro* by the spontaneous deletion of a single base pair in runs of nucleotides in the *ureD* gene<sup>28</sup>. This type of reversible mutation would free *Y. pestis* from the metabolic burden of producing proteins that are not required in its new flea/mammal life cycle, yet still give *Y. pestis* the potential to express them should a subsequent need arise. However, the loss of some enteropathogen virulence traits seem to be irreversible in *Y. pestis* as the gene pathways encoding them have been inactivated by multiple mutations. Motility and LPS biosynthesis are examples in which at least five genes in each pathway seem to no longer function in *Y. pestis* CO92 (REFS 10,29).

The *Y. pestis* CO92 genome also contains several pseudogenes of unknown function. Given that many of the familiar pseudogenes seem to be associated with a redundant enteric life cycle, identifying these sequences

#### UREASE

An enzyme that converts CO<sub>2</sub> and NH<sub>3</sub> to the alkali urea. It is important for enteropathogenic bacteria to survive in the acidic environment of the stomach during its enteric life cycle.

Table 4 | *Y. pestis* pseudogenes that might be important in central/intermediary metabolism

| Gene    | Name        | Function  | Mutation   |
|---------|-------------|---|------------|
| YPO0115 | <i>metB</i> | Cystathionine gamma-synthase  | F/S (7G)   |
| YPO0124 | <i>malP</i> | Maltodextrin phosphorylase  | F/S        |
| YPO0170 | <i>argD</i> | Acetylornithine aminotransferase  | IS100      |
| YPO0437 | <i>deoA</i> | Thymidine phosphorylase   | F/S        |
| YPO1012 |             | Putative peptide/polyketide synthase subunit                            | IS100      |
| YPO1130 | <i>aroG</i> | Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive              | IS100      |
| YPO1428 |             | 3-oxoacyl-(acyl-carrier-protein) synthase II                            | IS100      |
| YPO1761 | <i>hpaG</i> | 4-hydroxyphenylacetate degradation bifunctional isomerase/decarboxylase | F/S        |
| YPO1977 |             | 3-oxoacyl-(acyl-carrier-protein) synthase III                           | F/S        |
| YPO2171 | <i>purU</i> | Formyltetrahydrofolate deformylase                                      | IS285      |
| YPO2579 |             | Putative myo-inositol dehydrogenase                                     | F/S        |
| YPO3250 |             | Putative transaminase   | IS100      |
| YPO3281 | <i>pheA</i> | P-protein (includes chorismate mutase and prephenate dehydratase)       | IS100      |
| YPO3370 | <i>cysH</i> | Phosphoadenosine phosphosulphate reductase                              | F/S        |
| YPO3687 |             | Putative aldehyde dehydrogenase   | F/S (6G)   |
| YPO3870 | <i>gppA</i> | Guanosine-5'-triphosphate,3'-diphosphate pyrophosphatase                | F/S        |
| YPO3937 | <i>glpD</i> | Aerobic glycerol-3-phosphate dehydrogenase                              | Partial    |
| YPO4058 | <i>fdoG</i> | Formate dehydrogenase-O, major subunit                                  | Stop codon |

F/S, frameshift (parentheses denote runs of nucleotides that indicate the possibility of reversible switching of the expressed phenotype); IS, insertion sequence; stop codon, truncation of the expressed protein due to a premature stop codon.

in *Y. pestis* might reveal potential virulence determinants for investigation in the enteropathogenic yersiniae.

However, not all of the lost genes relate to putative virulence determinants — many relate to physiological functions (TABLE 4). Indeed, it is becoming increasingly evident that some genes increase the virulence of the organism when they are inactivated. When pathogenic *Shigella* strains developed from a non-pathogenic *Escherichia coli* ancestor, the loss of *ompT* and *cadA* genes (so-called 'black holes') might have contributed to their virulence and evolution<sup>30,31</sup>. More recently, this phenomenon has been shown in *Mycobacterium tuberculosis*, in which several experimentally designed knockout mutants are more virulent than the wild-type strain<sup>32</sup>. This might also be the case for *Y. pestis*, particularly with respect to the loss of bioenergetic functions, such as dicarboxylic amino-acid metabolism. It has been known for some time that all the *Y. pestis* strains tested so far lack enzymes that alter the catabolic flow of carbon, such as glucose-6-phosphate dehydrogenase and aspartase<sup>3,33,34</sup>. The reduction of unnecessary metabolic pathways might allow the organism to conserve energy and the newly evolved, streamlined organism might then contribute to the development of acute disease.

#### Other *Yersinia* genome projects

A second *Y. pestis* strain KIM10 (biovar Mediaevalis), which was originally isolated from a plague patient in Kurdistan, has recently been sequenced<sup>11</sup>. This provides the opportunity to compare the genome content of a modern-day plague strain with that of a more ancient strain, even though only a few hundred years separate the pandemic events by which these biovars are defined. Direct comparison showed that the two genomes share

more than 98% of their sequence, although the KIM10 genome is about 50 kb smaller than the CO92 genome as it has fewer IS elements and several small deletions. However, despite the genomes being so closely related, a remarkable amount of genome rearrangement has taken place. The differences seem to result from multiple inversions of the sequence at various insertion points. In particular, there are three regions where multiple inversions have taken place (FIG. 3). The most complicated multiple inversion region, MIR1, spans the replication origin and contains at least nine inversions<sup>11</sup> (FIG. 3).

The genome sequences of *Y. pseudotuberculosis* IP 32953 (serotype I) (see [Y. pseudotuberculosis genome sequence](#) in the Online links) and *Y. enterocolitica* 8081 (biogroup 1B, serotype O8) (see [Y. enterocolitica genome sequence](#) in the Online links) are also soon to be published. They too seem to have gained significant portions of their genome by lateral gene transfer. However, both genomes have fewer IS elements and pseudogenes than *Y. pestis* strains KIM10 or CO92 (REFS 35, 36). This indicates that the *Y. pseudotuberculosis* and *Y. enterocolitica* genomes are more stable than *Y. pestis*. Genome-sequence data confirm that *Y. pestis* and *Y. pseudotuberculosis* are closely related, with gene homology of nearly 97% and largely co-linear gene organization<sup>35</sup>. By contrast, *Y. enterocolitica* is more distantly related, and is about the same evolutionary distance from *Y. pseudotuberculosis* and *Y. pestis* as *E. coli* is from *Salmonella* species<sup>36</sup>.

Meanwhile, closer inspection of the disease syndromes of *Y. enterocolitica* and *Y. pseudotuberculosis* indicates that, although they seem similar, the two species do in fact cause different infections. Although both pathogens invade through M cells, *Y. enterocolitica* colonizes the Peyer's patches, whereas *Y. pseudotuberculosis* is

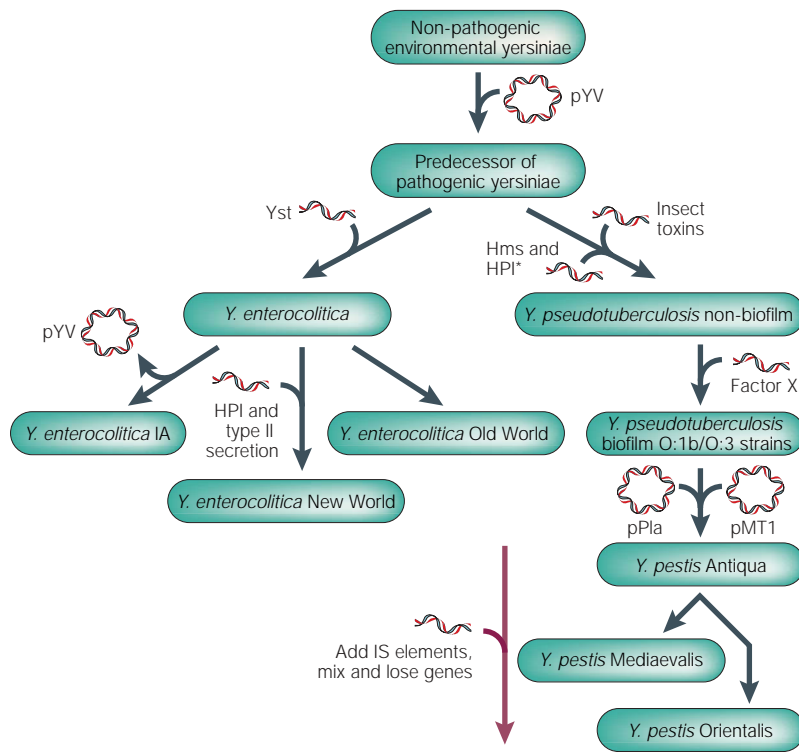


Figure 4 | Simplified model of *Yersinia* species evolution based on present knowledge of genome data. The non-pathogenic yersiniae gain the virulence plasmid pYV to form the predecessor of pathogenic yersiniae. *Y. enterocolitica* diverges from *Y. pseudotuberculosis* and forms three lineages: 1A, Old World and New World. *Y. pseudotuberculosis* gains the ability to parasitize insects and form biofilms in hosts before evolving into *Y. pestis* through the acquisition of the plasmids pPla and pMT1, genome mixing and decay. Hms, haemin storage; HPI and HPI\*, high-pathogenicity islands; IS, insertion sequence.

more widely disseminated and typically causes acute abdominal pain with mesenteric lymphadenitis of the small intestine. One distinguishing feature of *Y. enterocolitica* disease compared to *Y. pseudotuberculosis* is that it causes a more severe diarrhoea — a pronounced watery diarrhoea and occasionally bloody diarrhoea with fever in children. The heat-stable toxin Yst has been identified in all enteropathogenic *Y. enterocolitica*, but is absent in *Y. pseudotuberculosis*<sup>37</sup>. This could be one of the distinguishing genetic features responsible for this difference in symptoms. So, although diarrhoea is a common outcome, the diseases are different. This partly explains the ‘*Yersinia paradox*’, although it does not shed light on why *Y. pestis* causes such a different disease.

The evolution of pathogenic yersiniae  
The insights from genome analysis allow us to piece together a picture of how these three species might have evolved. It seems clear that *Y. enterocolitica* has evolved independently. As discussed, it can be separated into three lineages (FIG. 4) — the mostly avirulent biogroup 1A strains that lack the virulence plasmid pYV, the mouse-virulent Old-World strains (biogroups 2 to 5) and mouse-lethal New-World strains (biogroup 1B).

The New-World strains have acquired several elements by lateral gene transfer that contribute to their increased virulence compared with Old-World strains.

In particular, the New-World strains contain a ‘high-pathogenicity island’ (HPI) that encodes the synthesis of the siderophore yersiniabactin, an iron-sequestering low-molecular-weight compound that is invaluable in the iron-limiting environment of the host<sup>38</sup>. The importance of the HPI region to mouse virulence has been shown by transferring it from a New-World strain into an Old-World strain; the modified strain was lethal in mice<sup>39</sup>. The HPI region has also been found in other *Enterobacteriaceae*<sup>40</sup>, some of which might be candidates for donating the HPI region to the *Y. enterocolitica* New-World strains and also to *Y. pseudotuberculosis* and *Y. pestis*. A HPI region is also present in *Y. pseudotuberculosis* and *Y. pestis*, but sequence analysis reveals that it is significantly different to the HPI region that is present in *Y. enterocolitica*, indicating that the two regions might have been acquired independently<sup>40</sup>. More recently, another element has been identified that seems to occur exclusively in New-World strains — that is, an additional type-II secretion gene cluster<sup>2</sup>.

By contrast, *Y. pestis* is much more closely related to *Y. pseudotuberculosis*, indicating that it evolved from this species in a short amount of time. There is strong evidence that strains of the O:1b serogroup of *Y. pseudotuberculosis* are the most closely related to *Y. pestis* owing to the 98.9% sequence identity of the LPS region of this serogroup compared to that of *Y. pestis*<sup>29</sup>. However, O:1b strains are not clonal, and the ancestral O:1b strain for *Y. pestis* is unknown<sup>29</sup>. Other studies on the presence and sequence of the *fyuA* and *irp2* genes indicate that *Y. pseudotuberculosis* O:1 and O:3 strains belong to a distinct lineage compared with other *Y. pseudotuberculosis* strains and are closer to the *Y. pestis* lineage<sup>38</sup>.

Analysis of the *Y. pestis* sequence reveals a genome that has undergone a much more severe genetic flux than either *Y. pseudotuberculosis* or *Y. enterocolitica*. It has gained several loci by lateral gene transfer, but has also translocated large regions of its chromosome and now seems to be in the early stages of genome decay. As discussed, the mode of transmission of *Y. pestis* can be at least partly attributed to the acquisition of two plasmids (pPla and pMT1) and the *hms* locus. The first crucial step in *Y. pestis* evolution might have been the acquisition of pMT1 by an O:1 or O:3 strain of *Y. pseudotuberculosis* (FIG. 4). A serendipitous discovery during the sequencing of the genome of the recently isolated multidrug-resistant *Salmonella typhi* strain CT18 from Vietnam might shed light on the possible origin of this plasmid. The 96.2-kb pMT1 plasmid in *Y. pestis* has more than 50% sequence identity to the cryptic plasmid pHCM2 that is found in the *S. typhi* strain CT18 (REF. 42). Such high-sequence identity indicates the recent transfer of the plasmid, perhaps between *S. typhi* and *Y. pseudotuberculosis* and/or *Y. pestis* in a dually infected human host, or the gut of a flea vector that fed on multiple hosts. Given that *Salmonella* and *Yersinia* are gut pathogens, one possible evolutionary scenario is that *Salmonella* was the donor of this replicon (pMT1) in the gut of a rodent. Once the ancestral *Y. pseudotuberculosis* strain acquired pMT1, the combination of chromosomally encoded Hms proteins might have allowed for more efficient colonization of

insects and/or fleas<sup>5</sup>. Subsequent acquisition of pPla might then have enhanced the ability to disseminate after transmission to a mammalian host. Ensuing microevolution resulted in the three lineages given the biovar designations, Antiqua, Mediaevalis and Orientalis. Orientalis biovar is glycerol negative and nitrate positive, and Mediaevalis is glycerol positive and nitrate negative, which indicates that these biovars arose independently from the glycerol- and nitrate-positive Antiqua progenitor group<sup>9</sup>. Indeed, further analysis using subtractive hybridization has confirmed that Mediaevalis and Orientalis evolved independently from Antiqua<sup>43</sup>.

However, the presence of pPla and pMT1 is not sufficient to account for the extraordinary virulence of *Y. pestis*<sup>44,45</sup>. This seems to be due to lateral gene transfer into the *Y. pestis* chromosome. At least 21 regions of apparent lateral gene transfer have been identified by variation in G+C content within the CO92 genome. So when did *Y. pestis* acquire them? The answer is perhaps unexpected. In virtually all cases, comparative microarray hybridization analysis of dozens of *Y. pseudotuberculosis* and *Y. pestis* strains indicate that similar sequences are also present in the *Y. pseudotuberculosis* genome<sup>46</sup>, which indicates that these sequences were acquired by *Y. pseudotuberculosis* before *Y. pestis* started to diverge. For example, the sequences that might be related to the parasitism of insects (insecticidal toxin and baculovirus enhancin) are found in a wide range of *Y. pseudotuberculosis* strains. These sequences appear as pseudogenes in *Y. pestis*, but in *Y. pseudotuberculosis* they seem active, indicating that *Y. pseudotuberculosis* — as well as infecting mammals — is also widely found in soil and might have evolved the ability to kill insects, perhaps to gain access to nutrients. Rather than the adaptation of *Y. pestis* to the flea gut occurring in a single evolutionary event, *Y. pseudotuberculosis* might already have been associated with insect hosts for some time.

Recently, it has been found that all *Y. pestis* strains and about 20% of *Y. pseudotuberculosis* strains that have been tested so far form biofilms on the mouths of the nematode *Caenorhabditis elegans*, thereby preventing the worms from feeding<sup>47,48</sup>; however, *Y. enterocolitica* strains are unable to do this. It has been postulated that the ability to form biofilms on biotic surfaces might also have evolved originally in *Y. pseudotuberculosis* as a mechanism to prevent predation by nematodes that feed on bacteria<sup>47</sup>. The *hms* gene has been shown to be required for *Yersinia* biofilm formation in nematodes as well as in fleas<sup>47</sup>. A distinct phenotype of the *hms* locus is autoagglutination — that is, the tendency for *Y. pestis* cells to form clumps in liquid media. In fleas, this 'sticky' attribute is necessary to adhere to the cuticle-covered, proventricular spines, allowing the formation of a dense aggregate that is embedded in an extracellular matrix, reminiscent of a biofilm. So, once *Y. pseudotuberculosis* and/or *Y. pestis* extended its host range into insects, biofilm formation might have assumed another role — that is, blocking the foregut in fleas. This would have been a key step in allowing the initial transfer of the *Y. pestis* bacteria into mammals through a flea bite.

## Conclusions

Common themes to emerge from the genome analysis of more than 50 microbial pathogens that have been sequenced so far, include extensive lateral gene transfer (particularly among enteric pathogens), genome decay (among obligate intracellular pathogens) and extensive antigenic variation by gene shuffling or slipped-strand mispairing<sup>49</sup>. *Y. pestis* has all of these characteristics. It is an organism in an intermediate stage of genetic flux, in which the acquisition of novel sequences by lateral gene transfer seems to be counterbalanced by ongoing genome decay. Perhaps the most striking aspect of the evolution of *Yersinia* is the extremely rapid emergence of *Y. pestis* from *Y. pseudotuberculosis* and genome analysis shows us how this has happened. In terms of gene expansion, apart from the acquisition of pMT1 and pPla, there seems to be little difference between the two species. In other words, *Y. pseudotuberculosis* already has all the extra genes that *Y. pestis* needs for virulence. The key process that turned *Y. pseudotuberculosis* into *Y. pestis* seems instead to have been gene loss — for example, of the insect toxins that would have killed the insect host, and certain physiological functions that accentuate *Y. pestis* virulence in humans. This loss seems to have been triggered by the extensive expansion of IS elements, which caused significant genome rearrangements. Once *Y. pseudotuberculosis* had acquired certain crucial genes, the instability that was introduced by the IS elements was the main force to release its virulence potential — as *Y. pestis*.

Several questions remain regarding the pathogenesis of these *Yersinia* species and their evolution. For example, little is known about the expressed determinants that are important during pneumonic transfer of *Y. pestis*. Presumably, these are inactive in *Y. pseudotuberculosis*, which is incapable of causing infection by this route. Similarly, little is known about the relationship of *Y. pseudotuberculosis* and insects and/or fleas in nature. Why is *Y. pestis* exceptionally pathogenic compared with its recent relative *Y. pseudotuberculosis*? Perhaps the answer lies in the mode of transmission. For an enteropathogen, the most efficient method to ensure transmission to a new host is to cause diarrhoea. By contrast, as *Y. pestis* has to spread to the blood of a new host through a flea vector, the more severe the bacteraemia, the greater the chances of being transmitted through a flea bite. So, there is a strong selective pressure to cause severe disease. The factors influencing the rise and fall of plague pandemics also remain obscure. There will undoubtedly be many factors involved, but the genetic make-up of *Y. pestis* is likely to be important. It is possible that during the spread of an epidemic, passage through humans could allow *Y. pestis* to become more transmissible or more pathogenic, particularly during pneumonic transfer in which close human contact could aid transmission. Such a 'hypervirulent' strain of *Vibrio cholerae* was recently shown to have arisen during passage through humans in a cholera epidemic<sup>50</sup>. The flexible genome of *Y. pestis* makes it a likely candidate for such a mechanism.

Finally, what is the likely fate of *Y. pestis*? Has it finished altering its enteric genome content? Will its flexibility allow the bacterium to cause future pandemics? Or will continued genome decay see the species burn itself out in an evolutionary 'dead-end'? Irrespective

of the natural scenario for *Y. pestis*, it will remain an important organism to investigate both in terms of the development of measures to counteract the threat of bioterrorism, and as a model to study the evolution of pathogens that threaten mankind.

- Perry, R. D. & Fetherston, J. D. *Yersinia pestis*—etiologic agent of plague. *Clin. Microbiol. Rev.* **10**, 35–66 (1997). **Thorough account of the history, epidemiology, ecology, genetics and pathology of plague.**
- Naktin, J. & Beavis, K. G. *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. *Clin. Lab. Med.* **19**, 523–536 (1999).
- Brubaker, R. R. *Yersinia pestis* and bubonic plague, in *The prokaryotes, an evolving electronic resource for the microbiological community*, Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.H. & Stackebrandt, E. (eds), [online], (cited 31 Jul. 2003), <http://et.springer-ny.com:8080/prokPUB/index.htm> (2000).
- Carniel, E. *et al.* 2002. Enteropathogenic yersiniae, in *The prokaryotes, an evolving electronic resource for the microbiological community*, Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.H. & Stackebrandt, E. (eds), [online], (cited 31 Jul. 2003), <http://et.springer-ny.com:8080/prokPUB/index.htm> (2002).
- 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). **Use of molecular clock (MLST) analysis to show that *Y. pestis* evolved from *Y. pseudotuberculosis* between the last 1,500 and 20,000 years.**
- Gallmand, M. *et al.* Multidrug resistance in *Yersinia pestis* mediated by a transferable plasmid. *N. Engl. J. Med.* **337**, 677–680 (1997).
- Bacot, A. W. & Martin, C. J. Observations of the mechanisms of transmission of plague by fleas. *J. Hygiene* **13**, 423–439 (1914).
- Lahteenmaki, K., Virkola, R., Saren, A., Emody, L. & Korhonen, T. K. Expression of plasminogen activator pla of *Yersinia pestis* enhances bacterial attachment to the mammalian extracellular matrix. *Infect. Immun.* **66**, 5755–5762 (1998).
- Devignat, R. *Bulletin WHO* **4**, 247–263 (1951).
- Parkhill, J. *et al.* Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* **413**, 523–527 (2001). **First whole-genome sequence of a member of the *Yersinia*. It shows that the *Y. pestis* genome is highly fluid with evidence of recent acquisition, mixing of the genome and early stages of genome decay.**
- Deng, W. *et al.* Genome sequence of *Yersinia pestis* KIM. *J. Bacteriol.* **184**, 4601–4611 (2002). **Genome sequence of the second *Y. pestis* Mediaevalis KIM10 strain, providing the opportunity to compare the genome content of a modern-day plague strain to a more ancient strain.**
- 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).
- Ibrahim, A., Goebel, B. M., Liesack, W., Griffiths, M. & Stackebrandt, E. The phylogeny of the genus *Yersinia* based on 16S rDNA sequences. *FEMS Microbiol. Lett.* **114**, 173–177 (1993).
- Brubaker, R. R. Factors promoting acute and chronic diseases caused by yersiniae. *Clin. Microbiol. Rev.* **4**, 309–324 (1991).
- Sodeinde, O. A., Sample, A. K., Brubaker, R. R. & Goguen, J. D. Plasminogen activator/coagulase gene of *Yersinia pestis* is responsible for degradation of plasmid-encoded outer membrane proteins. *Infect. Immun.* **56**, 2749–2752 (1988).
- Davis, K. J. *et al.* Pathogenicity of experimental pneumonic plague produced by fraction 1-positive and fraction 1-negative *Yersinia pestis* in African Green monkeys. *Arch. Pathol. Lab. Med.* **120**, 156–163 (1996).
- Hinnebusch, B. J., Fischer, E. R. & Schwan, T. G. Evaluation of the role of the *Yersinia pestis* plasminogen activator and other plasmid-encoded factors in temperature-dependent blockage of the flea. *J. Infect. Dis.* **178**, 1406–1415 (1998).
- 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). **Shows a clear role for the so called 'murine toxin' on the acquired plasmid pMT1 as a phospholipase that is important for the survival of *Y. pestis* in the flea midgut.**
- 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).
- Waterfield, N. R., Bowen, D. J., Fetherston, J. D., Perry, R. D. & French-Constant, R. H. The *tc* genes of *Photobacterium*: a growing family. *Trends Microbiol.* **9**, 185–191 (2001).
- 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).
- Lobry, J. R. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* **13**, 660–665 (1996).
- Andersson, S. G. *et al.* The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**, 133–140 (1998).
- Brosch, R., Gordon, S. V., Eiglmeier, K., Garnier, T. & Cole, S. T. Comparative genomics of the leprosy and tubercle bacilli. *Res. Microbiol.* **151**, 135–142 (2000).
- Simonet, M., Riot, B., Fortineau, N. & Berche, P. Invasin production by *Yersinia pestis* is abolished by insertion of an IS200-like element within the *inv* gene. *Infect. Immun.* **64**, 375–379 (1996).
- Rosqvist, R., Skurnik, M. & Wolf-Watz, H. Increased virulence of *Yersinia pseudotuberculosis* by two independent mutations. *Nature* **334**, 522–524 (1988).
- Henderson, I. R., Owen, P. & Nataro, J. P. Molecular switches—the ON and OFF of bacterial phase variation. *Mol. Microbiol.* **33**, 919–932 (1999).
- 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).
- Skurnik, M., Peippo, A. & Ervela, E. Characterization of the O-antigen gene clusters of *Yersinia pseudotuberculosis* and the cryptic O-antigen gene cluster of *Yersinia pestis* shows that the plague bacillus is most closely related to and has evolved from *Y. pseudotuberculosis* serotype O:1b. *Mol. Microbiol.* **37**, 316–330 (2000).
- Maurelli, A. T., Fernandez, R. E., Bloch, C. A., Rode, C. K. & Fasano, A. 'Black holes' and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **95**, 3943–3948 (1998).
- Nakata, N. *et al.* The absence of a surface protease, OmpT, determines the intercellular spreading ability of *Shigella*: the relationship between the *ompT* and *kcpA* loci. *Mol. Microbiol.* **9**, 459–468 (1993).
- Parish, T. *et al.* Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis*. *Infect. Immun.* **71**, 1134–1140 (2003).
- Dreyfus, L. A. & Brubaker, R. R. Consequences of aspartase deficiency in *Yersinia pestis*. *J. Bacteriol.* **136**, 757–764 (1978).
- Mortlock, R. P. & Brubaker, R. R. Glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase activities of *Pasteurella pestis* and *Pasteurella pseudotuberculosis*. *J. Bacteriol.* **84**, 1122–1123 (1962).
- Garcia, E. Investigating the evolutionary relationship between *Yersinia pseudotuberculosis* and *Yersinia pestis* through whole-genome comparative sequencing. *Eighth International Symposium on Yersinia 16* (Turku, Finland, 2002).
- Parkhill, J. & Thompson, N. R. Genome sequence and comparison of *Yersinia pestis* and *Yersinia pseudotuberculosis*. *Eighth International Symposium on Yersinia 15* (Turku, Finland, 2002).
- Delor, I. & Cornelis, G. R. Role of *Yersinia enterocolitica* Yst toxin in experimental infection of young rabbits. *Infect. Immun.* **60**, 4269–4277 (1992).
- Pelludat, C., Rakin, A., Jacobi, C. A., Schubert, S. & Heesemann, J. The yersiniabactin biosynthetic gene cluster of *Yersinia enterocolitica*: organization and siderophore-dependent regulation. *J. Bacteriol.* **180**, 538–546 (1998).
- Pelludat, C., Hogardt, M. & Heesemann, J. Transfer of the core region genes of the *Yersinia enterocolitica* WA-C serotype O:8 high-pathogenicity island to *Y. enterocolitica* MRS40, a strain with low levels of pathogenicity, confers a yersiniabactin biosynthesis phenotype and enhanced mouse virulence. *Infect. Immun.* **70**, 1832–1841 (2002).
- Schubert, S., Rakin, A., Karch, H., Carniel, E. & Heesemann, J. Prevalence of the 'high-pathogenicity island' of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. *Infect. Immun.* **66**, 480–485 (1998).
- Iwobi, A. *et al.* Novel virulence-associated type II secretion system unique to high-pathogenicity *Yersinia enterocolitica*. *Infect. Immun.* **71**, 1872–1879 (2003).
- Prentice, M. B. *et al.* *Yersinia pestis* pFra shows biovar-specific differences and recent common ancestry with a *Salmonella enterica* serovar *Typhi* plasmid. *J. Bacteriol.* **183**, 2586–2594 (2001).
- Radnedge, L., Agron, P. G., Worsham, P. L. & Andersen, G. L. Genome plasticity in *Yersinia pestis*. *Microbiology* **148**, 1687–1698 (2002).
- Kutyrev, V. *et al.* Expression of the plague plasmidogen activator in *Yersinia pseudotuberculosis* and *Escherichia coli*. *Infect. Immun.* **67**, 1359–1367 (1999).
- Welkos, S. L., Davis, K. M., Pitt, L. M., Worsham, P. L. & Freidlander, A. M. Studies on the contribution of the F1 capsule-associated plasmid pFra to the virulence of *Yersinia pestis*. *Contrib. Microbiol. Immunol.* **13**, 299–305 (1995).
- Hinchcliffe, S. J. *et al.* Application of DNA microarrays to study the evolutionary genomics of *Yersinia pestis* and *Yersinia pseudotuberculosis*. *Genome Res.* (in the press).
- Darby, S., Hsu, J. W., Ghori, N. & Falkow, S. *Caenorhabditis elegans*: plague bacteria biofilm blocks food intake. *Nature* **417**, 243–244 (2002).
- Joshua, G. W. P. *et al.* A *Caenorhabditis elegans* model of *Yersinia* infection, biofilm formation on biotic surfaces. *Microbiology* (in the press).
- Wren, B. W. Microbial genome analysis: insights into virulence, host adaptation and evolution. *Nature Rev. Genet.* **1**, 30–39 (2000).
- Merrell, D. S. *et al.* Host-induced epidemic spread of the cholera bacterium. *Nature* **417**, 642–645 (2002).

Acknowledgements

Work in the author's laboratory is supported by the Biotechnology and Biological Sciences Research Council and the Defence Science and Technology Laboratory. I acknowledge J. Marchant and S. Hinchcliffe for critical reviews of the manuscript and anonymous reviewers for helpful comments.

Online links

DATABASES

The following terms in this article are linked online to: **Entrez**: <http://www.ncbi.nlm.nih.gov/Entrez/Y.pestis/C092> | **Y. pestis** KIM10 **Swiss-Prot**: <http://www.expasy.ch/Hms|Ymt>

FURTHER INFORMATION

**Brendan W. Wren's laboratory**: <http://www.lshim.ac.uk/pmbu/staff/bwren/bwren.html>  
**Y. enterocolitica** genome sequence: [http://www.sanger.ac.uk/Projects/Y\\_enterocolitica/](http://www.sanger.ac.uk/Projects/Y_enterocolitica/)  
**Y. pseudotuberculosis** genome sequence: <http://bbpr.llnl.gov/bbpr/html/microbe.html>  
 Access to this interactive links box is free online.