The Evolution of Infectious Disease

Why are some bacteria pathogenic to humans while other (closely-related) bacteria are not?

This question can be approached from two directions:

1. *From the point of view of the host.* What specific defense mechanisms of the host allow it to suppress infection (entry, attachment, invasion, replication) by certain pathogens and not others?
2. *From the point of view of the pathogen*. What are the differences between the agents that cause disease and those that do not?

Genomic insights into bacterial pathogenesis

What features enable certain bacteria to be pathogens?

How might it be possible to identify the particular gene or genes (termed “virulence factors” or “pathogenicity determinants”) that distinguish pathogenic from non-pathogenic bacteria.

Can these features be recognized by inspecting genome sequences?

The majority of sequencing projects have been directed towards determining the full genome sequences of bacterial pathogens, with the goal of identifying and understanding the genetic basis of pathogenicity and virulence.

Most research focuses on enteric bacteria

What are enteric bacteria?

The enterics (or the Enterobactericaea) form a group of related bacteria that were known to reside in, and were first isolated from, the mammalian intestine.

Why study enteric bacteria?

Enterics have been used as the model organism for bacterial genetics, allowing the experimental manipulation of their genomes to determine the gene function.

Enterics comprise species of widely different lifestyles and pathogenic potentials, allowing the comparisons of closely-related but ecologically distinct genomes.

Which bacteria are classified as enterics?

Escherichia - benign *E. coli* K-12 used in bacterial genetics; a normal constituent of intestinal flora; *some food-borne pathogens* (O157:H7)

Klebisiella - found in soil; some cause respiratory & other infections

Salmonella - causes typhoid fever, food poisoning, gastroenteritis; can be used as a bioweapon

Shigella - cause of bacillary dysentery; can be used as a bioweapon

Erwinia - a pathogen of plants that causes fireblight in pear and apple trees and soft rot of carrots and potatoes

Yersinia - found in soil, and as insect-borne pathogen of mammals, *e.g.*, *Y. pestis* causes bubonic plague

Proteus - found in soil; common saprophyte of decaying organic matter

What sort of genetic differences might lead to differences in pathogenic potential?

- Allelic differences in genes common to enteric bacteria
- Regulatory differences in genes common to enteric bacteria
- Absence of a virulence repressor in the pathogen
- Presence of pathogen-specific virulence determinants.

How is possible to identify the genes responsible for bacterial virulence?

1. Identify genes which, when knocked out, attenuate virulence

Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 6383–6387, August 1989
Microbiology

Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells

(cell invasion/bacterial pathogenesis/bacterial adhesion/Tn*phoA*)

JORGE E. GALÁN* AND ROY CURTISS III

ABSTRACT **Invasion of the intestinal epithelium is thought to be an important step in the pathogenesis of *Salmonella* infections. Using an *in vitro* system, we have isolated a genetic locus, *inv*, that confers to a noninvasive strain of *Salmonella typhimurium* the ability to penetrate tissue culture cells. Highly virulent *S. typhimurium* strains carrying *inv* mutations were defective for entry into Henle-407 cells while remaining unaffected in their ability to attach to cultured cells. When administered perorally to BALB/c mice, *inv* mutants of *S. typhimurium* had higher 50% lethal doses (LD_{50}) than their wild-type parent strains. To the contrary, there were no differences in the observed LD_{50} when strains were administered intraperitoneally. In addition, *inv* mutants presented decreased ability to colonize the Peyer's patches, the small intestinal wall, and the spleen when administered perorally, although when administered intraperitoneally, they showed no difference in their ability to colonize the spleen compared to the wild-type parent strain.**

How is possible to identify the genes responsible for bacterial virulence?

2. Identify genes that confer virulence properties upon a benign rel

A single genetic locus encoded by *Yersinia pseudotuberculosis* permits invasion of cultured animal cells by *Escherichia coli* K-12

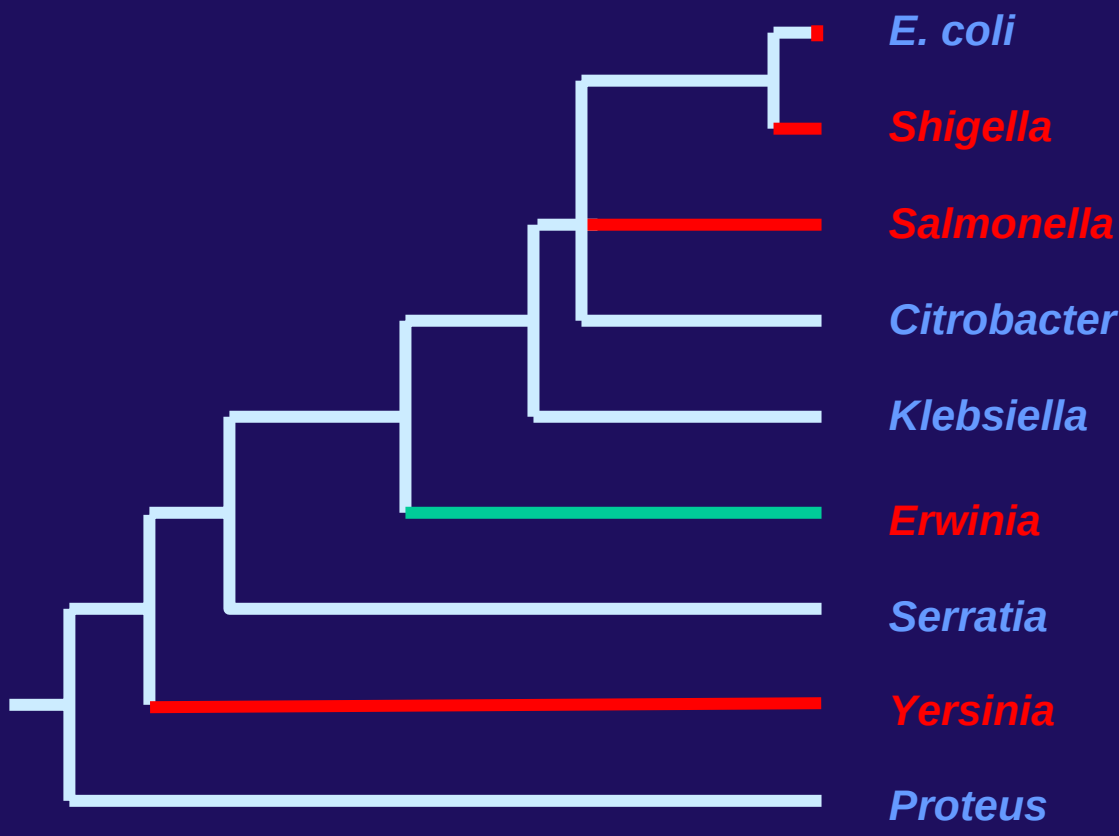
Ralph R. Isberg & Stanley Falkow

Department of Medical Microbiology, Stanford University Medical Center, Stanford, California 94305, USA

For many species of pathogenic bacteria, invasion and survival within animal cells is central to establishing a successful host–parasite relationship[1–3]. Localization within host cells protects the microorganism from host defences[4], or permits it to cross epithelial barriers and subsequently become systemically distributed[5]. The precise mechanisms that permit entry of bacteria into host tissues are unclear[6], therefore we have been studying the invasion of epithelial cells by *Yersinia pseudotuberculosis*[7,8]. As a first step towards identifying the factors required for this process, we report here the identification of a single genetic locus from this organism that is sufficient to convert the innocuous *Escherichia coli* K-12 strain into an organism capable of invading cultured animal cells.

NATURE VOL. 317 19 SEPTEMBER 1985

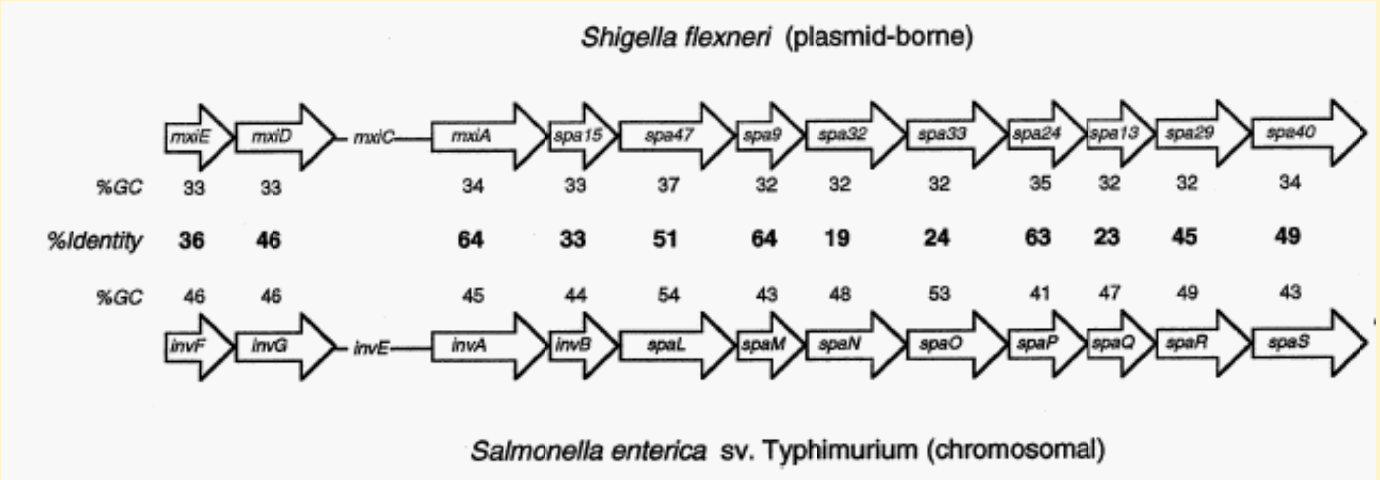
Distribution of Pathogenicity within Enteric Bacteria

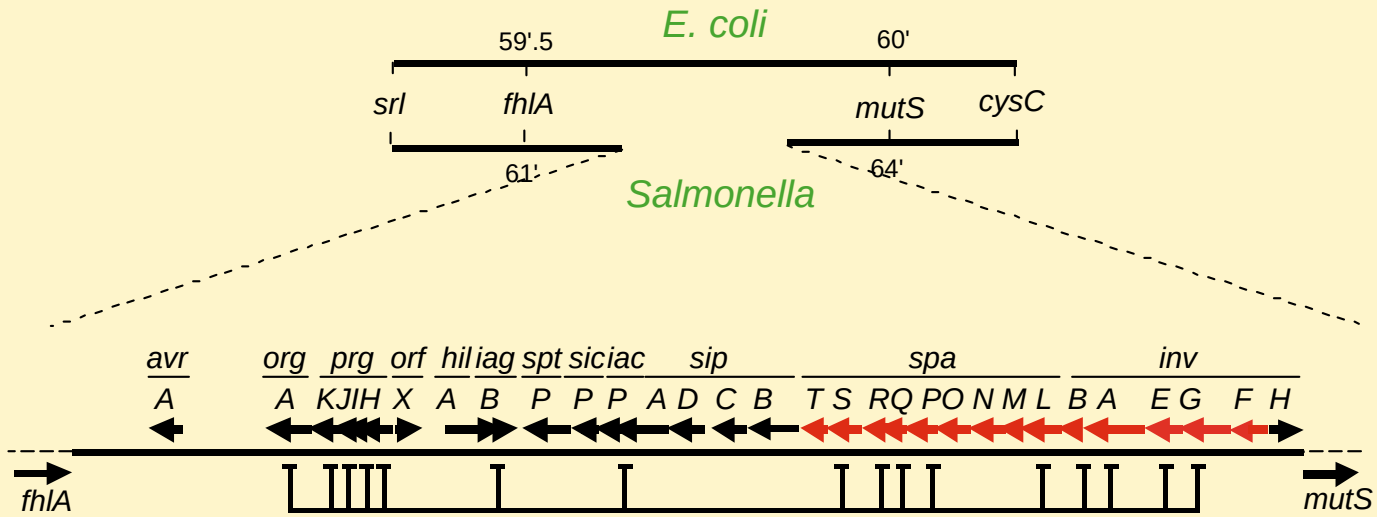


based on this distribution, virulence is the derived state

Pathogens have virulence genes not present in non-pathogenic relatives,
and this distribution suggests that bacteria evolve to become pathogens

A Salmonella-specific Region Required for Virulence





type III secretion system

PATHOGENICITY ISLANDS

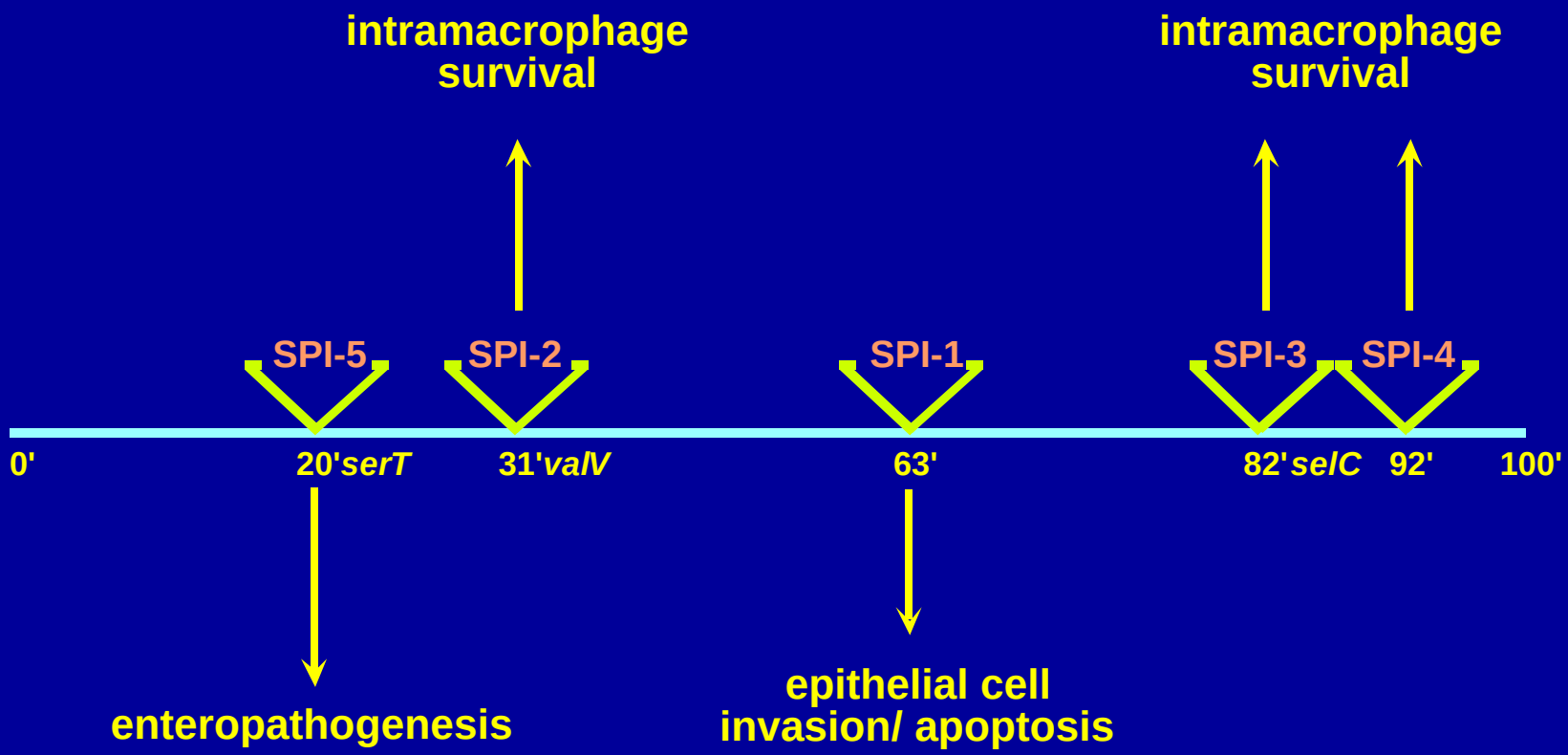
. Segments of the chromosome harboring large clusters of virulence genes

. Present in pathogenic strains but absent or sporadically distributed in related non-pathogenic species

. Typically have a G+C content different from that of the rest of the chromosome

. Often associated with tRNA genes and/or mobile genetic elements at their boundaries

PATHOGENICITY ISLANDS OF SALMONELLA ENTERICA

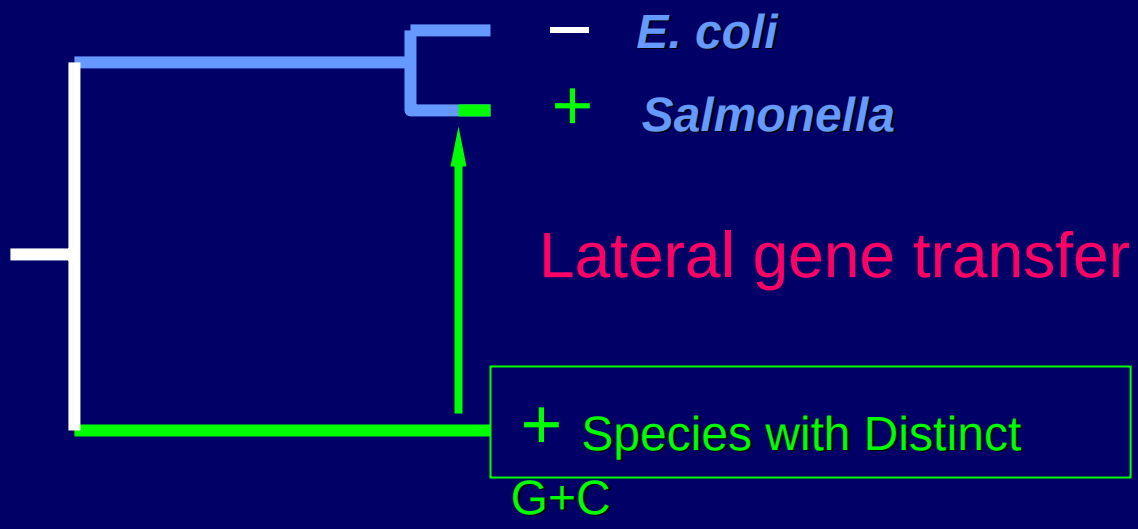


Why do pathogenicity islands have atypical G+C contents?

To understand the significance of this feature, you need to know something about bacterial genomes.

- *Bacterial genomes are tightly packed with genes and other functional elements. Their genomes range from 0.2-10 Mb (~200 to 10,000 genes) and contain very little repetitive, transposable, & non-coding DNA*
- *Base composition (G+C content) is relatively homogeneous over the entire chromosome, such that all genes have about the same overall G+C content*
- *Base composition varies among bacterial species from about 15% to 80% G+C & is similar in closely-related taxa*

Why do pathogenicity islands have atypical G+C contents?

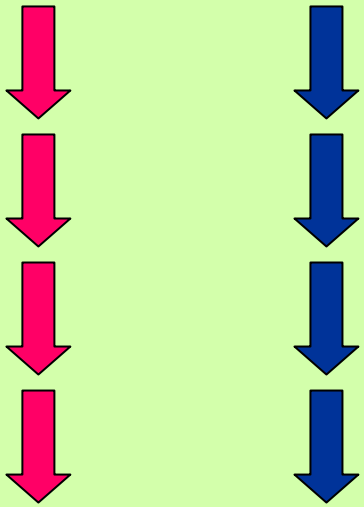


Lateral gene transfer is the source of "atypical" &"species-specific" genes

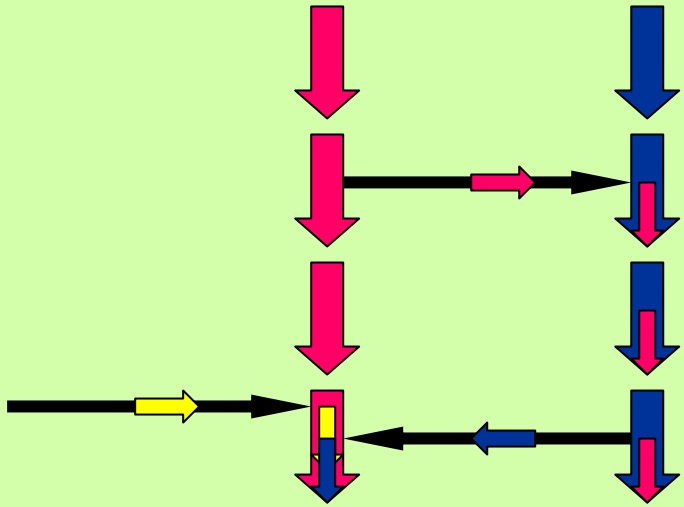
Why is this type of gene evolution considered "lateral"?

Lateral (or horizontal) gene transfer denotes any transfer, exchange or acquisition of genetic material that differs from the normal mode of transmission from parents to offspring (vertical transmission).

Vertical evolution

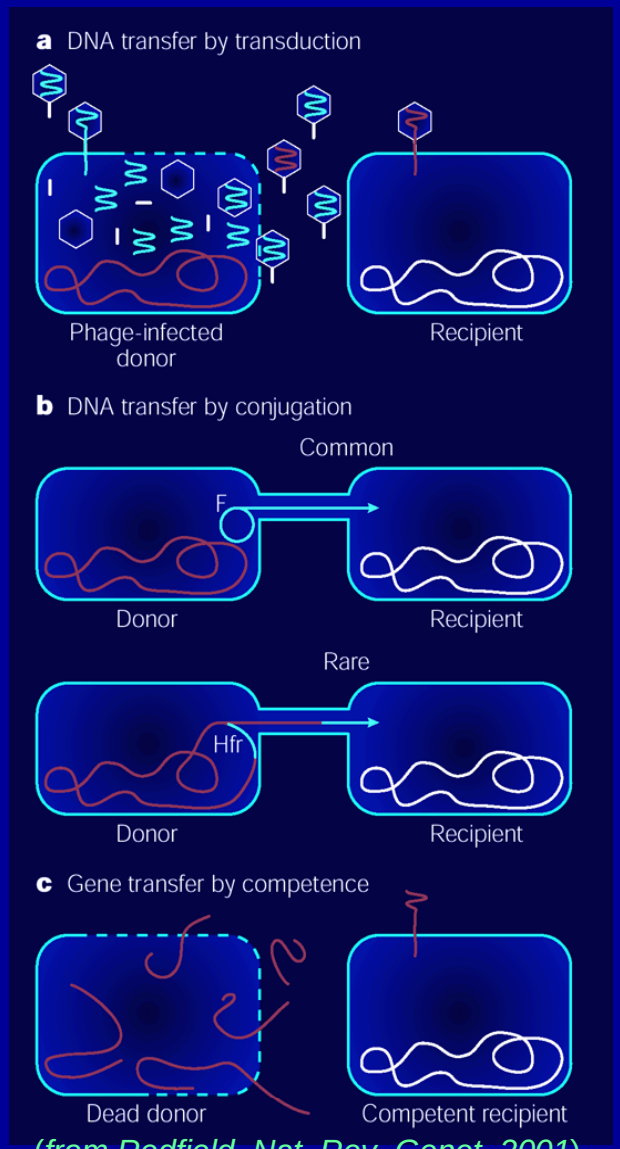


Horizontal evolution



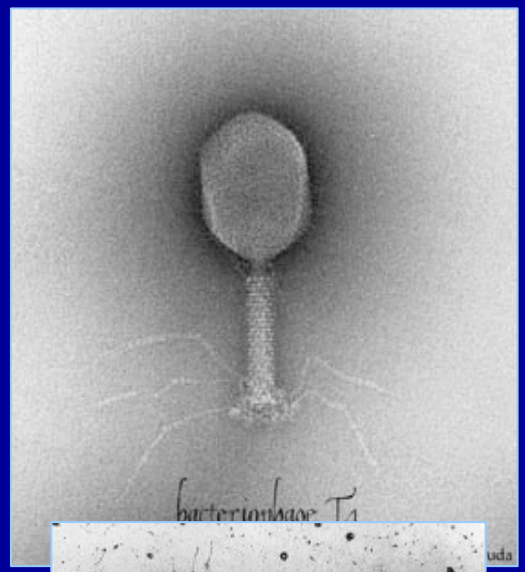
Lateral gene transfer (LGT) can occur by several mechanisms and cause the transfer/acquisition of genes within a genome, among members of the same species, or between members of very different taxa.

How do genes get transferred laterally?

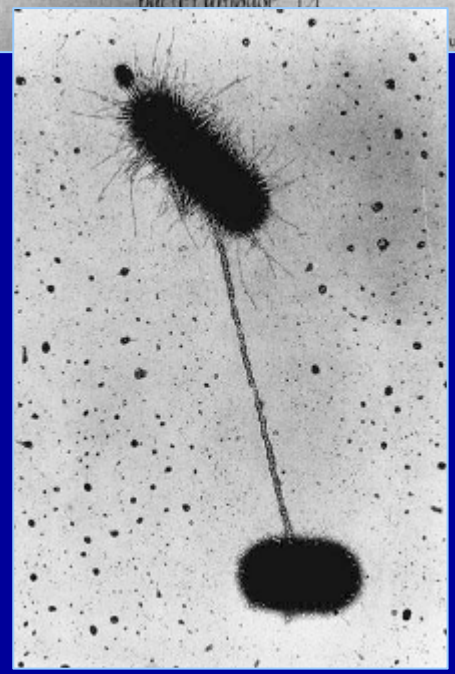


(*from Redfield, Nat. Rev. Genet. 2001*)

Transduction:
via bacteriophage

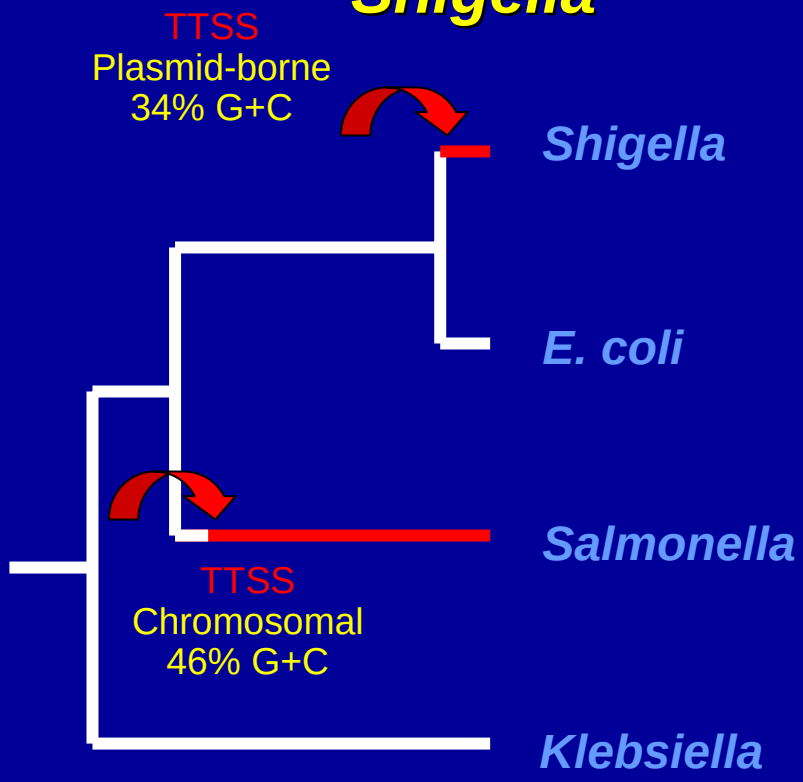


Conjugation:
direct contact



Transformation:
integrating free DNA
or plasmids

The genes for host cell invasion are the same, but were acquired independently by lateral gene transfer, in Salmonella and Shigella



The overall base composition of *E. coli, Shigella* & *Salmonella* is 52% G+C

The role of mobile elements in E. coli virulence

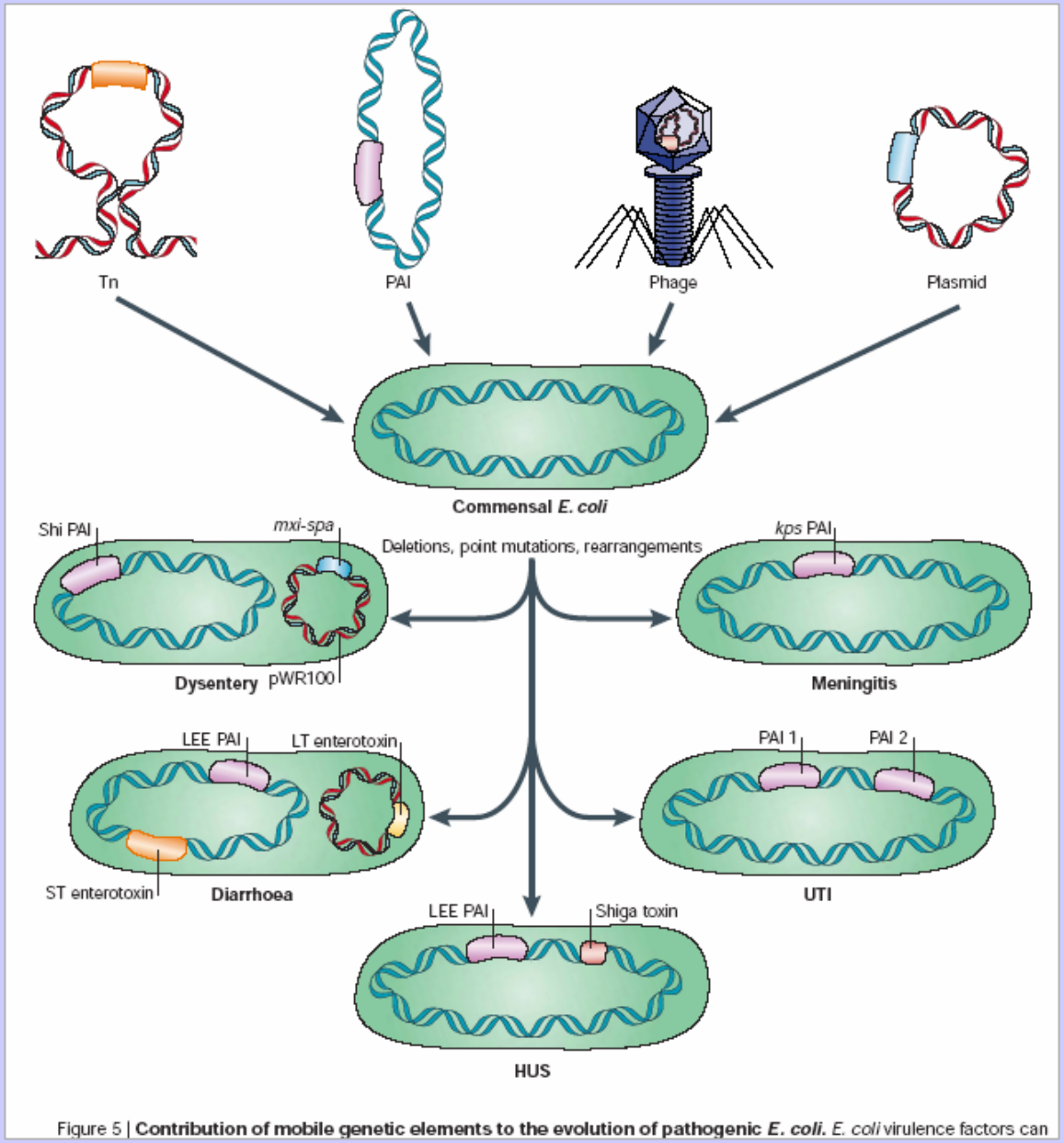


Figure 5 | **Contribution of mobile genetic elements to the evolution of pathogenic *E. coli*.** *E. coli* virulence factors can

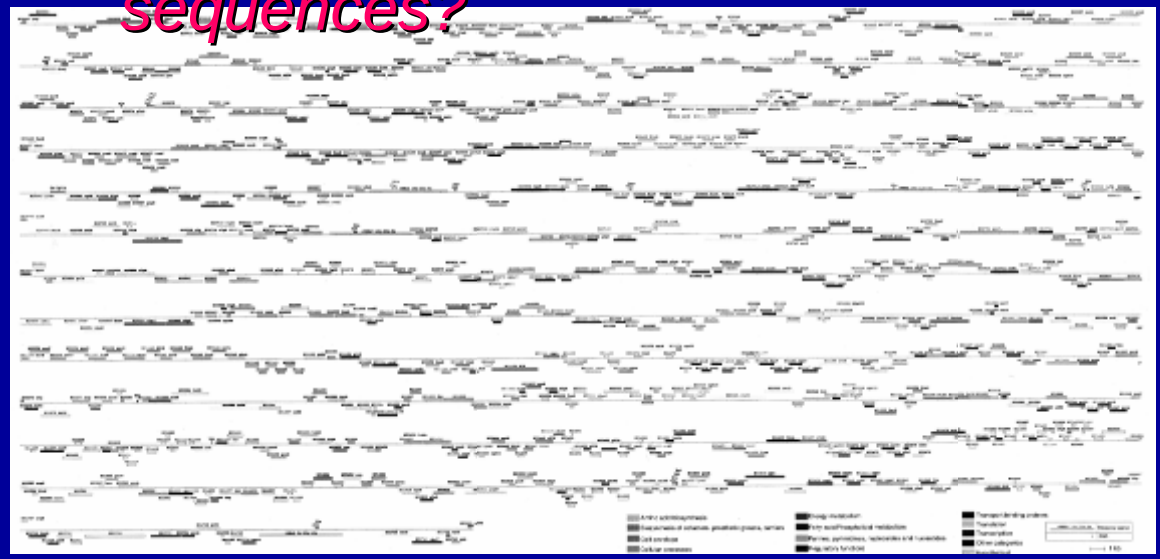
If genes acquired from distant sources by LGT have atypical G+C contents, shouldn't they be evident when examining genome sequences?

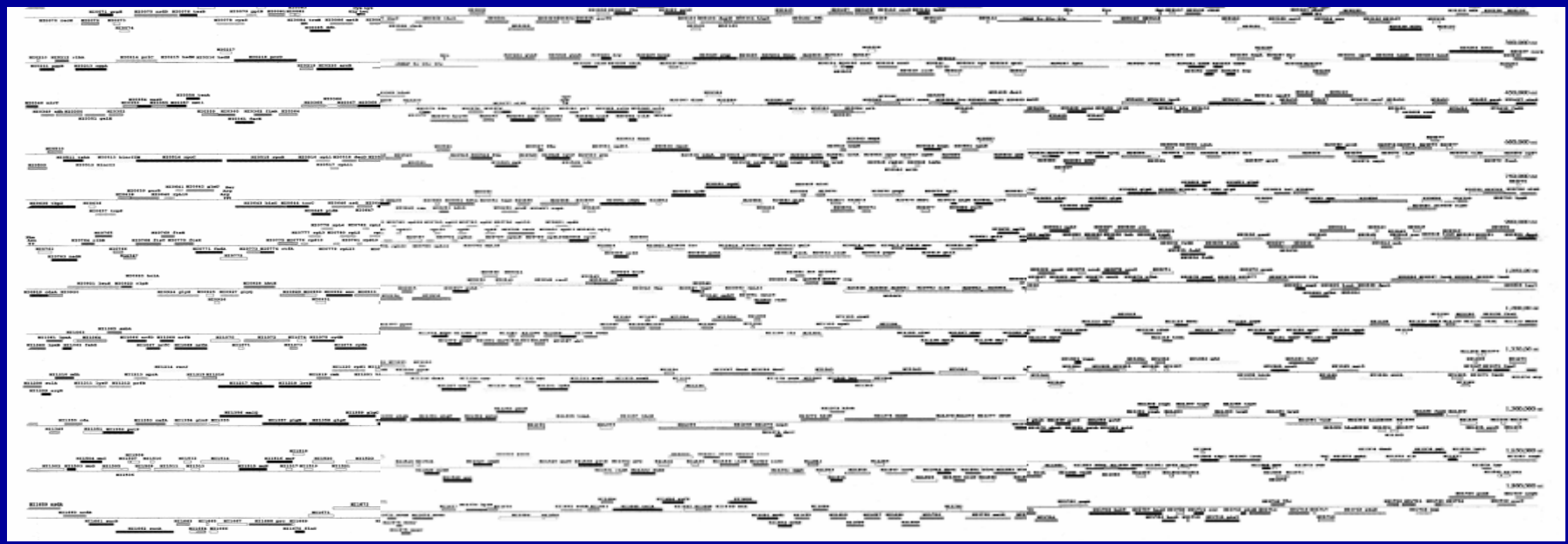
Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd

Robert D. Fleischmann, Mark D. Adams, Owen White, Rebecca A. Clayton, Ewen F. Kirkness, Anthony R. Kerlavage, Carol J. Bult, Jean-Francois Tomb, Brian A. Dougherty, Joseph M. Merrick, Keith McKenney, Granger Sutton, Will FitzHugh, Chris Fields,* Jeannine D. Gocayne, John Scott, Robert Shirley, Li-Ing Liu, Anna Glodek, Jenny M. Kelley, Janice F. Weidman, Cheryl A. Phillips, Tracy Spriggs, Eva Hedblom, Matthew D. Cotton, Teresa R. Utterback, Michael C. Hanna, David T. Nguyen, Deborah M. Saudek, Rhonda C. Brandon, Leah D. Fine, Janice L. Fritchman, Joyce L. Fuhrmann, N. S. M. Geoghagen, Cheryl L. Gnehm, Lisa A. McDonald, Keith V. Small, Claire M. Fraser, Hamilton O. Smith, J. Craig Venter†

An approach for genome analysis based on sequencing and assembly of unselected pieces of DNA from the whole chromosome has been applied to obtain the complete nucleotide sequence (1,830,137 base pairs) of the genome from the bacterium *Haemophilus influenzae* Rd. This approach eliminates the need for initial mapping efforts and is therefore applicable to the vast array of microbial species for which genome maps are unavailable. The *H. influenzae* Rd genome sequence (Genome Sequence DataBase accession number L42023) represents the only complete genome sequence from a free-living organism.

SCIENCE • VOL. 269 • 28 JULY 1995



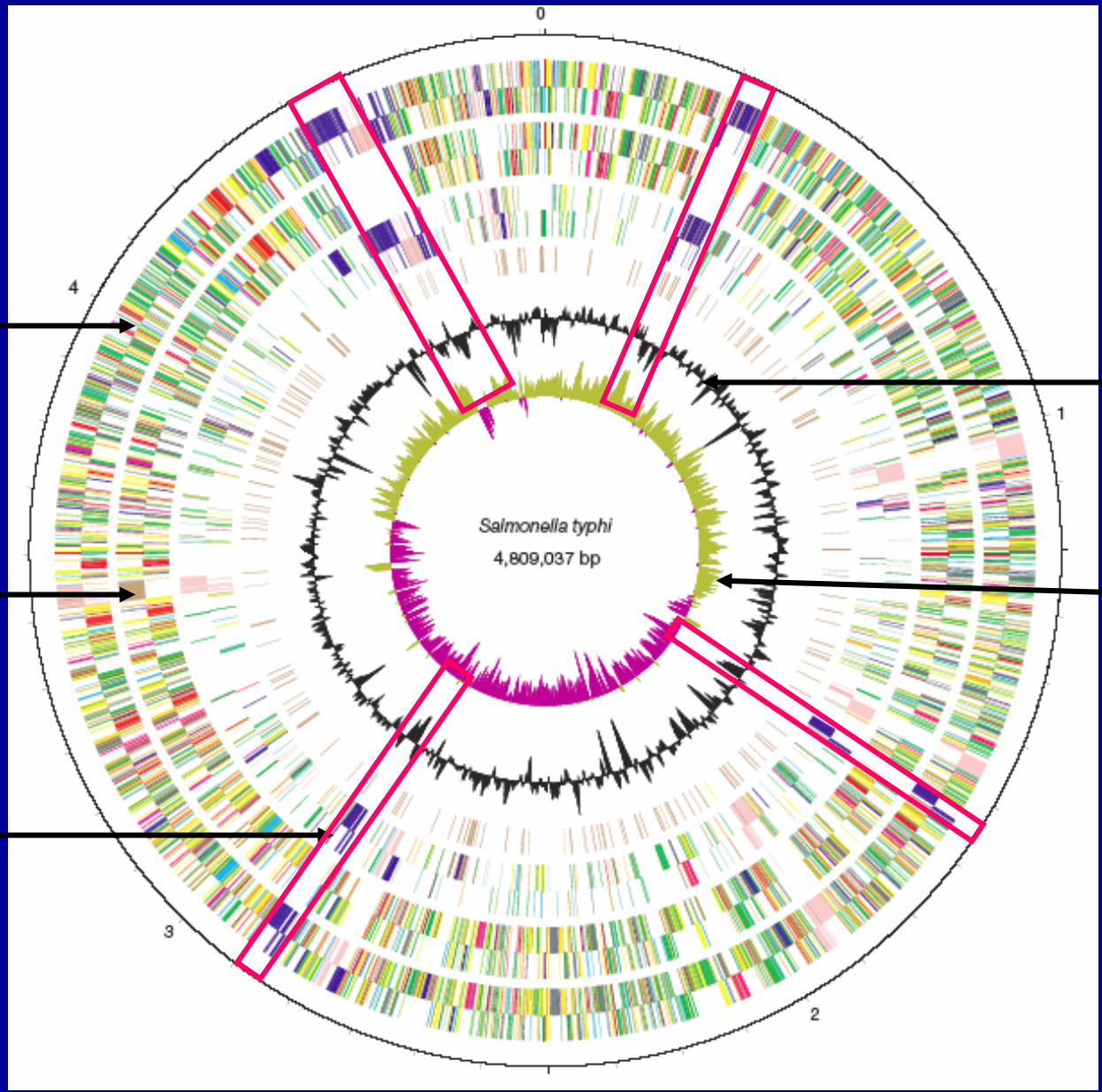


Depicting Bacterial Genome Sequences

Genes coded by location & function

Genes shared with *E. coli*

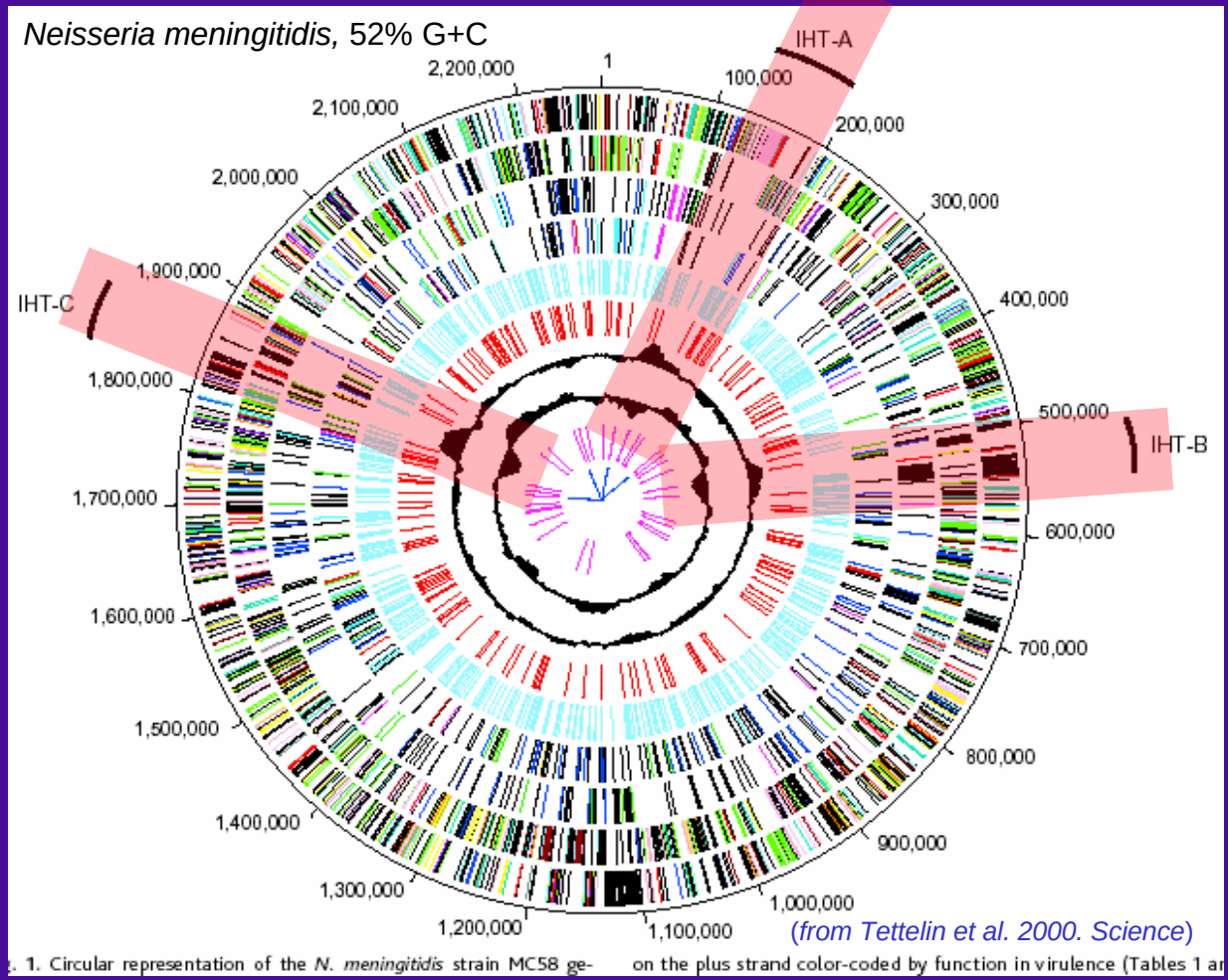
Genes unique to *S. typhi*



%G+C

GC skew (G-C)/G+C)

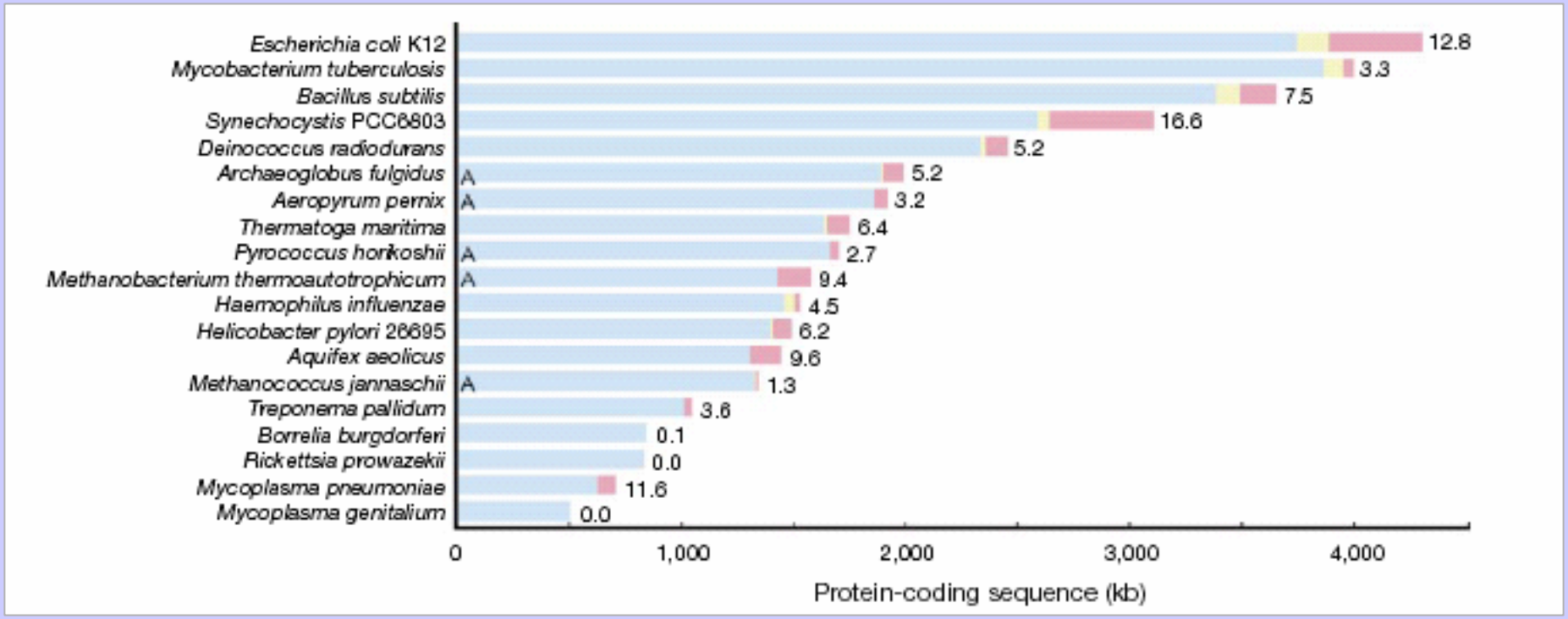
Inferrring lateral gene transfer (LGT) from sequence heterogeneity along the chromosome



(*from Tettelin et al. 2000. Science*)

. 1. Circular representation of the *N. meningitidis* strain MC58 ge- on the plus strand color-coded by function in virulence (Tables 1 ar

Amounts of 'atypical' (transferred) DNA in bacterial genomes



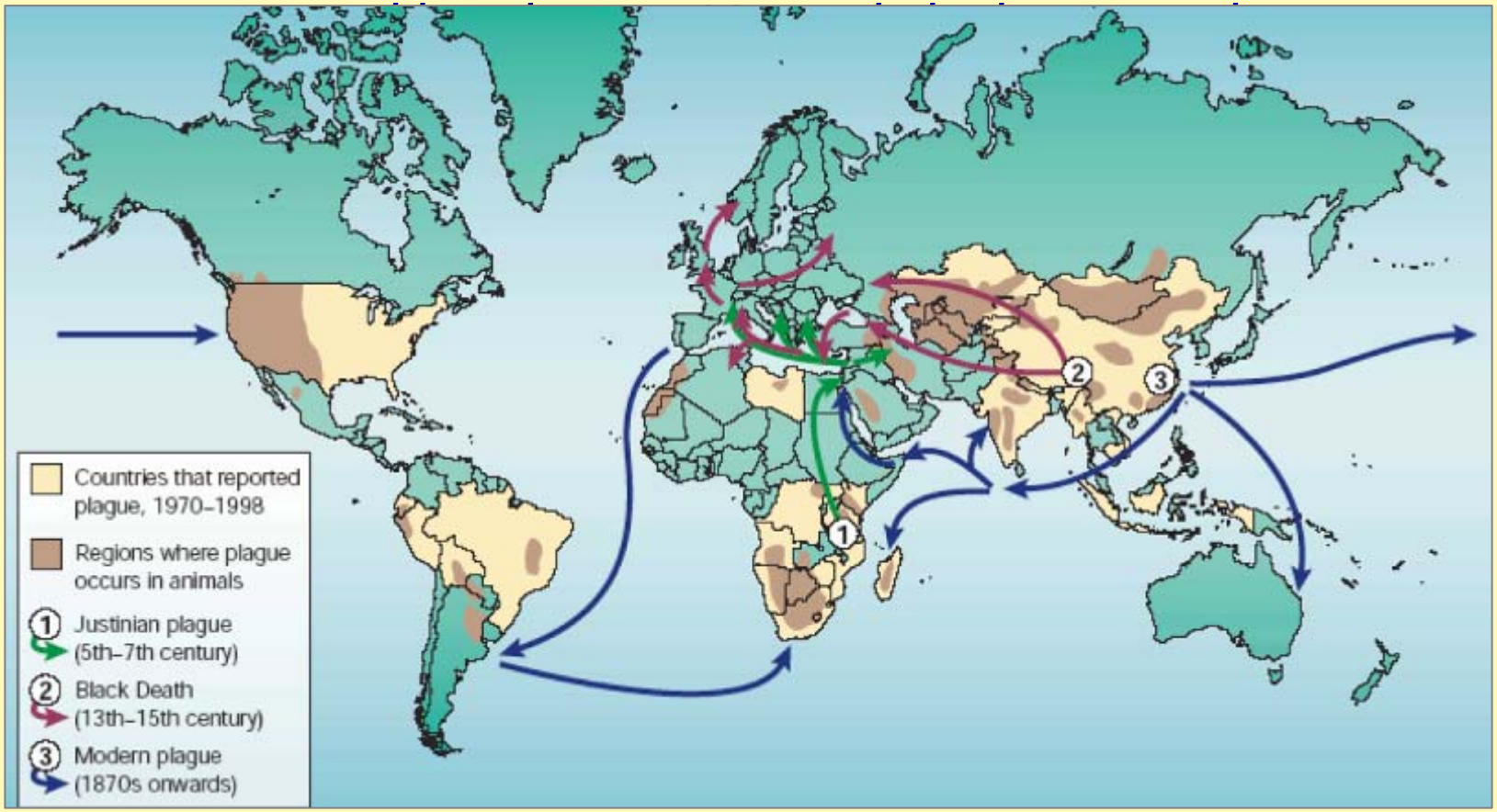
The story so far:

- ***Bacterial genomes are small and densely packed with genes.***
- ***Pathogenic bacteria often contain clusters of genes (PAIs) that are not present in related non-pathogenic bacteria.***
- ***Many of these virulence determinants were acquired by lateral gene transfer***
- ***Acquired genes have several features (G+C contents; association with plasmids or phage; sporadic distributions) that denote their ancestry***
- ***It is possible to recognize genes that arose by lateral gene transfer by simply examining genome sequences.***
- ***The amount of acquired DNA in many bacterial genomes can be substantial.***

Yersinia pestis: Rapid evolution of an enteric pathogen

Three (of the 11) species of Yersinia are pathogenic to humans:

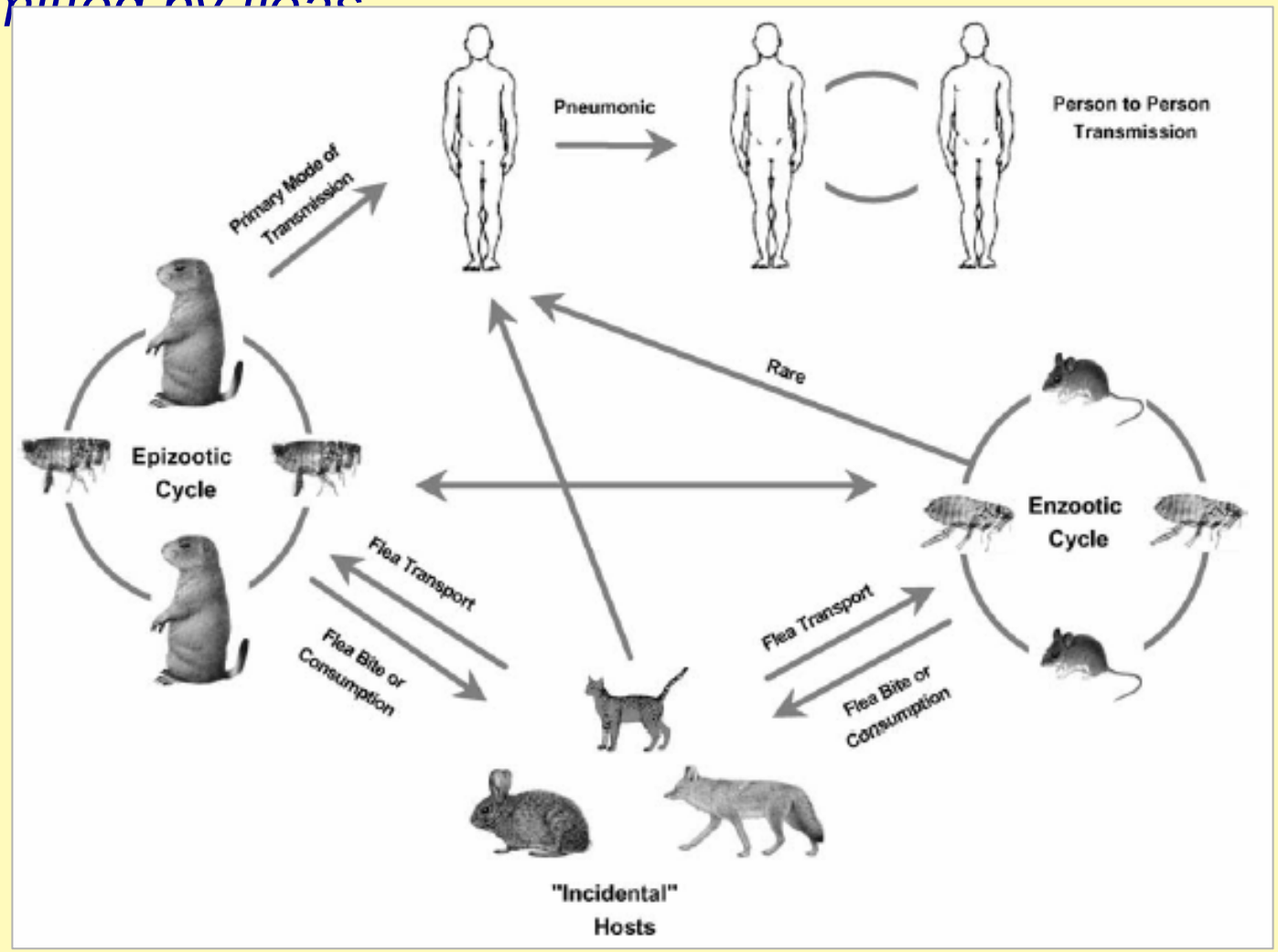
Y. enterocolitica & Y. pseudotuberculosis cause



Three known plague pandemics:

Justinian, 541-767; Black Death, 1346-1800s; Modern 1894-present

Y. pestis is primarily a disease of rodents & is usually transmitted by fleas



… whereas Y. enterocolitica & Y. pseudotuberculosis are food- & wat

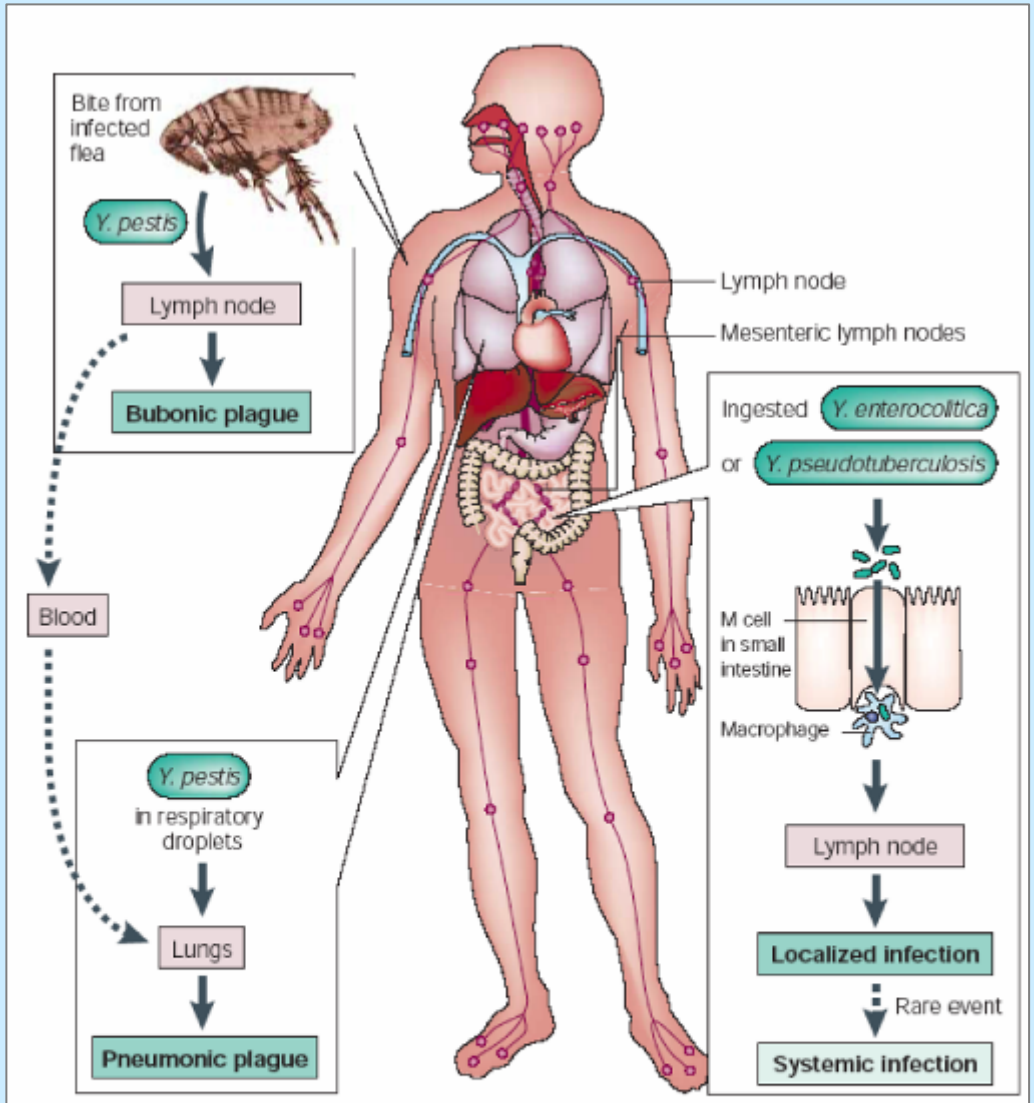


Figure 2 | **Steps in the transmission of the pathogenic yersiniae in humans.** *Y. pestis* has a rodent reservoir. The rodent's fleas acquire *Y. pestis* from a meal of infected blood, and transmit the bacterium primarily to other rodents, or occasionally to humans — causing bubonic plague in humans. 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.

Y. pestis pathogenesis has several unique features including:

1. Mammalian reservoir

- Has enzootic (maintenance, resistant) as well as epizootic (spreading) hosts.

2. Flea-mediated transmission

-

- *hms* product makes bacteria form aggregates that block the foregut of infected flea. (Blocked flea regurgitates infected blood back into bite site.)

- *ymt* locus needed to survive in flea midgut

3. Causes systemic infections

-

- expresses capsular antigen to resist phagocytosis and kill monocytes

Y. pestis evolved from Y. pseudotuberculosis only 2000-20000 years ago

Genome comparisons suggest that the transition from enteric pathogen to flea-borne pathogen involved at least three steps:

1. Plasmid acquisition.

- All three yersinae species harbor a 70-kb virulence plasmid (pYV) needed for toxicity and to overcome host immune system; but there are two *Y. pestis*-specific plasmids that were recently acquired by horizontal gene transfer.
- pPCP1 (9.6 kb) contain plasminogen activator (a surface molecule that provides proteolytic, adhesive and invasive functions) and allows dissemination from intradermal site of infection; also a bacteriocin and an immunity protein.
- pMT1 or pFra (96.2 kb) - capsular antigen (blocks phagocytosis) and murine toxin (Ymt) needed to survive in flea.

2. Acquisition of PAIs and recruitment of endogenous chromosomal genes for ne

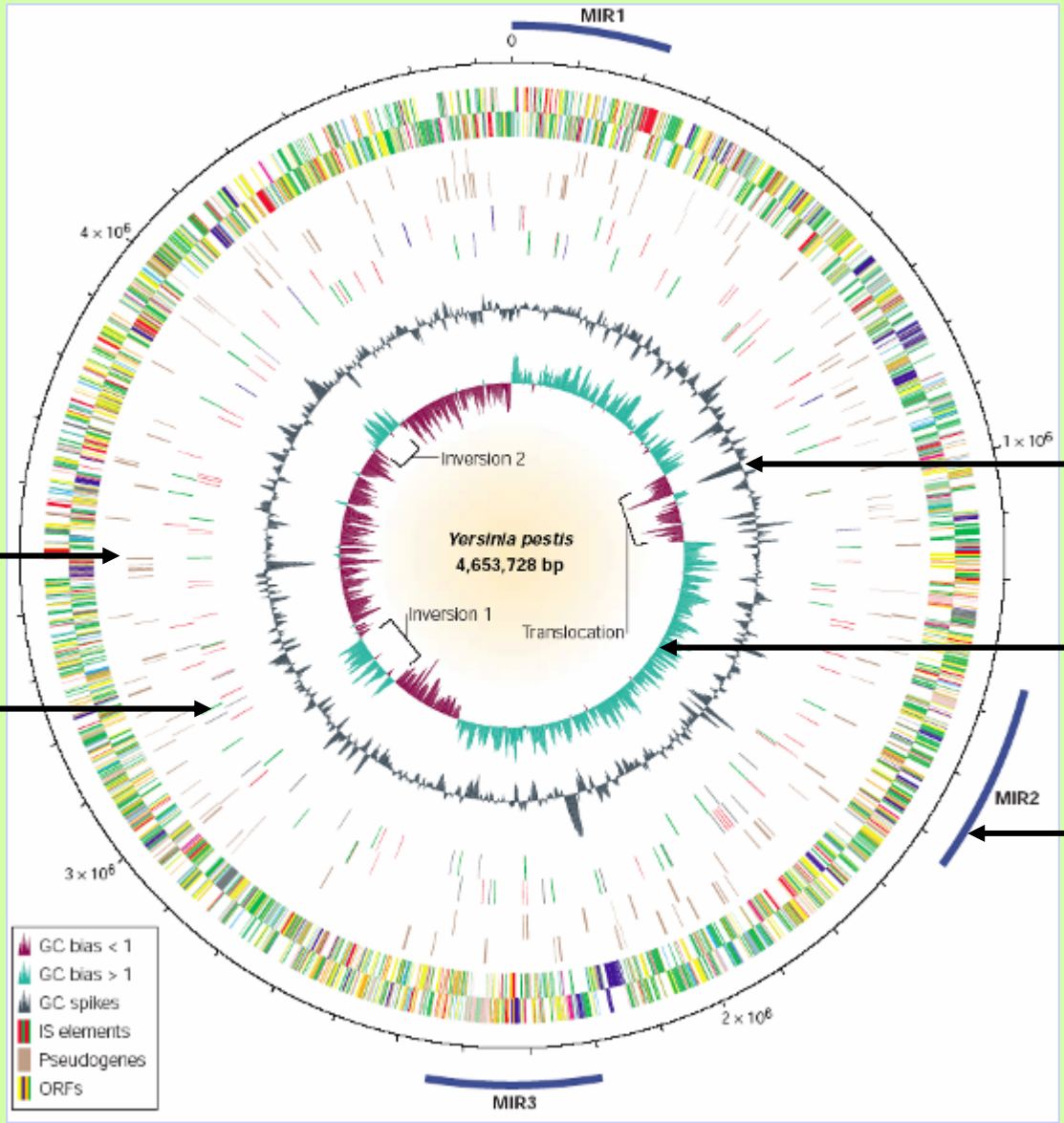
Table 2 | **Selected putative pathogenicity islands from *Y. pestis* (*Y. pseudotuberculosis*)**

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.

3. Genome rearrangements, transposon amplification, and gene degr

(whose direct effects on *Y. pestis* virulence are still unknown)



pseudogenes

IS elements

% G+C

GC skew

multiple inversion regions