J. Gen. Appl. Microbiol., **51**, 11–19 (2005)

Full Paper

Genetic variations in the *pgm* locus among natural isolates of *Yersinia pestis*

Zongzhong Tong,[2,4] Dongsheng Zhou,[1,4] Yajun Song,[1] Ling Zhang,[1] Decui Pei,[1] Yanping Han,[1] Xin Pang,[1] Min Li,[3] Baizhong Cui,[3] Jin Wang,[1] Zhaobiao Guo,[1] Zhizhen Qi,[3] Lixia Jin,[3] Junhui Zhai,[1] Zongmin Du,[1] Xiaoyi Wang,[1] Jian Wang,[2] Peitang Huang,[1] Huanming Yang,[2,*] and Ruifu Yang[1,*]

[1] *Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Beijing 100071, P. R. China*
[2] *Beijing Genomics Institute, Chinese Academy of Sciences, Beijing 101300, P. R. China*
[3] *Qinghai Institute for Endemic Diseases Prevention and Control, Xining 811602, P. R. China*

(Received December 22, 2003; Accepted October 12, 2004)

A PCR-based screening method was used to study the genetic variations of the *pgm* locus among natural isolates of *Yersinia pestis* from China. Our results indicate that genetic variations in the *pgm* locus are well correlated with biovars of *Y. pestis* and plague foci, suggesting that the *pgm* locus plays a role in *Y. pestis* adaptation to its environment. The gene encoding two-component regulatory system sensor kinase became a pseudogene in all strains of biovar *Orientalis* due to a thymidine deletion, while it is intact in all the strains of the other biovars. Only strains from Foci H and L are the same as *Yersinia pseudotuberculosis* in that they have an intact transmembrane helix in the sensor kinase protein, which is lost in all the other strains because of the 18 bp in-frame deletion. The IS*100* element that flanks the 3′ terminus of the *pgm* locus was inserted into the chromosome during the within-species microevolution of *Y. pestis*, which is absent in strains from Foci G, H and L and also in *Y. pseudotuberculosis*. This fact indicates that the strains from these three foci are of an older lineage of Chinese *Y. pestis*. It is this IS*100* element's absence that maintained high stability of the *pgm* locus in the *Y. pestis* strains from these three foci. The IS*285* element insertion in the pigmentation segment and the IS*100* element insertion in the downstream flanking region of the *pgm* locus are only present in strains from Foci H and L. The flanking region outside the 5′ terminus of the upstream IS*100* element is identical in the strains from these two foci, which is different in the other strains. All of these unique characteristics suggest that they are of a special lineage of Chinese *Y. pestis*.

Key Words——genetic variation; *pgm* locus; plague; *Yersinia pestis*

Introduction

Yersinia pestis, the causative agent of plague, is a Gram-negative bacterium that evolved from *Yersinia pseudotuberculosis* 1,500–20,000 years ago, shortly before the first known human plague pandemic (Achtman et al., 1999). Strains of *Y. pestis* were historically classified into three biovars based on their ability to reduce nitrate and ferment glycerol. Biovar *Antiqua* is positive for both characteristics: *Mediaevalis* ferments

These authors contributed equally to this work.

* Address reprint requests to: Dr. Ruifu Yang, Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Beijing 100071, P. R. China and Dr. Huanming Yang, Beijing Genomics Institute, Chinese Academy of Sciences, Beijing 101300, P. R. China.

E-mail: yangrf@nic.bmi.ac.cn (R.Y.)
yanghm@genomics.org.cn (H.Y.)

glycerol but does not generate nitrite from nitrate, and *Orientalis* synthesizes nitrite but does not ferment glycerol. These three biovars are thought to be responsible for the three recorded human plague pandemics: the Justinian plague, the Black Death and the modern plague (Perry and Fetherston, 1997).

Plague is a natural focus-based disease; human beings play no role in the long-term survival of the pathogen. The long-term survival and the periodical epidemic of animal plague occur in the plague's natural foci. Epidemics of human plague have existed in China for a long time, and it has been successfully controlled since the 1950s. However, there are eleven typical natural plague foci still remaining in China, covering more than 277 counties in 19 provinces with an area of more than 1 million square kilometers (Ji et al., 1990a; Liu et al., 2001). The epizootics of animal plague in these natural foci pose a great threat to public health. Chinese researchers have developed an ecotyping system which utilizes several biochemical features including glycerol utilization, utilization of various carbohydrates, nitrate reduction, nutrient requirement, mutation from Pgm^+ to Pgm^-, pesticin I production, and pesticin I sensitivity, the fraction 1 antigen content and SDS-PAGE for water-soluble proteins to group the Chinese isolates of *Y. pestis* into 18 ecotypes (Ji et al., 1990b).

The pigmentation phenotype (Pgm^+) was originally defined as the ability of *Y. pestis*, grown at 26°C, to absorb sufficient hemin or Congo red to form dark brown or red 'pigmented' colonies on solidified media (Jackson and Burrows, 1956). A chromosomal region that was designated as the *pgm* locus encodes this phenotype. The *pgm* locus is a 102-kb unstable DNA region embedded between two IS*100* elements in the same orientation (Buchrieser et al., 1999; Fetherston et al., 1992). The recombination between the two IS*100* elements will delete this locus, which accounts for the negative pigmentation phenotype (Pgm^-). The *pgm* locus is an established virulence-related gene cluster composed of several distinct parts: the high-pathogenicity island (HPI), a fimbriae gene cluster, the genes for a BvgAS-like two-component regulatory system and the *hms* (hemin storage) locus. The HPI carries virulence-related genes involved in iron acquisition. The *hms* locus is important for the transmission of *Y. pestis* by the flea vector (Hinnebusch et al., 1996).

The recent determination of the sequences of the whole genomes of *Y. pestis* strains CO92, KIM and 91001 and the whole *pgm* locus of strain 6/69M provided a good opportunity to study the genetic variation in the *pgm* locus (Buchrieser et al., 1999; Deng et al., 2002; Parkhill et al., 2001; Song et al., 2004). The in silico comparison and PCR-based screening of the genetic variation of the *pgm* locus among Chinese natural isolates of *Y. pestis* revealed that the *pgm* locus plays a role in the adaptation of *Y. pestis* to its environment. The absence of the IS*100* element that flanks the 3′ terminus of the *pgm* locus maintained its stability, and led to the permanent Pgm^+ phenotype in *Y. pestis* from Foci G, H and L. This suggests that they are of an older lineage of Chinese *Y. pestis*. The strains from Foci H and L may be located in a special lineage with specific genetic variations.

Materials and Methods

Bacterial strains. There are 11 natural plague foci in China. In the past 50 years, a large number of strains of *Y. pestis* were isolated from various host, vectors, and human patients from 10 of these foci excluding the *Marmota sibirica* plague focus of the Hulun Buir Plateau of Inner Mongolia. There is no collection of the bacterial strains from this focus at present. In this study four subfoci of the *Marmota himalayana* plague focus of the Qinghai-Tibet Plateau were investigated separately and referred to as Foci A, B, C and J. Two-hundred and sixty isolates of *Y. pestis* from the 10 natural plague foci in China, covering all of the three biovars (Table 3), and 7 strains of *Y. pseudotuberculosis* (Table 4) were used in this study. All the *Y. pestis* strains were collected by the Qinghai Center for Disease Prevention & Control, China. The *Y. pseudotuberculosis* strains were originally from the China Medical Culture Collection except for ATCC29833, which was purchased from the American Type Culture Collection. All the strains were grown in Luria-Bertani broth. Genomic DNA was extracted using the conventional phenol-chloroform extraction method.

In silico comparative analysis and primer design. The sequences of the *pgm* locus from *Y. pestis* CO92 (AL590842), 6/69M (AL031866) and KIM (AE009952) were downloaded from GenBank (Buchrieser et al., 1999; Deng et al., 2002; Parkhill et al., 2001). The whole-genome sequence of *Y. pestis* 91001 was recently completed in our laboratory (AE017042) (Song et al., 2004). The alignment and comparison of

the *pgm* locus and its flanking regions from *Y. pestis* CO92, KIM and 91001 were performed using BLAST, CLUSTAL X (1.8) and DNASTAR, and assisted with a Perl-based program to find all the minute genetic variations. Both *Y. pestis* 6/69M and CO92 are *Orientalis*, so the *pgm* locus sequence of *Y. pestis* 6/69M was not included in the comparison. All the primers for PCR-based screening of the genetic variations among the 267 bacterial strains were designed using Primer Premier 5.0 (Table 1).

Allele-specific PCR. The genomic DNA samples from all the bacterial strains were arrayed in a 96-well plate. Each primer pair was pre-tested to ensure the specificity of amplification with the 91001 genomic DNA as template. Thirty microliters of PCR reaction mixture contained 50 mmol/L KCl, 10 mmol/L Tris-HCl (pH 8.0), 2.5 mmol/L $MgCl_2$, 0.001% gelatin, 0.1% BSA, 100 μmol/L each of dATP, dCTP, dGTP and dTTP, 0.3 μmol/L of each primer, 1 unit of *Taq* DNA polymerase (MBI), and 10 ng of template DNA. The parameters for amplification were as follows: predenaturation at 95°C for 5 min, followed by 30 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and elongation at 72°C for 40 s, and a final extension at 72°C for 5 min to insure the complete extension of the amplicons. After amplification, 15 μl of each PCR product was subjected to electrophoresis on 1.2% agarose gel and visualized by ethidium bromide staining and UV irradiation.

Table 1. Primers used in this study.

Name	Sequence
91001-pgm-F1	5′-TGTCCATCCAGATAGGTC-3′
91001-pgm-R2	5′-CGAACTGACTAACTCTCG-3′
hutC-F	5′-AGCGCAGCAATGGACGAC-3′
hutC-R	5′-GACCATAGAGCTGATAAC-3′
IS100-c-L	5′-ACTCATTCCCTGCTTGTG-3′
IS100-c-R	5′-CACTGACCTCAGCGATGC-3′
IS285-c-L	5′-ATCTCAATACATTCTTCGGC-3′
IS285-c-U	5′-AACGCTGTTTCGACGGTAAGC-3′
pgm-F1	5′-TGAGGCAGGTAGGTATCG-3′
pgm-F2	5′-AGGCGCATACTCATCATC-3′
pgm-m-F	5′-TGGCCAGTGTCATTATAC-3′
pgm-m-R	5′-TGTTTCACCTACCGTAGC-3′
pgm-R1	5′-CAAGATGGAATTCGTCAG-3′
pgm-R2	5′-ATAACGCTTTAGATCTAC-3′
YP1666-L1M	5′-CAGGTTTTCTTTACGTAC-3′
YP1666-L1W	5′-CAGGTTTTCTTTACGTAG-3′
YP1666-LM	5′-TGAAGCCCCAGACACCAG-3′
YP1666-LW	5′-TGAAGCCCCAGACACCAT-3′
YP1666-R	5′-CTCCACTCACTGGTTAGG-3′
YP1666-R1	5′-ATGGCATTATCATCGCGG-3′
ybtT-L	5′-GTGCATCCCGCTGTGG-3′
ybtT-R	5′-GTACTGGGCTGTTTTTGC-3′
hmsS-L	5′-CGCCCCTGATTTTTACGG-3′
hmsS-R	5′-CATCCCTGGCGTAAATGG-3′

Relative positions of primers except for YP1666, *ybtT* and *hmsS* primers are shown in Fig. 1.

Results and Discussion

In silico comparison and allele-specific PCR

The genome structure of the *pgm* locus is well conserved among the three *Y. pestis* strains CO92, KIM and 91001. However, there are still some minute variations (single nucleotide mutations and insertions/deletions). Most of these mutations occur in the intergenic regions (data not shown). Table 2 lists all of the non-synonymous mutations and fragment deletions that occurred in the coding sequences.

YP1666 and YP1667 in 91001 encode the components of a two-component regulatory system similar to the BvgAS system of *Bordetella pertussis* (Buchrieser et al., 1999). There is a transmembrane helix between the 21st and 43rd amino acids predicted by TMHMM2.0 (Krogh et al., 2001) in the protein encoded by YP1666. An 18 bp (from the 70th to 87th) in-frame deletion occurred in the alleles of YP1666 in *Y. pestis* CO92 (YPO1923) and KIM (y2387), resulting in the loss of six amino acids in the transmembrane helix from the gene products. There is another frameshift in the allele in CO92 due to a thymidine deletion at the 1406th nucleotide.

There is an IS*285* insertion in the *pgm* locus of 91001 (Fig. 1), interrupting YP1700 that encodes a hypothetic protein. But this insertion is not present in the *pgm* locus of CO92 and KIM, and the corresponding alleles YPO1956 and y2354 are intact respectively.

The *pgm* locus is embedded between two IS*100* elements in the same orientation in the chromosome of CO92 and KIM. However, there is no IS*100* flanking the 3′ terminus of the *pgm* locus in 91001. Except for the presence or absence of this IS*100* insertion, the genetic content that immediately flanks the 3′ terminus is identical in the three strains. However, there is a novel IS*100* insertion in another position in the downstream flanking region of the *pgm* locus in the opposite orientation to the upstream one in 91001. The down-

Table 2. Variations among the alleles in *Y. pestis* 91001, CO92 and KIM that lead to nonsynonymous mutations or gene inactivations.

Allele in 91001	Allele in CO92	Allele in KIM	Product	Variation (91001/CO92/KIM)
YP1647	YPO1905	y2406	hypothetical protein	nt 40 (A/T/T)
YP1649	YPO1906	y2404	pesticin/yersiniabactin receptor protein	nt 1540 (G/A/A)
YP1653	YPO1910	y2400	yersiniabactin biosynthetic protein	nt 454 (A/C/C)
YP1656	YPO1913	y2397	permease and ATP-binding protein of yersiniabactin-iron ABC transporter	nt 1333 (A/G/A)
YP1659	YPO1916	y2394	putative salicylate synthetase	nt 661 (T/C/C)s
YP1666	YPO1923	y2387	putative histidine protein kinase sensor	nt 70–87 were deleted from CO92 and KIM; T at nt 1406 was deleted from CO92
YP1684	Unpredicted	Unpredicted	hypothetical protein	nt 241–246 (6As/7As/7As)
YP1695	YPO1951	y2359	hemin storage system, HmsH protein	nt 2038 (A/G/G)
YP1697	YPO1953	y2357	hemin storage system, HmsR protein	nt 188 (T/G/T)
YP1700	YPO1956	y2354	hypothetical protein	YP1700 was disrupted by IS*285*
YP1712	YPO1967	y2344	outer membrane protein (porin)	YPO1967 and y2344 were disrupted by IS*100*

Unpredicted means that the homologous sequence exists but has not been predicted as CDS. The positions of the variations are indicated by the nucleotide numbers counted from the 5′ terminus of the respective CDSs in 91001.

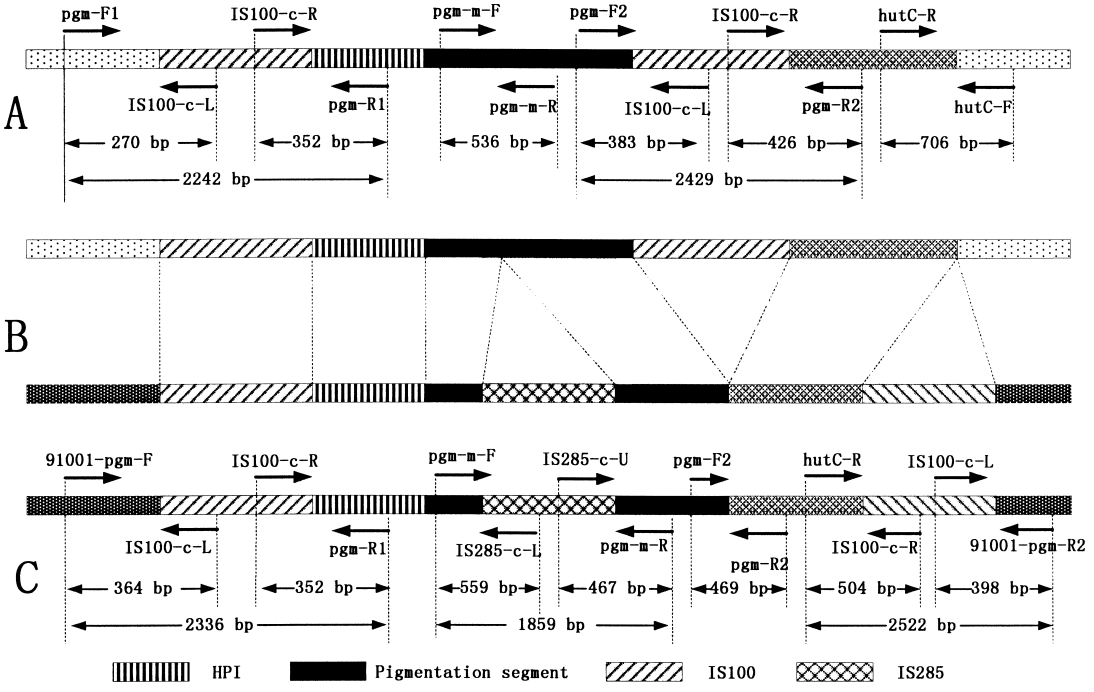

Fig. 1. Schematic diagrams of the *pgm* loci of *Y. pestis* CO92 (A) and 91001 (C) and their flanking regions, relative locations of primers and the lengths of amplification products.

B depicts linear genomic comparison of the *pgm* locus of *Y. pestis* CO92 and 91001, the upper line stands for the *pgm* locus of CO92 and the lower one stands for the *pgm* locus of 91001. The regions between dashed lines represent DNA-DNA similarities (BLASTN matches) between the two sequences.

stream region of this IS*100* is different to the downstream regions of the corresponding insertion site in CO92 and KIM (Fig. 1).

The flanking region outside the 5′ terminus of the IS*100* element of the *pgm* locus is identical in CO92 and KIM, but different in 91001 (Fig. 1). The 5′ flanking region in 91001 was found to be identical to another chromosomal region in CO92 or KIM. It is suggested that the rearrangement of this region, mediated by the homologous recombination of the IS*100* elements, occurred either in CO92 and KIM or in 91001 after the acquisition of the *pgm* locus through horizontal gene transfer.

Primers were designed to screen the distribution of the genetic variations including the 18 bp in-frame deletion and the thymidine deletion in YP1666, the IS*285* insertion in YP1700, the deletion of the downstream IS*100*, the IS*100* insertion in the 3′ terminal flanking region, and the flanking region outside the 5′ terminal IS*100* element (Fig. 1, Table 1).

The BvgAS-like two-component system appears to be nonfunctional in Orientalis strains

Bacteria have evolved sophisticated sensory mechanisms and intracellular signal transduction pathways to respond to a large number of extracellular signals in their continuously changing surroundings. The two-component regulatory system is a basic stimulus-response coupling mechanism used by bacteria to sense and respond to changing environmental conditions (Stock et al., 2000). The YP1666/YP1667 gene couple located in the pigmentation segment shows high homology to the *bvgS*/*bvgA* gene couple encoding a two-component regulatory system of *B. pertussis*, in which virulence gene expression is regulated (Stibitz et al., 1989). Our PCR analyses with primer pairs J and K (Table 3) indicate that the allele of YP1666 became a pseudogene due to a thymidine deletion in *Orientalis* strains from Focus M, though it is intact in all the other strains, which is in accord with a previous study (Buchrieser et al., 1999). Although the gene for the BvgS-like sensor protein became a pseudogene in *Orientalis* strains, the response regulator encoded by the remaining active gene may still find its cross-talking partner from the other two-component systems in *Y. pestis*. A similar cascade has been found in *Escherichia coli* (Verhamme et al., 2002).

Unique features of the pgm locus in Microtus strains

There are two *Microtus*-related natural plague foci in China: Foci H and L (the isolates from these two foci are referred to as *Microtus* strains in this article). Human plague resulting from *Microtus* strains has never been reported, but the epizootic of the *Microtus* plague occurs almost every year in these two natural foci. Strain 91001 from Focus H has been proven to be avirulent to human beings by subcutaneous inoculation (Fan et al., 1994). Comparative genomic analysis revealed that *Microtus* strains have undergone deletions of certain genomic regions and accumulation of point mutations in individual genes. Strains from these two foci were proposed as a novel biovar *Microtus* for their unique pathogenic, biochemical and molecular features (Zhou et al., 2004b).

The molecular features unique to *Microtus* strains in the *pgm* locus and its flanking sequences were clarified in this work by the allele-specific PCR methods (Tables 3 and 5). The IS*285* insertion in YP1700 (as shown by experiment using primer pair D) and the IS*100* insertion in the 3′ terminal flanking region of the *pgm* locus (primer pairs D1, E1 and F1) are only present in the *Microtus* strains, which show the Pgm^+ phenotype. The IS*285* insertion does not affect the Pgm^+ phenotype for which probably only the *hmsHFRS* cluster is essential.

Based on the amplification result with primer pairs L and M, it can be inferred that the 18 bp in-frame deletion in the allele of YP1666 is present in all the *Y. pestis* strains analyzed in this study, except the *Microtus* strains and the seven *Y. pseudotuberculosis* strains (Tables 3 and 4). The 18 bp in-frame deletion results in the loss of the transmembrane helix of the gene product. The protein encoded by YP1666 and its allele is a sensor kinase of the BvgAS-like two-component system. So only the *Microtus* strains and the *Y. pseudotuberculosis* strains contain the sensor kinase protein with the intact transmembrane helix. It cannot be determined at present whether the sensor kinase in other strains is nonfunctional due to the lost transmembrane helix, because not all orthodox kinases of two-component regulatory systems are membrane bound proteins with transmembrane helices. The chemotaxis kinase CheA and the nitrogen regulatory kinase NtrB are examples of soluble cytoplasmic kinases. Soluble kinases can be regulated by intracellular stimuli and/or interactions with cytoplasmic domains of other membrane proteins (Stock et al.,

Table 3. Amplification results of 260 strains of *Y. pestis* isolated from Chinese natural foci (sorted by focus).

Focus	Numbers of strains	Primer pair													Del
		A	B	C	D	E	F	G	H	I	J	K	L	M	
Focus A[A]	16	1	1	1	1[A]	1	1[C]	1	1	0	1	0	0	1	1
Focus B[A]	37	1	1	1	1[A]	1	1[C]	1	1	0	1	0	0	1	2
Focus C[A]	11	1	0	1	1[A]	1	1[C]	1	1	0	1	0	0	1	1
Focus D[A]	30	1	0/1	0/1	0/1[A]	0/1	0/1[C]	1	1	0	0/1	0	0	0/1	9
Focus E[A]	10	1	1	1	1[A]	1	1[C]	1	1	0	1	0	0	1	0
Focus F[A]	11	1	1	1	1[A]	1	1[C]	1	1	0	1	0	0	1	1
Focus G[A]	30	1	1	1	1[A]	0	1[D]	0	1	0	1	0	0	1	0
Focus H[M]	10	1	0	1	1[B]	0	1[D]	0	0	1	1	0	1	0	0
Focus I[M]	21	1	0/1	0/1	0/1[A]	0/1	0/1[C]	1	1	0	0/1	0	0	0/1	8
Focus J[M]	13	1	0	0/1	0/1[A]	0/1	0/1[C]	1	1	0	0/1	0	0	0/1	10
Focus K[M]	20	1	1	1	1[A]	1	1[C]	1	1	0	1	0	0	1	0
Focus L[M]	29	1	0	1	1[B]	0	1[D]	0	0	1	1	0	1	0	0
Focus M[O]	22	1	1	1	1[A]	1	1[C]	1	1	0	0	1	0	1	4

Foci are: Focus A, *Marmota himalayana* Plague Focus of the Qilian Mountain; Focus B, *Marmota himalayana* Plague Focus of the Qinghai-Gansu-Tibet Grassland; Focus C, *Marmota himalayana* Plague Focus of the Gangdisi Mountains; Focus D, *Spermophilus dauricus* Plague Focus of the Song-Liao Plain; Focus E, *Apodemus chevrieri-Eothenomys miletus* Plague Focus of the highland of Northwestern Yunnan Province; Focus F, *Marmota caudate* Plague Focus of the Pamirs Plateau; Focus G, *Marmota baibacina-Spermophilus undulates* Plague Focus of the Tianshan Mountains; Focus H, *Microtus brandti* Plague Focus of the Xilin Gol Grassland; Focus I, *Meriones unguiculatus* Plague Focus of the Inner Mongolian Plateau; Focus J, *Marmota himalayana* Plague Focus of the Kunlun Mountains; Focus K, *Spermophilus dauricus alaschanicus* Plague Focus of the Loess Plateau in Gansu and Ningxia provinces; Focus L, *Microtus fuscus* Plague Focus of the Qinghai-Tibet Plateau; Focus M, *Rattus flavipectus* Plague Focus of the Yunnan-Guangdong-Fujian provinces; Foci A, B, C and J are the sub-foci of the *Marmota himalayana* Plague Focus of the Qinghai-Tibet Plateau.

[A], biovar *Antiqua*; [M], biovar *Mediaevalis*; [O], biovar *Orientalis*.

Primer pairs used were: A, pgm-F1/IS100-c-L; B, pgm-F1/pgm-R1; C, pgm-R1/IS100-c-R; D, pgm-m-F/pgm-m-R; E, pgm-F2/IS100-c-L; F, pgm-F2/pgm-R2; G, pgm-R2/IS100-c-R; H, hutC-F/hutC-R; I, 91001-pgm-F1/IS100-c-L; J, YP1666-LW/YP1666-R; K, YP1666-LM/YP1666-R; L, YP1666-L1W/YP1666R1; M, YP1666-L1M/YP1666R1.

Results are indicated as follows: 0, negative; 1, positive; 0/1, some strains were negative and some strains were positive; 1[A], product length was 0.5 kb; 1[B], product length was 1.9 kb; 1[C], product length was 2.4 kb; 1[D], product length was 0.5 kb bp.

Del, the numbers of strains from each focus with deletion of the whole *pgm* locus.

Table 4. Amplification results of 7 strains of *Y. pseudotuberculosis*.

Strains	Primer pairs									
	A*	B	C	D	F	H	J	K	L	M
CMCC53518	0	0	1	1[A]	1[B]	1	0	1	1	0
CMCC53519	0	0	0	1[A]	1[B]	1	0	1	1	0
CMCC53520	0	0	0	1[A]	1[B]	1	0	1	1	0
CMCC53521	0	0	0	1[A]	1[B]	1	0	1	1	0
CMCC53522	0	0	0	1[A]	1[B]	1	0	1	1	0
ATCC29833	0	0	0	1[A]	1[B]	1	0	1	1	0
CMCC53502	0	0	0	1[A]	1[B]	1	0	1	1	0

* A to M stands for the primer pairs as in the footnote in Table 3.

0, negative; 1, positive; 1[A], product length was 0.5 kb; 1[B], product length was 1.9 kb.

2000). Loss of the transmembrane helix may not influence the regulatory function of this two-component regulatory system. Therefore, further study is needed to determine whether the sensor inactivation or the transmembrane helix deletion is responsible for the difference in virulence among *Y. pestis* strains of different biovars or different origins.

Absence of the downstream IS100 element may account for the stability of pigmentation phenotype

The *pgm* locus is flanked with two direct repeats of the IS*100* element in the chromosome. Spontaneous deletion of the *pgm* locus is due to the homologous recombination of these two IS*100* elements. The downstream IS*100* element cannot be found in the *Y. pestis* isolates from Focus G, the *Microtus* strains and *Y. pseudotuberculosis* strains as shown by PCR analyses using primer pairs E, F and G (Tables 3 and 4). *Y. pestis* is a clone that evolved from *Y. pseudotuberculosis* 1,500–20,000 years ago (Achtman et al., 1999). It can be postulated that the downstream IS*100* element was inserted into the chromosome during the within-species microevolution of *Y. pestis*. *Y. pestis* strains from Focus G and *Microtus* strains are older than the other Chinese isolates, which was also confirmed by the DNA microarray-based evolutionary genomic analysis of the natural populations of *Y. pestis* in China (Zhou et al., 2004a).

The biochemical tests revealed that the Pgm^+ phenotype of the strains from Focus G and *Microtus* strains were very stable even after the long-term and successive passages of bacterium in vitro (Song, 1990). Indeed, all the strains from these three foci tested in this study harbor the whole *pgm* locus, except for one losing the HPI region (see below). The stability of the *pgm* locus in these strains is probably due to the absence of the downstream IS*100* element that can undergo homologous recombination with the upstream IS*100* in the same orientation. In the *Microtus* strains, there is an IS*100*-insertion in the 3′ terminal flanking region of the *pgm* locus, but its orientation is opposite to the upstream IS*100* element as shown by PCR analyses using primer pairs D1, E1 and F1 (Table 5). The homologous recombination mediated by inverted repeats leads to the inversion of the *pgm* locus but not its deletion.

Variation found in the flanking region outside the 5′ terminal IS100 element of the pgm locus

Table 5. Amplification results of *Y. pestis* strains from Foci H, L and A.

Natural focus	Primer pairs					
	A1*	B1	C1	D1	E1	F1
Focus H	1	1	1	1	1	1
Focus L	1	1	1	1	1	1
Focus A	0	0	0	0	0	0

*A1 to F1 represents the primer pairs: A1, 91001-pgm-F1/pgm-R1; B1, pgm-m-F/IS285-c-L; C1, pgm-m-R/IS285-c-R; D1, hutC-R/91001-pgm-R2; E1, hutC-R/IS100-c-R; F1, 91001-pgm-R2/IS100-c-L.

0, negative; 1, positive.

Strains from Focus A were used as a negative control.

The upstream IS*100* element is conserved in all *Y. pestis* and it presents in one of seven *Y. pseudotuberculosis* too (the other six lost HPI). However, the in silico comparative analysis revealed that the flanking region outside the 5′ terminal IS*100* element of the *pgm* locus is divergent between 91001 and CO92 (CO92 and KIM have the same 5′ flanking region). Two sense primers were designed based on the divergence. These two primers, coupled with a common antisense primer annealing to the 5′ terminal of IS*100* (Primer pairs A and I) and of the *pgm* locus (Primer pairs B and A1), were used in allele-specific PCR to screen for the distribution of the variation in the 260 natural isolates (Tables 3 and 4). The amplicon size of each *Microtus* strain is the same as that of 91001, indicating the uniformity of the upstream flanking region of the *pgm* locus among *Microtus* strains. Interestingly, both the primer pairs B and I gave no product for the strains from Foci C and J, indicating novel genetic content in the upstream flanking region of the *pgm* locus. However, the PCR results of all the remaining strains of *Y. pestis* are the same as those of CO92 and KIM. Various genetic contents of the flanking regions may be the consequence of the *pgm* locus inserted into different positions on the chromosome in different strains. For example, the high-pathogenicity island of *Y. pseudotuberculosis* can be inserted into any of the three chromosomal *asn* tRNA genes (Buchrieser et al., 1998b). Considering the conservation of the 3′-flanking region, however, it is more likely that the divergence was the consequence of intra-chromosome rearrangements including the intra-genomic inversion

of the 5′ terminal flanking region mediated by the homologous recombination of the IS*100* elements. Such rearrangements probably occurred in some of the strains after the integration of the *pgm* locus into the chromosome.

Instability of the pgm locus in natural isolates of Y. pestis

The GC content analysis suggested that the *pgm* locus is composed of two distinct parts: the pigmentation segment and high-pathogenicity island (HPI) (Fig. 1). These two parts were acquired individually through horizontal gene transfer (Buchrieser et al., 1998a). Proteins encoded by three genes located at both extremities of HPI show high homology to phage proteins, indicating that the HPI locus was acquired by phage mediated horizontal gene transfer (Buchrieser et al., 1999). In this study, the *ybtT* gene in the HPI and the *hmsS* gene in the pigmentation segment were arbitrarily selected to represent the respective entire relevant locus. PCRs targeting these two genes were performed to screen for the distribution of the HPI and the pigmentation segment in the 260 *Y. pestis* strains and the seven *Y. pseudotuberculosis* strains. Six of the seven strains of *Y. pseudotuberculosis* lost HPI while their pigmentation segment still remained. Thirty-six of the 260 *Y. pestis* isolates lost both *ybtT* and *hmsS*, and probably underwent deletion of the entire *pgm* locus (Table 3). One underwent the selective deletion of HPI (data not shown). Most of the *pgm*-incomplete isolates of *Y. pestis* are distributed in three natural foci: I, D and J. This is consistent with a previous discovery that the frequencies of Pgm^+ to Pgm^- mutation differ among strains from different natural foci in China and the mutation frequencies in the strains from Foci F, I, J and M are higher than those from the other foci (Song, 1990).

The *pgm* locus plays important roles in the pathogenicity and transmission of *Y. pesits* (Hinnebusch et al., 1996). Interestingly, the *pgm*-incomplete strains were found in natural isolates of *Y. pestis*. Moreover, all *Microtus* strains tested harbor the whole *pgm* locus, although these strains are thought to be avirulent to human beings (Fan et al., 1994). Genomic islands (e.g. the *pgm* locus) are present not only in the genomes of the pathogenic bacteria but also those of their nonpathogenic counterparts. Genomic islands contribute to the pathogenic potency and function as pathogenicity islands in pathogenic bacteria, while in nonpathogenic bacteria they are ecological islands that increase the bacterial fitness to the specific ecological niches (Hacker and Carinel, 2001). Therefore we propose that, for the strains of *Y. pestis* of different origins, the *pgm* locus plays different roles in the bacterial adaptation to the niches through complex interactions with other *Y. pestis* strains, animal reservoirs, flea vector and natural environment.

Acknowledgments

This work was supported by National 863 program (contract no. 2001AA223061) and National Natural Science fund (contract no. 30371284). We thank Dr. David Bastin (Tianjing Biochip Co., Tianjing, China) and Mr. Qi Guo for critical reading of the manuscript.

References

Achtman, M., Zurth, K., Morelli, G., Torrea, G., Guiyoule, A., and Carniel, E. (1999) *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc. Natl. Acad. Sci. USA,* **96**, 14043–14048.

Buchrieser, C., Prentice, M., and Carniel, E. (1998a) The 102-kilobase unstable region of *Yersinia pestis* comprises a high-pathogenicity island linked to a pigmentation segment which undergoes internal rearrangement. *J. Bacteriol.,* **180**, 2321–2329.

Buchrieser, C., Roland, B., Sandrine, B., Annie, G., and Elisabeth, C. (1998b) The high-pathogenicity island of *Yersinia pseudotuberculosis* can be inserted into any of the three chromosomal asn tRNA genes. *Mol. Microbiol.,* **30**, 965–978.

Buchrieser, C., Rusniok, C., Frangeul, L., Couve, E., Billault, A., Kunst, F., Carniel, E., and Glaser, P. (1999) The 102-kilobase *pgm l*ocus of *Yersinia pestis*: Sequence analysis and comparison of selected regions among different *Yersinia pestis* and *Yersinia pseudotuberculosis* strains. *Infect. Immun.*, **67**, 4851–4861.

Deng, W., Burland, V., Plunkett, G., 3rd, Boutin, A., Mayhew, G. F., Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C., Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plano, G. V., Straley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S., Blattner, F. R., and Perry, R. D. (2002) Genome sequence of *Yersinia pestis* KIM. *J. Bacteriol.*, **184**, 4601–4611.

Fan, Z., Luo, Y., Li, F., Zhang, C., Su, X., and Wang, S. (1994) The pathogenicity of *Yersinia pestis* of Xilin Gol grassland type. *Chin. J. Control Endemic Dis.,* **9**, 340–342 (in Chinese).

Fetherston, J. D., Schuetze, P., and Perry, R. D. (1992) Loss of the pigmentation phenotype in *Yersinia pestis* is due to the spontaneous deletion of 102 kb of chromosomal DNA which is flanked by a repetitive element. *Mol. Microbiol.*, **6**,

2693–2704.

Hacker, J. and Carinel, E. (2001) Ecological fitness, genomics islands and bacterial pathogenicity. *EMBO Rep.*, **2**, 376–381.

Hinnebusch, B. J., Perry, R. D., and Schwan, T. G. (1996) Role of the *Yersinia pestis* hemin storage (*hms*) locus in the transmission of plague by fleas. *Science*, **273**, 367–370.

Jackson, S. and Burrows, T. W. (1956) The pigmentation of *Pasteurella pestis* on a defined medium containing haemin. *Br. J. Exp. Pathol.*, **37**, 570–576.

Ji, S., He, J., Teng, Y., Zhan, X., Lei, C., and Wang, W. (1990a) The discovery and research of plague natural foci in China. *Chin. J. Epidemiol.*, **11**(Suppl.), 1–41 (in Chinese).

Ji, S., Zhang, H., Liu, Y., He, Y., Huang, J., Song, Y., and Zhu, J. (1990b) A study on typing of *Yersinia pestis* in China and its ecologioepidemiological significance. *Chin. J. Epidemiol.*, **11**, 60–66 (in Chinese).

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.*, **305**, 567–580.

Liu, Z., Hai, R., Li, F., Li, C., Cong, X., Gao, C., Wang, L., Wei, B., Chen, H., Chen, H., Yu, X., Li, C., Shi, Y., Xu, G., and Li, G. (2001) The discovery and study of *Microtus fuscus* nature plague foci in Qinghai-Tibet Plateau. *Chin. J. Control Endemic Dis.*, **16**, 321–327.

Parkhill, J., Wren, B. W., Thomson, N. R., Titball, R. W., Holden, M. T., Prentice, M. B., Sebaihia, M., James, K. D., Churcher, C., Mungall, K. L., Baker, S., Basham, D., Bentley, S. D., Brooks, K., Cerdeno-Tarraga, A. M., Chillingworth, T., Cronin, A., Davies, R. M., Davis, P., Dougan, G., Feltwell, T., Hamlin, N., Holroyd, S., Jagels, K., Karlyshev, A. V., Leather, S., Moule, S., Oyston, P. C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., and Barrell, B. G. (2001) Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature*, **413**, 523–527.

Perry, R. D. and Fetherston, J. D. (1997) *Yersinia pestis*—Etiologic agent of plague. *Clin. Microbiol. Rev., **10**, 35–66.*

Song, Y. (1990) Research on some biological characters, biotypes and epidemiological significance of Chinese strains of *Y. pestis. Chin. J. Epidemiol.,* **11**(Suppl.), 67–76 (in Chinese).

Song, Y., Tong, Z., Wang, J., Wang, L., Guo, Z., Han, Y., Zhang, J., Pei, D., Zhou, D., Qin, H., Pang, X., Han, Y., Zhai, J., Li, M., Cui, B., Qi, Z., Jin, L., Dai, R., Chen, F., Li, S., Ye, C., Du, Z., Lin, W., Wang, J., Yu, J., Yang, H., Wang, J., Huang, P., and Yang, R. (2004) Complete genome sequence of *Yersinia pestis* strain 91001, an isolate avirulent to human. *DNA Res.*, **11**, 179–197.

Stibitz, S., Aaronson, W., Monack, D., and Falkow, S. (1989) Phase variation in *Bordetella pertussis* by frameshift mutation in a gene for a novel two-component system. *Nature*, **338**, 266–269.

Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-component signal transduction. *Annu. Rev. Biochem.*, **69**, 183–215.

Verhamme, D. T., Arents, J. C., Postma, P. W., Crielaard, W., and Hellingwerf, K. J. (2002) Investigation of in vivo cross-talk between key two-component systems of *Escherichia coli. Microbiology*, **148**, 69–78.

Zhou, D., Han, Y., Song, Y., Tong, Z., Wang, J., Guo, Z., Pei, D., Pang, X., Zhai, J., Li, M., Cui, B., Qi, Z., Jin, L., Dai, R., Du, Z., Bao, J., Zhang, X., Yu, J., Huang, P., and Yang, R. (2004a) DNA microarray analysis of genome dynamics in *Yersinia pestis*: Insights into bacterial genome microevolution and niche adaptation. *J. Bacteriol.*, **186**, 5138–5146.

Zhou, D., Tong, Z., Song, Y., Han, Y., Pei, D., Pang, X., Zhai, J., Li, M., Cui, B., Qi, Z., Jin, L., Dai, R., Du, Z., Wang, J., Guo, Z., Huang, P., and Yang, R. (2004b) Genetics of metabolic variations between *Yersinia pestis* biovars and the proposal of a new biovar, microtus. *J. Bacteriol.*, **186**, 5147–5152.