HAL open science

Delineation and Analysis of Chromosomal Regions Specifying Yersinia pestis

Anne Derbise, Viviane Chenal-Francisque, Christèle Huon, Corinne Fayolle, Christian Demeure, Béatrice Chane-Woon-Ming, Claudine Médigue, B. Joseph Hinnebusch, Elisabeth Carniel

To cite this version:

Anne Derbise, Viviane Chenal-Francisque, Christèle Huon, Corinne Fayolle, Christian Demeure, et al.. Delineation and Analysis of Chromosomal Regions Specifying Yersinia pestis. Infection and Immunity, 2010, 78 (9), pp.3930-3941. 10.1128/IAI.00281-10 . pasteur-03260318

HAL Id: pasteur-03260318

https://pasteur.hal.science/pasteur-03260318

Submitted on 3 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Delineation and analysis of chromosomal regions specifying *Yersinia pestis*

Anne Derbise[1*], Viviane Chenal-Francisque[1], Christèle Huon[1], Corinne Fayolle[1], Christian E. Demeure [1], Béatrice Chane-Woon-Ming[2], Claudine Médigue[2], B. Joseph Hinnebusch[3] and Elisabeth Carniel[1]

[1]*Yersinia* Research Unit, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France.

[2] Commissariat à l'Energie Atomique (CEA), Direction des Sciences du Vivant, Institut de Génomique, Genoscope & CNRS-UMR 8030, Laboratoire d'Analyse Bioinformatique en Génomique et Métabolisme, 2 rue Gaston Crémieux, 91057 Evry Cedex, Evry cedex F-91006, France.

[3] Laboratory of Zoonotic Pathogens, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT 59840, USA.

Key words: plague, horizontal transfer, Yersinia, evolution, genomic

Running title: Chromosomal regions acquired by *Y. pestis*

To whom correspondence should be addressed.

Dr. Anne Derbise

Institut Pasteur,

Yersinia Research Unit

28, rue du Dr. Roux,

25 75724 Paris Cedex 15

26 France.

27 Tel: (33-1)-45-68-84-48.

28 Fax: (33-1)-45-68-89-54.

29 E-mail: anne.derbise@pasteur.fr

Abstract

Yersinia pestis, the causative agent of plague, has recently diverged from the less virulent enteropathogen *Y. pseudotuberculosis*. Its emergence has been characterized by massive genetic loss and inactivation, and limited gene acquisition. Acquired genes include two plasmids, a filamentous phage and a few chromosomal loci. The aim of this study was to characterize the chromosomal regions acquired by *Y. pestis*. Following in silico comparative analysis and PCR screening of 98 strains of *Y. pseudotuberculosis* and *Y. pestis*, we found that eight chromosomal loci (six regions: R1pe-R6pe and two coding sequences : CDS1pe-CDS2pe) specified *Y. pestis*. Signatures of integration by site specific or homologous recombination were identified for most of them. These acquisitions and the loss of ancestral DNA sequences were concentrated in a chromosomal region opposite to the origin of replication. The specific regions were acquired very early during *Y. pestis* evolution and were retained during its microevolution, suggesting that they might bring some selective advantages. Only one region (R3pe), predicted to carry a lambdoid prophage, is most likely no longer functional because of mutations. With the exception of R1pe and R2pe, which have the potential to encode a restriction/modification and a sugar transport system, respectively no functions could be predicted for the other *Y. pestis*-specific loci. To determine the role of the eight chromosomal loci in the physiology and pathogenicity of the plague bacillus, each of them was individually deleted from the bacterial chromosome. None of the deletants exhibited defects during growth in vitro. Using the *Xenopsylla cheopis* flea model, all deletants retained the capacity to produce a stable and persistent infection, and to block fleas. Similarly, none of the deletants caused any acute flea toxicity. In the mouse model of infection, all deletants were fully virulent upon subcutaneous or aerosol infections. Therefore, our results suggest that acquisition of new chromosomal materials has not been of major

importance in the dramatic change of life cycle that has accompanied the emergence of *Y. pestis*.

57 **Introduction**

58 *Yersinia pestis*, the causative agent of plague, is a highly uniform clone that has recently
59 diverged from the enteric pathogen *Y. pseudotuberculosis* (2). Despite their close genetic
60 relationship, the plague bacillus and its recent ancestor differ radically in their pathogenicity
61 and transmission. *Y. pseudotuberculosis* is transmitted by the fecal-oral route and causes
62 enteritis of moderate intensity in Humans (48). In contrast, *Y. pestis* is transmitted by fleabites
63 or aerosols and causes highly severe infections, i.e. bubonic and pneumonic plague (37).

64 The transformation of *Y. pseudotuberculosis* into *Y. pestis* has been accompanied by the
65 acquisition of two plasmids, pFra and pPla. The 101-kb pFra plasmid is required for
66 colonization of and survival in the flea midgut and appears to increase the efficacy of
67 infection after flea bites (46). However, this plasmid is dispensable for *Y. pestis* pathogenicity
68 upon subcutaneous infection in mice (18) or aerosol infection in African green monkeys (9).
69 The 9.6-kb pPla plasmid, which encodes the plasminogen activator and protease Pla, is
70 critical for the virulence of some *Y. pestis* strains (55). It was shown to be required for the
71 translocation of the bacteria from their intradermal site of inoculation to the draining lymph
72 node (49), but not for their direct dissemination to the blood stream (47). Nonetheless, recent
73 studies have shown that *Y. pseudotuberculosis*, which is naturally deprived of pPla, can reach
74 the lymph node draining the intradermal site of inoculation as efficiently as *Y. pestis* (19).
75 Therefore, the stage of the infectious process during which pPla exerts its activity remains
76 unclear. Furthermore, acquisition of pPla is not sufficient to account by itself for the dramatic
77 rise in virulence of *Y. pestis* because there are natural isolates of *Y. pestis* that lack pPla and
78 still are highly virulent to mice after subcutaneous injection (45, 55), and introduction of pPla
79 in *Y. pseudotuberculosis* does not increase its virulence in mice (31, 38).

80 Chromosomal determinants are thus likely to be involved in the modifications in
81 pathogenicity and life cycle that accompanied the emergence of *Y. pestis*. Surprisingly, it

82 appears that the transformation of the enteropathogen *Y. pseudotuberculosis* into the plague
83 bacillus has been mostly characterized by extensive loss of genetic material and functions (8).
84 Ten chromosomal loci conserved in the species *Y. pseudotuberculosis* have been lost by *Y.*
85 *pestis* during the course of its differentiation (39, 40), and 208 genes present in the two
86 species have been inactivated during the transformation process (8). The natural inactivation
87 of one of these genes has recently been shown to have participated to the change in *Y. pestis*
88 mode of transmission (51).

89 In contrast, only a few new chromosomal genetic materials, designated here as *Y. pestis*-
90 specific, have been acquired during the transformation of an enteropathogen into the hyper
91 virulent plague bacillus. One of these loci encodes a filamentous phage (YpfΦ), which is
92 functional and participates (albeit moderately) to *Y. pestis* pathogenicity (11). Previous
93 comparative genomic analyses of *Y. pestis* and *Y. pseudotuberculosis* have identified
94 additional *Y. pestis*-specific genes (8, 24, 42, 56, 57). Some of these studies involved large
95 arrays of *Y. pestis* strains (up to 260), but the number of *Y. pseudotuberculosis* isolates was
96 always much more limited (nine or less). Similarly, while the number of *Y. pestis* genome
97 sequences available is increasing rapidly
98 (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?database=386656), only two genomes
99 of *Y. pseudotuberculosis* have been published until now (8, 14). Nonetheless, since *Y.*
100 *pseudotuberculosis* is a much older species, it displays a level of genetic polymorphism much
101 higher than that of the relatively clonal species *Y. pestis* (2). Therefore, the delineation of *Y.*
102 *pestis*-specific chromosomal regions based on the comparison with a few *Y.*
103 *pseudotuberculosis* isolates could not be accurate.

104 In this study, a large panel of *Y. pestis* and *Y. pseudotuberculosis* strains was screened to
105 identify chromosomal regions that specify the plague bacillus. The genetic organisation,
106 potential mode of integration into the ancestral bacterial chromosome, putative function and

107 microevolutionary step at which the plague bacillus acquired these regions was studied.
108 Finally, the role of this new genetic materials on *Y. pestis* growth, transmission and
109 pathogenicity was evaluated.
110

111 Materials and Methods

112 **Bacterial strains, plasmids and culture conditions.** The natural *Y. pestis* and *Y.*
113 *pseudotuberculosis* isolates used in this study are listed in Supplementary Tables S1 and S2.
114 They were taken from the collection of the *Yersinia* Research Unit (Institut Pasteur). The
115 various *Y. pestis* CO92 derivatives constructed here are described in Supplementary Table S3.
116 All bacteria were grown in Luria-Bertani (LB), or M63S broth (KH_2PO_4, 0.1M; $(NH_4)_2SO_4$,
117 0.2%; $MgSO_4$, 0.02%; $FeSO_4$, 0.00005%) with 0.2% glucose and 0.2% casaminoacids, and on
118 LB agar plates supplemented with 0.002% (w/v) hemin at 28°C or 37°C. Chloramphenicol
119 (25 µg/ml), kanamycin (Km, 30 µg/ml), or irgasan (0.1 µg/ml) were added to the media when
120 necessary. The pigmentation phenotype of *Y. pestis* was determined after 4 days of growth on
121 Congo red agar plates (52). All experiments involving *Y. pestis* strains were performed in a
122 BSL3 laboratory.
123

124 **In silico analysis of *Yersinia* genomes.** The genome sequences of *Y. pestis* CO92 (36) and *Y.*
125 *pseudotuberculosis* IP32953 (8) were compared pairwise using the Artemis Comparison Tool
126 (ACT) (http://www.sanger.ac.uk/Software/ACT). Re-annotation was performed using the
127 AMIGene software (Annotation of MIcrobial Genes)(6). DNA and protein sequences were
128 analyzed with Blast programs (3). Search for syntenies was performed with the tools
129 implemented in the MicroScope platform (54) and available via the MaGe interface
130 (https://www.genoscope.cns.fr/agc/mage/YersiniaScope).
131

132 **DNA manipulations.** Bacterial genomic DNA was extracted with the Isoquick kit (ORCA
133 Research Inc.). All primers used for PCR amplification are listed in Supplementary Table S4.
134 PCR were performed with 1 unit of Taq polymerase (Roche) or 1 unit of a 3:1 mixture of Taq
135 and Pfu (Stratagene) polymerases in the supplier's buffer. PCR amplification reaction
136 mixtures contained 10 µM of each primer and 1 mM dNTPs. The PCR program involved one
137 step at 94°C for 5 min, followed by 35 cycles of amplification of three steps: (i) 94°C for 30s,
138 (ii) 55°C for 30s and (iii) 72°C for 1 min. PCR products were maintained at 72°C for 7 min,
139 electrophoresed in 1% agarose gels, and stained with ethidium bromide. DNA sequencing
140 was performed on PCR fragments by the Millegene company.
141
142 **RNA isolation and transcript analysis.** RNA extraction was performed with the TRI
143 Reagent® kit (Ambion) on 4 ml of *Y. pestis* grown to stationary phase in LB at 28°C. DNA
144 was removed from the RNA samples by adding 4 U DNase, as described in the DNA-free kit
145 instructions (Ambion). Total RNA was quantified by optical density at OD_{260} and 200 ng
146 were used for reverse transcription and PCR analysis using primer pairs 623/624 (for control
147 of RNA quality and absence of DNA) and 703A/B (for detection of YPO2469 specific
148 mRNA transcript, Table S3).
149
150 **Mutagenesis.** Deletion of *Y. pestis* specific sequences was done by allelic exchange with a
151 Km (*km*) resistance cassette, following the Short Flanking Homology procedure (12) for $R1_{pe}$
152 to $R5_{pe}$ and $CDS2_{pe}$, and following the Long Flanking Homology procedure (12) for $R6_{pe}$ and
153 $CDS1_{pe}$. The Km resistance cassette was obtained by PCR amplification using plasmid
154 pGP704N-*km* as template and primer pairs listed in Table S3. The PCR products were
155 introduced by electroporation into *Y. pestis* CO92(pKOBEG-*sacB*), as described previously
156 (12). Recombinant colonies were selected on Km agar plates. Correct insertion of the

157 antibiotic cassettes was verified by PCR with primers located on each side of the inserted
158 fragment (Table S3).
159
160 **Animal infections.** Prior to infection, the presence of the three plasmids was systematically
161 checked by PCR with primer pairs 157A/157B (*caf1*), 159A/159B (*pla*) and 160A/160B
162 (*yopM*). The presence of the High-Pathogenicity Island was determined by plating the *Y.*
163 *pestis* clones on Congo red agar plates. The animals used for subcutaneous (sc) and aerosol
164 infections were 5 weeks old female OF1 mice (Charles River). The animals were infected sc
165 with 10 cfu/ml of bacterial suspensions from cultures grown for 48h at 28°C on LB agar
166 plates. Aerosol infections were induced by exposing mice for 10 min to an atmosphere
167 containing droplets of a *Y. pestis* suspension (10^8 CFU/ml of saline). Droplets (1 μm median
168 size) were generated by a Raindrop® nebulizer (Tyco, http://www.tyco.com) and were
169 introduced in an all-glass aerosol chamber deviced for a nose-only exposure of animals (34).
170 Lethality was recorded daily for 3 weeks. All animal experiments were performed in a BSL3
171 animal facility.
172
173 **Flea infections.** *Xenopsylla cheopis* fleas were infected by allowing them to feed on blood
174 containing ~5 x 10^8 CFU/ml of the various *Y. pestis* CO92 derivatives, using a previously
175 described artificial feeding system (25, 26). A sample of 20 female fleas was collected
176 immediately after the infectious blood meal, placed at -80° C, and subsequently used for CFU
177 plate count determinations of the average infectious dose per flea (25, 26, 30). An additional
178 group of 76 to 112 fleas (approximately equal numbers of males and females) that took an
179 infectious blood meal were maintained at 21 °C for four weeks, during which time they were
180 allowed to feed on uninfected mice on days 6, 9, 13, 16, 20, 23, and 27. Immediately after
181 each of these feedings all fleas were individually examined under a dissecting microscope to

182 determine how many had taken a normal blood meal (indicated by the presence of fresh red
183 blood filling the midgut) and how many were blocked (indicated by the presence of fresh red
184 blood only in the oesophagus, anterior to the proventriculus) (26). A matched group of
185 uninfected control fleas that fed on blood containing no bacteria was routinely included in
186 each infection experiment. Excess mortality (% mortality of infected fleas) – (% mortality of
187 uninfected control fleas) was recorded during the four week period after the infectious blood
188 meal. After four weeks, a final sample of 20 surviving female fleas was collected to determine
189 the infection rate and average CFU per infected flea (25, 26).
190

191 **RESULTS AND DISCUSSION**

192 **Delineation of the *Y. pestis*-specific chromosomal loci.** A previous *in silico* comparative
193 analysis of the genomes of *Y. pestis* (strains CO92 and KIM10+) and *Y. pseudotuberculosis*
194 (strain IP32953) identified 21 chromosomal clusters specific to the two *Y. pestis* genomes (8).
195 These loci were either single coding sequences (CDS) or clusters of two or more adjacent
196 genes (designated R_{pe}, for *Y. pestis*-specific regions). When the presence of these 21 clusters
197 was previously tested in a panel of 22 *Y. pestis* and *Y. pseudotuberculosis* isolates, only six
198 clusters were found to be present in all *Y. pestis* and absent from all *Y. pseudotuberculosis*
199 strains examined (8).

200 In this study we wanted to perform an in depth analysis of chromosomal regions that
201 specify *Y. pestis* in order to better understand how they might have been acquired and whether
202 they participated to the change in life cycle and pathogenicity of the plague bacillus. These
203 regions could have been acquired by *Y. pestis* after its divergence from *Y. pseudotuberculosis,*
204 but it is also possible that a few of them were horizontally acquired by the specific *Y.*
205 *pseudotuberculosis* strain from which the plague bacillus emerged, because their presence
206 was a prerequisite for *Y. pestis* emergence. Furthermore, the *Y. pestis* genome is known to

207 undergo rearrangements that lead to loss of genetic material, as previously shown for the High
208 Pathogenicity Island (HPI)(10) and the filamentous phage YpfΦ (11). Taking all these
209 considerations into account, we considered here as *Y. pestis*-specific, a locus present in most
210 *Y. pestis* strains and absent from almost all *Y. pseudotuberculosis* isolates studied.

211 To identify all potentially critical regions, we proceeded in four steps: (i) the *in silico*
212 comparative analysis of the genomes of CO92 and IP32953 was re-run, (ii) new primers were
213 designed and used to screen the same set of 22 *Y. pestis* and *Y. pseudotuberculosis* isolates
214 previously examined (8) to confirm the specificity of the regions studied, (iii) less stringent
215 criteria were used to select chromosomal loci that might have participated to the emergence of
216 the plague bacillus, and (iv) the specificity and conservation of the selected loci were then
217 tested across a large panel of 76 additional strains of the two species. Since one of these loci
218 corresponds to the *Y. pestis* filamentous phage YpfΦ which has already been the subject of an
219 in depth analysis (11), this genomic region was not included in the specific regions analyzed
220 here.

221 Our *in silico* genomic comparison of CO92 and IP32953 allowed us to identify, in addition
222 to the 21 clusters previously found, two additional *Y. pestis* CDS, YPO1899 and YPO2469,
223 that were absent from the genome of *Y. pseudotuberculosis* IP32953. These two CDS were
224 not retained in the initial comparative analysis (8), probably because their predicted coding
225 potential was low. However, since their functionality is unknown, we decided to keep them
226 for further screening. The first PCR screening was then done to confirm the presence or
227 absence of these 23 loci in the 13 *Y. pestis* belonging to the evolutionary branches 1 and 2 (1),
228 and the 9 *Y. pseudotuberculosis* of serotypes I to V previously analyzed, with 51 pairs of
229 primers newly designed and located within one or more ORFs of each locus (Table S1). We
230 considered that the absence of a locus in a few *Y. pestis* strains was not sufficient to exclude it
231 from the list of specific regions. As explained above, the transformation of *Y.*

232 *pseudotuberculosis* into *Y. pestis* was probably gradual and continuous, and consequently the
233 presence of a locus in a single isolate of *Y. pseudotuberculosis* was not considered as a
234 redhibitory condition for selection. Based on these criteria, 10 of the 23 initial loci were not
235 further considered as *Y. pestis*-specific (Table S1). Despite the presence of two non-specific
236 ORFs (YPO3946 and YPO3947) in locus 21, this region was kept for further analysis because
237 it also contained two specific ORFs (YPO3945 and YPO3948). Altogether, 13 loci were thus
238 retained for the second screening (Table S1). These included the five *Y. pestis*-specific
239 clusters previously identified (8) as well as eight additional loci, of which two corresponded
240 to the two CDS newly detected by our *in silico* comparative analysis.

241 The specificity of these 13 loci was further investigated using the same set of primers to
242 screen by PCR a panel of 46 *Y. pestis* and 30 *Y. pseudotuberculosis* additional strains (Table
243 S2). This screen showed that the genome of some *Y. pestis* strains (for instance IP539, IP537,
244 IP540 and IP566) was prone to frequent rearrangements leading to a significant loss of
245 genetic material. These strains were isolated during the 1950's, and their genomic
246 rearrangements most likely occurred upon prolonged storage. It also appeared that some loci
247 (especially YPO2469, YPO2484 and YPO3616) were frequently lost (while others were
248 systematically conserved), suggesting both that they are located in an unstable chromosomal
249 region and that these regions do not carry genes essential for *in vitro* survival of *Y. pestis*.
250 Since the number of isolates was large in the second screen, we decided to retain for further
251 physiological analyses the loci for which ≥90% of the 39 *Y. pseudotuberculosis* strains
252 studied during the first and second screens gave a negative or weak signal. As shown in Table
253 S2, eight loci corresponding to six regions (designated $R1_{pe}$ to $R6_{pe}$) and two CDS ($CDS1_{pe}$
254 and $CDS2_{pe}$) met this criterion.

255 After our screening was performed, the genome sequences of three other *Y.*
256 *pseudotuberculosis* strains (IP31758 (14), PB1+

257 (http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genomeprj&cmd=ShowDetailView&TermTo
258 Search=28745), and YPIII
259 (http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genomeprj&cmd=ShowDetailView&TermTo
260 Search=28743) became available. Blast analyses were performed between these sequenced
261 genomes and the nucleotide sequences of the eight *Y. pestis*-specific loci selected. This
262 genome analysis confirmed the results of the PCR screen, except for YPO3945 in $R6_{pe}$ (Table
263 S2). Indeed, a similar sequence (97% nt identity) was found in the genomes of *Y.*
264 *pseudotuberculosis* IP31758 and YPIII, although it lacked the start codon present in the *Y.*
265 *pestis* sequence. A portion of YPO3945 was also found in the PB1+ genome. Sequence
266 analysis of YPO3945 revealed that one of the primers we used to amplify this locus was
267 located in a region exhibiting a certain degree of polymorphism, which could explain our
268 negative PCR results. This agrees with the results of the previous study (in which other
269 primer pairs were used), showing the presence of this sequence in several *Y.*
270 *pseudotuberculosis* strains (8). Therefore, YPO3945 should not be considered as *Y. pestis*-
271 specific. However, based on certain pecularities of $R6_{pe}$ that will be described below, we
272 decided to keep the entire region, and not only YPO3948, for future analyses.
273
274 **Organization and putative functions of the *Y. pestis*-specific chromosomal regions.** The
275 results of Blast analyses performed on the deduced amino acid sequences of each region and
276 CDS indicated that the majority of the coding sequences are predicted to encode hypothetical
277 proteins or proteins of unknown functions (Table S5). An in depth *in silico* analysis of the
278 organization and putative function of each specific chromosomal locus, along with a
279 comparison with the homologous *Y. pseudotuberculosis* regions in which each *Y. pestis*-
280 specific locus has inserted itself, were then performed.

281 CDS1$_{pe}$ (YPO1899) is located at one extremity of the *Y. pestis* High Pathogenicity Island
282 (HPI) (7), and more specifically at the 5' end of the IS*100* copy (Fig. 1A). This locus most
283 likely corresponds to region A, previously found to be specific for the *Y. pestis* HPI (21). Its
284 putative start codon is situated within the inverted repeat of IS*100.* The YPO1899 predicted
285 product shares no significant similarity with proteins in data bases, but its DNA sequence
286 contains a 159 bp region similar (70% nt identity) to the iso-IS*630* transposase of *Salmonella*
287 *choleraesuis* (Genbank D10689). To determine whether YPO1899 could be a functional
288 coding sequence, we performed a reverse transcription assay (RT-PCR) on total RNA
289 extracted from *Y. pestis* CO92. No mRNA was detected (data not shown), arguing against its
290 functionality. It is thus possible that this CDS is a remnant of an insertion sequence that, like
291 IS*100*, inserted itself into a hot spot of IS integration. This is reminiscent of the local hopping
292 of IS*3* elements found at one extremity of the *Y. enterocolitica* HPI (43).

293 CDS2$_{pe}$ (YPO2469) is located within a region known as Clustered Regularly Interspaced
294 Short Palindromic Repeats (CRISPRs, Fig. 1B), which consists in 28 bp-long repeated
295 sequences separated by 32-33 bp unique spacers (28). Addition of new spacers renders these
296 structures highly plastic and generates a polymorphism which can be used for phylogenetic
297 analyses (41). The fact that several *Y. pestis* strains were negative during our PCR screens
298 (Table S2) suggests that they had lost (or not acquired) the corresponding spacers. In CO92,
299 the CRISPR contains nine repeats whereas in IP32953 there are 18 repeats (Fig. 1B).
300 However, the sequences in the CO92 and IP32953 spacers are different, indicating a different
301 origin. The YPO2469 open reading frame (ORF) overlaps several 28 bp-long repeated
302 sequences and its start and stop codons are located inside specific spacers (Fig. 1B). Because
303 of the peculiar features of YPO2469, we also performed RT-PCR on total RNA extracted
304 from CO92 to see whether this putative ORF is transcribed. A transcript of the expected size
305 was indeed detected (data not shown), thus confirming that, despite its unusual location,

306 $CDS2_{pe}$ is transcribed. A putative function could not be attributed to $CDS2_{pe}$ because it has no
307 homologs in data bases. CRISPR were shown to provide resistance against phage infections
308 in prokaryotes (5), but such a function has not been demonstrated in *Y. pestis*.

309 $R1_{pe}$ is a 11.7 kb-long region which carries 11 ORFs (YPO0387-0397) in the two
310 orientations (Fig. 1C). The loci on each side of $R1_{pe}$ are not adjacent on the *Y.*
311 *pseudotuberculosis* IP32953 chromosome. It is known that despite the very close genetic
312 relationship between *Y. pestis* and *Y. pseudotuberculosis*, their chromosomes are not colinear
313 (8), most probably because of frequent rearrangements between the numerous IS present on
314 the *Y. pestis* chromosome. The presence of an IS*100* copy adjacent to $R1_{pe}$ (or in $R1_{pe}$)
315 suggests that the integration event was followed by IS-mediated chromosomal rearrangements
316 in *Y. pestis.* The very low GC content of the central ORFs in $R1_{pe}$ is suggestive of a mosaic
317 structure resulting from successive integration events. This hypothesis is reinforced by the
318 presence of remnants of transposable elements (IS*R1* and IS*21*) in this region (Table S5).
319 Most of the genes on $R1_{pe}$ are predicted to encode products with low or no homology with
320 known proteins. However, three may play a role in DNA restriction/modification. One of
321 them, located in the low GC% portion of $R1_{pe}$, would be a methylase (YPO0391), while two
322 adjacent sequences (YPO0387-0388) at the left-hand extremity of $R1_{pe}$ have the potential to
323 encode a complete restriction/modification system similar to the McrB system of *E. coli*
324 (Table S5). Remarkably, R1 appears to have been exchanged with a *Y. pseudotuberculosis*-
325 specific region (YPTB0535-0537) coding also for a restriction/modification system. This
326 region was shown to play a major role in *Y. pseudotuberculosis* pathogenicity (40).

327 $R2_{pe}$ is a 7.8 kb region composed of five CDSs (YPO1668-1672, Fig. 1D). This region has
328 inserted itself between YPTB2403 and YPTB2404 in the genome of the *Y.*
329 *pseudotuberculosis* ancestor. No mobile elements adjacent to $R2_{pe}$ could explain its
330 integration site into the *Y. pestis* genome. The YPO1672 ORF located at the right-hand

331 extremity of $R2_{pe}$ is predicted to encode a large protein of 1,268 aa which has two paralogs on
332 the CO92 chromosome: YPO0765 (82% id. in 1,206 aa) and YPO0309 (81% id. in 1,199 aa).
333 This identity concerns the last two thirds of the encoded products while their NH2 extremities
334 are different. The conserved COOH portion of these proteins exhibits homology with a
335 domain PL2-passenger and an autotransporter beta-domain, suggesting that they encode
336 autotransporters proteins of a type V secretion system. Many proteins secreted by type V
337 secretion systems have been linked to virulence in Gram-negative bacteria (22, 23). Recently,
338 one *Y. pestis* autotransporter adhesin (YapE encoded by YPO3984) has been shown to be
339 required for efficient colonization of the host during bubonic plague (33). The other ORFs
340 (YPO1668 to YPO1671) carried by $R2_{pe}$ share similarity (63% to 85% id.) with CDSs of
341 *Proteus mirabilis* that exhibit the same gene organization and that are putatively forming a
342 sugar transport system (Table S5). Interestingly, one of the *Y. pseudotuberculosis*-specific
343 regions ($R5_{\mathbf{pst}}$) that was shown to have been lost during the emergence of *Y. pestis,* is also
344 putatively involved in sugar transport (39). However, this *Y. pseudotuberculosis*-specific
345 region was not found to be important for the physiology or pathogenicity of this species.

346 $R3_{\mathbf{pe}}$ is a very large region (41.7 kb) composed of 52 ORFs (YPO2084-2136), which is
347 predicted to be a lambda prophage (Fig. 1E and Table S5). This phage inserted itself between
348 YPTB2404 and YPTB2405 in the genome of the *Y. pseudotuberculosis* ancestral strain,
349 probably using its own machinery of insertion via a mechanism of site specific
350 recombination. Indeed, an identical 31 nt repeat
351 (ATATTGAATAAGCCGAGTTGGGCTAAATAAC) is found at each extremity of the
352 CO92 prophage. This 31 nt sequence is found in two copies flanking the prophage in all *Y.*
353 *pestis* sequenced genomes while it is present in a single copy in the *Y. pseudotuberculosis*
354 genomes. This suggests that the 31 nt sequence served as an *att* site for phage integration and,
355 as classically observed during chromosomal insertion of lambda phages, resulted in a

356 duplicated sequence upon integration. This putative prophage also carries at its left-hand
357 extremity two genes predicted to form the phage insertion/excision machinery, i.e. an
358 integrase (YPO2084) and an excisionase (YPO2087) (Table S5). However, the *int* gene is
359 interrupted by the insertion of an IS*100* element which most likely inactivates it, suggesting
360 that this phage is now stabilized in the *Y. pestis* chromosome. Another IS element (IS*285*) is
361 found at the other extremity of the prophage and has most likely inserted itself secondarily
362 into the phage genome. In addition to *int*, another prophage-borne gene (YPO2103) predicted
363 to encode a terminase, which is required for cleavage of the *cos* site and therefore for the
364 encapsidation of the phage DNA, is interrupted by an IS*285* insertion. Finally, the YPO2106
365 CDS (of unknown function) is truncated because of a frameshift in its nt sequence. By
366 comparing the genetic organization of the *Y. pestis* lambdoid prophage with that of the well
367 studied *E. coli* lambda phage, it appears that the regions coding for phage morphogenesis are
368 similar, while the loci involved in lambda replication, regulation and recombination are absent
369 or divergent in the *Y. pestis* phage. This along with the presence of several truncated genes
370 and the fact that the size of the phage is drastically smaller in some *Y. pestis* strains such as
371 Antiqua (28.6 kb), Angola (23.6 kb) or Microtus (8.9 kb), suggest that the *Y. pestis* lambda
372 prophage is non-functional and is undergoing a process of reductive evolution. Interestingly,
373 the lambdoid phage the closest to that of *Y. pestis* is found in the genome of *Y. enterocolitica*
374 Ye8081 (53). Indeed, 25 of the $R3_{pe}$ ORFs have homologs in a region of the Ye8081 genome
375 predicted to be a prophage. Despite a high level of conservation of some genes, some other
376 genes are much less conserved and some regions are different in the two prophages (Table
377 S5), suggesting that they belong to the same phage family but that they are different or
378 diverged long before infecting independently the *Y. enterocolitica* and *Y. pestis* branches.
379

380 $R4_{pe}$ is a 2.7 kb region composed of four ORFs (YPO2483-2486), all in the same
381 orientation (Fig. 1F). $R4_{pe}$ is flanked by two 98% identical 250 nt-long DNA repeats that are
382 inside the 5' portions of two coding sequences : one in YPO2486 at the left-hand extremity of
383 $R4_{pe}$, and the other one in YPO2482, bordering the right-hand extremity of $R4_{pe}$. Remarkably,
384 a segment sharing 97% and 98% nt identity with the left and right 250 nt repeats, respectively
385 is also found within the YPTB2524 coding sequence on the *Y. pseudotuberculosis*
386 chromosome (Fig. 1F). This is highly suggestive of the insertion of a circular DNA molecule
387 by homologous recombination between the 250 nt segment in YPTB2524 and the same
388 sequence carried by the incoming replicon. This integration event has probably been
389 accompanied or followed by secondary rearrangements at the site of integration since a single
390 recombination event would not be sufficient to explain the genetic organization of this region
391 in *Y. pestis.* As could be expected for an integration by a single crossing over, two fusion
392 ORFS were generated at each extremity: one composed of the entire YPTB2524 sequence
393 followed by a sequence brought by the replicon, creating YPO2486, and another mosaic ORF
394 (YPO2482) composed of a portion of YPTB2524 fused to a portion of the downstream
395 YPTB2523 CDS (Fig. 1F). Of note, another copy of the 250 nt DNA segment is found within
396 another CDSs in *Y. pestis* (YPO2490) and in *Y. pseudotuberculosis* (YPTB2527). Both these
397 CDSs have the characteristics of being located in the vicinity of the $R4_{pe}$ integration site and
398 of coding for very large proteins predicted to be hemolysins. YPO2486 also shares 38%
399 similarity with an hemolysin/hemagglutinin-like protein HecA precursor from *Erwinia*
400 *chrysanthemi* (44). The other three ORFs carried by $R4_{pe}$ (YPO2483-2485) are of unknown
401 functions. They all have a GC% lower than 39 and share little or no homology with proteins
402 in data bases, suggesting that they originate from a distantly related donor strain.

403 $R5_{pe}$ is a 3 kb-long region composed of two ORFs (YPO3436-3437). Elements involved in
404 gene mobility are found on each side: YPO3438 at the right-hand extremity encodes a

405 putative integrase, and at the left-hand extremity YPO3436 is truncated by the insertion of an
406 IS*100* element, which is itself followed by an IS*1541* copy (Fig. 1G). As already observed for
407 $R1_{pe}$, the loci on each side of $R5_{pe}$ are not adjacent on the *Y. pseudotuberculosis* IP32953
408 genome. Possibly, $R5_{pe}$ was part of a larger mobile element and its integration into the *Y.*
409 *pestis* chromosome was associated with intra-genomic rearrangements via IS*100*
410 recombination. Similarly to the other two CDS, the putative function of $R5_{pe}$ remains
411 unknown because its amino acid sequence shares no significant similarity with proteins in
412 data bases.

413 $R6_{pe}$ is 1.9 kb-long and is composed of four small ORFs (YPO3945-3948) in the two
414 orientations (Fig. 1H). This region inserted itself between YPTB3789 and YPTB3790 in the
415 ancestral *Y. pseudotuberculosis* strain. Interestingly, a perfectly conserved 15 nt-long
416 sequence (GCGAATTCTGCTGTA) is found both between YPTB3789 and YPTB3790 on
417 the *Y. pseudotuberculosis* IP32953 chromosome, and at the two extremities of $R6_{pe}$ on the *Y.*
418 *pestis* CO92 chromosome (Fig. 1H). This is consistent with the insertion of a foreign DNA by
419 a mechanism of site specific recombination. The analysis of the three other *Y.*
420 *pseudotuberculosis* sequenced genomes indicates that all three have an element inserted at the
421 same site, flanked by the 15 nt repeat (Fig. S1). This argues for the 15 nt sequence being a hot
422 spot of integration of foreign DNA. This 15 nt repeat is identical to that of *Y. pestis* on one
423 side, and is slightly divergent (caGtAATCTGCTGTA) on the other side. The imported DNA
424 has a common backbone (represented by strain YPIII), in which different DNA sequences
425 have inserted themselves, either before or after integration at the 15 nt site. Remarkably, in *Y.*
426 *pseudotuberculosis* IP31758, the incoming DNA has created a large ORF, either by fusion
427 with the two bordering sequences, or more likely by allelic replacement between these
428 sequences and a larger ORF having the same extremities. Also remarkable is the observation
429 that in *Y. pestis* YPO3945, a potential start codon is located just upstream of the 15 nt repeat

430 of *Y. pestis,* while in *Y. pseudotuberculosis*, the slight divergence in the repeat introduces a
431 stop codon which leads to a shorter ORF. Therefore, despite the conservation of the region
432 overlapping YPO3945 in *Y. pestis* and several *Y. pseudotuberculosis* strains, the coding
433 capacity and function of this DNA sequence might be different. No function could be
434 predicted for the four ORFs carried by $R6_{pe}$ based on Blast analyses. From the above data, we
435 can thus hypothesize that $R6_{pe}$ corresponds to a mobile element which can insert itself into
436 bacterial chromosomes by site specific recombination, and which carries a conserved core
437 sequence and a variable region in which different CDSs are found inserted. Whether the *Y.*
438 *pestis* region was horizontally acquired after its divergence from *Y. pseudotuberculosis* or
439 whether it was vertically acquired from an ancestor that had an identical region cannot be
440 determined. However, the fact that none of the screened *Y. pseudotuberculosis* strains
441 harbored YPO3948 and that the 15 nt right-hand repeat is different between the two species
442 argues for the first hypothesis.

443 Altogether, it thus appears that, based on the scars left by the insertion of the eight *Y.*
444 *pestis*-specific chromosomal loci, a mechanism of acquisition can be predicted for most of
445 them. The position of these loci on the *Y. pestis* CO92 chromosome indicates that six out of
446 nine (when the YpfΦ genome is considered) are located in the region opposite to the origin of
447 replication (Fig. 2). Similarly, most of the regions lost by *Y. pestis* (i.e. *Y. pseudotuberculosis*-
448 specific) are also located in this part of the chromosome (Fig. 2). The genetic polymorphism
449 in bacterial chromosomes is known to be the highest in this portion of the chromosome,
450 which contains the *dif* site where resolution of chromosomal dimers takes place during
451 bacterial division (50). Indeed, more than one third of the acquisition or loss of genetic
452 materials (from $R4_{pe}$ to $R2_{pst}$, Fig. 2) in *Y. pestis* are concentrated in a region which represents
453 only 5% of the entire *Y. pestis* chromosome. In two instances, the acquisition of the *Y. pestis*-
454 specific loci seems to have occurred by exchange with a *Y. pseudotuberculosis*-specific region

455 ($R3_{pst}$ replaced by YpfΦ, and $R6_{pst}$ by $R1_{pe}$). In terms of functions, with the exception of $R1_{pe}$,
456 $R2_{pe}$, and $R3_{pe}$, which putatively encode a restriction/modification system, a sugar transporter
457 and a lambdoid phage, respectively, no specific function could be attributed to the other five
458 loci. It is noteworthy that no putative virulence factor was identified on any of the eight *Y.*
459 *pestis*-specific chromosomal loci studied.
460
461 **History of the acquisition and modification of the specific regions during *Y. pestis***
462 **microevolution.** We wondered whether the eight specific regions studied here were acquired
463 early or late after *Y. pestis* emergence from *Y. pseudotuberculosis*. A previous molecular
464 phylogenetic study allowed us to design a microevolution tree for *Y. pestis* and to delineate a
465 set of Single Nucleotide Polymorphisms (SNPs) that could be used to position a new isolate
466 on this tree (1). Today, the complete genome sequences of 21 *Y. pestis* strains (Microtus
467 91001, pestoides A, pestoides F, Angola, Antiqua B42003004, Nepal 516, KIM10,
468 K1973002, Antiqua, UG05-0454, E1979001, D106004, D182038, IP275, CA88-4125,
469 F1991016, MG05-1020, FV-1, India 195, PEXU 2 and CO92) are available (listed in
470 http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?database=386656), but their position
471 along the microevolutionary tree is not known for most of them. To draw an updated tree, a
472 subset of 16 SNPs defining the main branches and sub-branches were selected among those
473 previously identified (1), and their presence/absence was checked in the 21 genomes. This
474 allowed us to position all strains for which no branching was yet assigned in the *Y. pestis*
475 updated microevolution tree (Fig. 3). The position of *Y. pestis* pestoides A was confirmed
476 using a second set of SNPs along the 0.PE4 branch. Furthermore, based on the presence or
477 absence of the s1 SNP, the 1.ANT strains could be differentiated into two sub-branches (the
478 $1.ANT_1$ and $1.ANT_2$, Fig. 3). The presence of the eight *Y. pestis*-specific loci was then
479 examined in these 21 *Y. pestis* genomes.

480 $CDS2_{pe}$ is a peculiar case because it is located within a CRISPR, which has the
481 characteristics of being highly plastic and of generating a polymorphism (41). We could
482 nonetheless observe in our PCR screens that this CDS is present in only a portion of Antiqua
483 and Medievalis strains, while it is found in all 23 Orientalis strains analyzed (Table S2). All
484 seven other specific loci were present in the genomes of all 21 sequenced *Y. pestis* isolates,
485 including Pestoides strains which were not available for our PCR screens (Fig. 3). This
486 indicates that these seven loci were acquired very early during *Y. pestis* evolution, before the
487 split of the most ancient branch (0.PE4) along branch 0. It also indicates that these regions
488 were stably kept during the microevolution of the plague bacillus.

489 In a previous analysis of *Y. pestis* housekeeping genes, we were unable to identify a single
490 polymorphic nt at 21,881 synonymous sites within 36 *Y. pestis* strains because this species is
491 remarkably monomorphic (2). We wondered whether this was also the case for the seven
492 horizontally acquired regions, or whether peculiar mutations could specify some evolutionary
493 branches. A detailed analysis of these regions at the gene level indicated a very high degree of
494 nt conservation, with nonetheless some polymorphisms. $CDS1_{pe}$ was completely
495 monomorphic in the 21 genomes analyzed and $R5_{pe}$ was also identical in 20 genomes except
496 in Angola, which is known to be a very atypical strain (15). Among various reasons that could
497 explain this absence of polymorphism, one could be that these two regions were subjected to
498 a strong selective pressure because they encode functions important for *Y. pestis* life cycle.
499 However, another and more likely explanation is that, because of their small size, the chances
500 of acquiring point mutations or deletions/insertions are lower.

501 For the other regions, a few sequence variations were observed, and some of these
502 variations were informative in terms of *Y. pestis* microevolution. Indeed several
503 SNPs/deletion in $R1_{pe}$ and $R3_{pe}$ were characteristic of branch 0 and of the 0.PE4, and 0.PE2
504 branches that split off along this branch early during evolution (Fig. 3 and Table S6). SNPb in

505 $R1_{pe}$ occurred before the separation between branches 1 and 2 and thus characterizes all
506 classical *Y. pestis* strains of branches 1 and 2. Along branch 2, one SNP in $R3_{pe}$ defines the
507 2.MED sub-branch. On branch 1, one SNP in $R2_{pe}$ (Table S6) specified the 11 strains (three
508 $1.ANT_2$ and eight 1.ORI) the most recently emerged along this branch 1 and confirmed the
509 1.ANT split evidenced by the s1 SNP (Fig. 3). Curiously, the older $1.ANT_1$ sub-branch was
510 characterized by mutations, including inactivating deletions (Table S6), in three specific
511 regions ($R1_{pe}$, $R3_{pe}$ and $R6_{pe}$, Fig. 3).

512 Overall, $R1_{pe}$ and $R3_{pe}$ were the two specific regions in which the highest rates of genetic
513 polymorphisms were observed. The fact that these regions are the two largest in size could
514 explain at least in part their higher polymorphism. $R3_{pe}$ corresponds to the lambdoid prophage
515 genome. The fact that deletions of large regions predicted to encode phage functions required
516 to infect a bacterial host occurred in different branches (YPO2095 to YPO2135 in branch
517 0.PE4, and YPO2087 to YPO2103 in $1.ANT_1$ sub-branch, Fig. 3) suggests a single
518 acquisition event and further argues for a secondary process of reductive evolution. This,
519 along with our above observation that some genes necessary for the phage physiology are
520 inactivated in all strains support the hypothesis that a functional phage is not necessary for *Y.*
521 *pestis* life cycle and/or virulence. The polymorphism of $R1_{pe}$ was the highest but
522 corresponded mostly to SNPs that occurred in different genes without inactivating them
523 (Table S6). Even the small 6 nt deletion that arose early along branch 0 (but which could also
524 be a 6 nt insertion in branch 0.PE4) was in frame and therefore did not disrupt the coding
525 sequence. Except for the presence of a putative DNA restriction/modification system, most of
526 the genes carried by $R1_{pe}$ have no predicted functions. The absence of inactivating mutations
527 in this region, whatever the genes involved and the evolutionary branch, supports the
528 hypothesis that this region may be useful for *Y. pestis* physiological or pathological processes.
529

530 **Mutagenesis of the *Y. pestis* acquired regions and impact on bacterial physiological and**
531 **pathogenic properties.** To investigate the role of the eight specific chromosomal loci on the
532 ability of *Y. pestis* to grow *in vitro*, to infect fleas and to kill mice, each locus was individually
533 deleted by allelic exchange with a kanamycin cassette. The ORFs encompassed by the
534 deletions and the primers used to check that the replacements occurred as expected are
535 indicated on Fig. 1 and in Table S4. The resulting deletion strains with their designation are
536 listed in Table S3. To perform the deletions using the lambda red technology, the pKOBEG-
537 *sacB* plasmid was introduced into *Y. pestis* strain CO92, yielding CO92p (Table S3). Since
538 the chromosome of *Y. pestis* is prone to frequent rearrangements (20), and since we found that
539 the presence of pKOBEG-*sacB* had no significant impact on *Y. pestis* pathogenicity, we
540 decided not to remove this plasmid in the deletion mutants, in order to limit as much as
541 possible subcultures which would increase the occurrence of genomic reassortments. During
542 our previous analysis of *Y. pseudotuberculosis*-specific regions, we found that two CDSs
543 (YPTB1495 and YPTB3368) were not deletable because they are essential for IP32953
544 survival (40). This was not the case for the *Y. pestis*-specific loci as the eight specific regions
545 were successfully deleted.

546 To determine whether some of these loci play a role in the ability of *Y. pestis* to multiply *in*
547 *vitro*, the CO92p parental strain and the eight deletants were cultured in LB at 28°C (optimal
548 temperature for *Yersinia* growth) and 37°C (*in vivo* temperature). No growth defects were
549 observed in the various mutants as compared to the CO92p parental strain (data not shown).
550 Similar results were previously obtained with the various *Y. pseudotuberculosis* strains
551 deleted of each specific region (40). However, when these *Y. pseudotuberculosis* deletants
552 were grown in the biochemically defined medium M63S, a growth defect could be observed
553 for several mutants. The ability of the eight *Y. pestis* deletants and the parental CO92p strain
554 to grow in the M63S medium at 28°C and 37°C was thus tested and compared. The growth

555 curves of all strains were comparable, both at 28°C and 37°C (data not shown). Therefore,
556 none of the eight *Y. pestis*-acquired chromosomal loci are required for optimal bacterial
557 growth, at least under the experimental conditions used here.

558 A major difference during the transformation of *Y. pseudotuberculosis* into *Y. pestis* has
559 been a switch in the mode of transmission, from the oral route to flea bites. It could thus be
560 expected that some of the chromosomal regions that specify *Y. pestis* participated to its
561 adaptation to its new transmission cycle. This has been demonstrated for the two *Y. pestis*-
562 specific plasmids: pFra is necessary for an efficient colonization of and survival in the flea
563 midgut (27), as well as for an efficient transmission by fleas (46), whereas pPla is important
564 for the dissemination of the bacteria from the intradermal inoculation site to the draining
565 lymph node (47, 49). A key feature of *Y. pestis* that facilitates flea-borne transmission, and
566 that differentiates it from *Y. pseudotuberculosis,* is its capacity to block fleas. This blockage is
567 due to the formation of a biofilm produced by the products of the chromosomal *hms* genes
568 (29). Although *Y. pseudotuberculosis* harbors similar *hms* genes, this species does not block
569 fleas (16), indicating that other genetic determinants participate to this activity. To determine
570 whether the acquired regions could play a role in the capacity of *Y. pestis* to infect and block
571 fleas, groups of 76 to 112 *X. cheopis* were artificially fed on blood containing similar inocula
572 ($1.5x10^8$ to $1x10^9$ CFU/ml) of each of the eight deletants or CO92p, and monitored for
573 blockage during a four week period after infection. As shown in Fig. 4A, blockage developed
574 in all of the flea groups at a rate expected for the *X. cheopis*-*Y. pestis* infection model (26, 27),
575 indicating that none of the *Y. pestis*-specific loci is required for biofilm development in the
576 flea proventriculus. Because fleas that become blocked can no longer take a blood meal, they
577 rapidly starve to death. Thus, excess mortality can be used as an indirect indicator of
578 blockage. In keeping with their ability to block fleas, all groups of infected fleas had elevated
579 excess mortality rates of 15-58%, which is typical for *X. cheopis* fleas infected with *Y. pestis*

580 (27). The eight mutants also were able to produce a stable, persistent infection in the flea at
581 rates equivalent to that of CO92p (Fig. 4B). The average bacterial load in infected fleas after 4
582 weeks was high for all the *Y. pestis* strains tested, ranging from $1.2x10^4$ to $7.9x10^5$ CFU/flea.
583 In *X. cheopis*, it has been shown that *Y. pestis* has gained the capacity to colonize the flea
584 gut without causing major damages during the early phase of infection, whereas its *Y.*
585 *pseudotuberculosis* progenitor exerts acute toxic effects on these insects (17). None of the
586 eight deletants or CO92p caused any acute toxicity to the fleas that took an infectious blood
587 meal, indicating that none of the *Y. pestis*-specific chromosomal loci play an essential role in
588 the lower toxicity of *Y. pestis* for fleas.
589 One of the major differences between *Y. pestis* and its ancestor *Y. pseudotuberculosis* is the
590 acquisition by the former of an exceptional pathogenicity potential. Although the horizontally
591 acquired pPla plasmid plays a major role in the virulence of most *Y. pestis* isolates, the fact
592 that it is dispensable in some strains (32, 45, 55) and that its introduction into *Y.*
593 *pseudotuberculosis* is not accompanied by an increase in virulence (31, 38) clearly indicates
594 that other factors are required. One of them has been shown to be the YpfΦ filamentous phage
595 (11), but other *Y. pestis*-acquired chromosomal loci may have also played a role in this gain of
596 virulence. However, using the mouse model of subcutaneous infection to mimic bubonic
597 plague, the eight mutants were found to be as virulent as the parental strain CO92p (Table 1).
598 *Y. pestis* has also acquired the capacity to be transmitted by aerosols and to cause pneumonic
599 plague. The capacity of the eight mutants to produce a lethal infection upon aerosol
600 inoculation was tested by exposing mice to aerosolized droplets of a suspension containing
601 10^8 CFU/ml (three times the LD_{50}). All deletants killed 100% of the infected animals with a
602 mean time to death similar to that of the CO92p strain (Table 1), Altogether, these results
603 suggest that the acquired regions and CDSs studied here did not participate in the increased

604 pathogenicity of *Y. pestis,* at least in our experimental mouse models of sc and aerosol
605 infections.
606
607 **Conclusions.** Our study has identified eight chromosomal loci that were acquired soon before
608 or early after the divergence of *Y. pestis* from *Y. pseudotuberculosis*. Surprisingly, we were
609 unable to evidence an impact of these regions under *in vitro* and *in vivo* conditions. It is
610 possible that acquisition of these regions provided no benefit to the bacterial host and,
611 because *Y. pestis* is a recently emerged bacterium, that they did not have yet the time to be
612 inactivated and eliminated by gradual mutations and/or deletions. This could be the case for
613 the lambdoid phage ($R3_{pe}$) in which several genes predicted to encode physiological functions
614 are interrupted, as well as for $CDS1_{pe}$ which is probably a remnant of an insertion element and
615 which is no longer transcribed. This is also true for many other genes in *Y. pestis* whose
616 genome is characterized by a massive gene inactivation (8). However, the fact that the other
617 regions, which were acquired early during *Y. pestis* microevolution, are conserved and
618 potentially functional in the various evolutionary sub-branches, suggests that they have been
619 subjected to a positive selective pressure for their maintenance, and thereby that they should
620 confer advantages at some steps of the *Y. pestis* biological cycle. If so, our inability to
621 attribute a physiological role to these loci may have several reasons. One possible explanation
622 would be that the functions of two or more acquired regions are redundant. Multiple deletions
623 of the *Y. pestis*-acquired loci would thus be needed to observe a defect. In the flea model,
624 although no effect on toxicity or proventriculus blockage was observed, it is still possible that
625 the deletion of some loci had an impact on the efficiency of flea transmission, as recently
626 observed for the F1 capsule (46). Furthermore, since *Y. pestis* can be transmitted by many
627 different flea species, some of which do not develop proventricular blockage to the same
628 extent as *X. cheopis* (13), testing the various mutants in other flea models might also be of

629 interest. The absence of virulence impairment in the various deletants indicates that none of
630 the *Y. pestis*-specific regions play a major role in *Y. pestis* pathogenicity. However, because of
631 the extreme pathogenicity of the plague bacillus in the mouse model, minor virulence defects
632 might be difficult to evidence. The use of more resistant animal models could potentially
633 unable the identification of subtle virulence impediments. Finally, *Y. pestis* has been shown to
634 be able to survive in naturally contaminated soils for several weeks (4) and in burrows of wild
635 rodents for several years (35). The assumption that the acquired regions could participate to
636 the long-term persistence of the plague bacillus outside an insect vector or a mammalian host
637 may also be considered. Whatever the reasons for the absence of identified roles, our results
638 suggest that, unlike what might have been assumed, acquisition of new chromosomal
639 materials has not been of major importance in the dramatic change of life cycle that has
640 accompanied the emergence of *Y. pestis.*
641

642 **Acknowledgements**

643 This study was funded in part by contract DGA 99 01110004709450 from the French Defense
644 Ministry and a fellowship for A. D. from the Pasteur-Weissman foundation.
645

646 **References**

647 1. **Achtman, M., G. Morelli, P. Zhu, T. Wirth, I. Diehl, B. Kusecek, A. J. Vogler, D.
648 M. Wagner, C. J. Allender, W. R. Easterday, V. Chenal-Francisque, P.
649 Worsham, N. R. Thomson, J. Parkhill, L. E. Lindler, E. Carniel, and P. Keim.**
650 2004. Microevolution and history of the plague bacillus, *Yersinia pestis*. Proc Natl
651 Acad Sci U S A **101:**17837-17842.

652 2. **Achtman, M., K. Zurth, G. Morelli, G. Torrea, A. Guiyoule, and E. Carniel.**
653 1999. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia*
654 *pseudotuberculosis*. Proc Natl Acad Sci U S A **96:**14043-14048.
655 3. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and**
656 **D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein
657 database search programs. Nucleic Acids Res **25:**3389-3402.
658 4. **Ayyadurai, S., L. Houhamdi, H. Lepidi, C. Nappez, D. Raoult, and M. Drancourt.**
659 2008. Long-term persistence of virulent *Yersinia pestis* in soil. Microbiology
660 **154:**2865-2871.
661 5. **Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D.**
662 **A. Romero, and P. Horvath.** 2007. CRISPR provides acquired resistance against
663 viruses in prokaryotes. Science **315:**1709-1712.
664 6. **Bocs, S., S. Cruveiller, D. Vallenet, G. Nuel, and C. Medigue.** 2003. AMIGene:
665 Annotation of MIcrobial Genes. Nucleic Acids Res **31:**3723-3726.
666 7. **Buchrieser, C., M. Prentice, and E. Carniel.** 1998. The 102-kilobase unstable region
667 of *Yersinia pestis* comprises a high-pathogenicity island linked to a pigmentation
668 segment which undergoes internal rearrangement. J Bacteriol **180:**2321-2329.
669 8. **Chain, P. S., E. Carniel, F. W. Larimer, J. Lamerdin, P. O. Stoutland, W. M.**
670 **Regala, A. M. Georgescu, L. M. Vergez, M. L. Land, V. L. Motin, R. R.**
671 **Brubaker, J. Fowler, J. Hinnebusch, M. Marceau, C. Medigue, M. Simonet, V.**
672 **Chenal-Francisque, B. Souza, D. Dacheux, J. M. Elliott, A. Derbise, L. J. Hauser,**
673 **and E. Garcia.** 2004. Insights into the evolution of *Yersinia pestis* through whole-
674 genome comparison with *Yersinia pseudotuberculosis*. Proc Natl Acad Sci U S A
675 **101:**13826-13831.

676 9. **Davis, K. J., D. L. Fritz, M. L. Pitt, S. L. Welkos, P. L. Worsham, and A. M.**
677 **Friedlander.** 1996. Pathology of experimental pneumonic plague produced by
678 fraction 1-positive and fraction 1-negative *Yersinia pestis* in African green monkeys
679 *Cercopithecus aethiops*. Arch Pathol Lab Med **120:**156-163.
680 10. **de Almeida, A. M., A. Guiyoule, I. Guilvout, I. Iteman, G. Baranton, and E.**
681 **Carniel.** 1993. Chromosomal *irp2* gene in Yersinia: distribution, expression, deletion
682 and impact on virulence. Microb Pathog **14:**9-21.
683 11. **Derbise, A., V. Chenal-Francisque, F. Pouillot, C. Fayolle, M. C. Prevost, C.**
684 **Medigue, B. J. Hinnebusch, and E. Carniel.** 2007. A horizontally acquired
685 filamentous phage contributes to the pathogenicity of the plague bacillus. Mol
686 Microbiol **63:**1145-1157.
687 12. **Derbise, A., B. Lesic, D. Dacheux, J. M. Ghigo, and E. Carniel.** 2003. A rapid and
688 simple method for inactivating chromosomal genes in Yersinia. FEMS Immunol Med
689 Microbiol **38:**113-116.
690 13. **Eisen, R. J., S. W. Bearden, A. P. Wilder, J. A. Montenieri, M. F. Antolin, and K.**
691 **L. Gage.** 2006. Early-phase transmission of *Yersinia pestis* by unblocked fleas as a
692 mechanism explaining rapidly spreading plague epizootics. Proc Natl Acad Sci U S A
693 **103:**15380-15385.
694 14. **Eppinger, M., M. J. Rosovitz, W. F. Fricke, D. A. Rasko, G. Kokorina, C.**
695 **Fayolle, L. E. Lindler, E. Carniel, and J. Ravel.** 2007. The complete genome
696 sequence of *Yersinia pseudotuberculosis* IP31758, the causative agent of Far East
697 scarlet-like fever. PLoS Genet **3:**e142.
698 15. **Eppinger, M., P. L. Worsham, M. P. Nikolich, D. R. Riley, Y. Sebastian, S. Mou,**
699 **M. Achtman, L. E. Lindler, and J. Ravel.** Genome sequence of the deep-rooted

700 *Yersinia pestis* strain Angola reveals new insights into the evolution and pangenome
701 of the plague bacterium. J Bacteriol.

702 16. **Erickson, D. L., C. O. Jarrett, B. W. Wren, and B. J. Hinnebusch.** 2006. Serotype
703 differences and lack of biofilm formation characterize *Yersinia pseudotuberculosis*
704 infection of the Xenopsylla cheopis flea vector of *Yersinia pestis*. J Bacteriol
705 **188:**1113-1119.

706 17. **Erickson, D. L., N. R. Waterfield, V. Vadyvaloo, D. Long, E. R. Fischer, R.**
707 **Ffrench-Constant, and B. J. Hinnebusch.** 2007. Acute oral toxicity of *Yersinia*
708 *pseudotuberculosis* to fleas: implications for the evolution of vector-borne
709 transmission of plague. Cell Microbiol **9:**2658-2666.

710 18. **Friedlander, A. M., S. L. Welkos, P. L. Worsham, G. P. Andrews, D. G. Heath,**
711 **G. W. Anderson, Jr., M. L. Pitt, J. Estep, and K. Davis.** 1995. Relationship
712 between virulence and immunity as revealed in recent studies of the F1 capsule of
713 *Yersinia pestis*. Clin Infect Dis **21 Suppl 2:**S178-181.

714 19. **Guinet, F., P. Ave, L. Jones, M. Huerre, and E. Carniel.** 2008. Defective innate cell
715 response and lymph node infiltration specify *Yersinia pestis* infection. PLoS One
716 **3:**e1688.

717 20. **Guiyoule, A., B. Rasoamanana, C. Buchrieser, P. Michel, S. Chanteau, and E.**
718 **Carniel.** 1997. Recent emergence of new variants of *Yersinia pestis* in Madagascar. J
719 Clin Microbiol **35:**2826-2833.

720 21. **Hare, J. M., A. K. Wagner, and K. A. McDonough.** 1999. Independent acquisition
721 and insertion into different chromosomal locations of the same pathogenicity island in
722 *Yersinia pestis* and *Yersinia pseudotuberculosis*. Mol Microbiol **31:**291-303.

723 22. **Henderson, I. R., and J. P. Nataro.** 2001. Virulence functions of autotransporter
724 proteins. Infect Immun **69:**1231-1243.

725 23. **Henderson, I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez, and D.**
726 **Ala'Aldeen.** 2004. Type V protein secretion pathway: the autotransporter story.
727 Microbiol Mol Biol Rev **68:**692-744.
728 24. **Hinchliffe, S. J., K. E. Isherwood, R. A. Stabler, M. B. Prentice, A. Rakin, R. A.**
729 **Nichols, P. C. Oyston, J. Hinds, R. W. Titball, and B. W. Wren.** 2003. Application
730 of DNA microarrays to study the evolutionary genomics of *Yersinia pestis* and
731 *Yersinia pseudotuberculosis*. Genome Res **13:**2018-2029.
732 25. **Hinnebusch, B. J., K. L. Gage, and T. G. Schwan.** 1998. Estimation of vector
733 infectivity rates for plague by means of a standard curve-based competitive
734 polymerase chain reaction method to quantify *Yersinia pestis* in fleas. Am J Trop Med
735 Hyg **58:**562-569.
736 26. **Hinnebusch, B. J., R. D. Perry, and T. G. Schwan.** 1996. Role of the *Yersinia pestis*
737 hemin storage (hms) locus in the transmission of plague by fleas. Science **273:**367-
738 370.
739 27. **Hinnebusch, B. J., A. E. Rudolph, P. Cherepanov, J. E. Dixon, T. G. Schwan, and**
740 **A. Forsberg.** 2002. Role of Yersinia murine toxin in survival of *Yersinia pestis* in the
741 midgut of the flea vector. Science **296:**733-735.
742 28. **Jansen, R., J. D. van Embden, W. Gaastra, and L. M. Schouls.** 2002. Identification
743 of a novel family of sequence repeats among prokaryotes. OMICS **6:**23-33.
744 29. **Jarrett, C. O., E. Deak, K. E. Isherwood, P. C. Oyston, E. R. Fischer, A. R.**
745 **Whitney, S. D. Kobayashi, F. R. DeLeo, and B. J. Hinnebusch.** 2004. Transmission
746 of *Yersinia pestis* from an infectious biofilm in the flea vector. J Infect Dis **190:**783-
747 792.

748 30. **Jarrett, C. O., F. Sebbane, J. J. Adamovicz, G. P. Andrews, and B. J.**
749 **Hinnebusch.** 2004. Flea-borne transmission model to evaluate vaccine efficacy
750 against naturally acquired bubonic plague. Infect Immun **72:**2052-2056.
751 31. **Kutyrev, V., R. J. Mehigh, V. L. Motin, M. S. Pokrovskaya, G. B. Smirnov, and**
752 **R. R. Brubaker.** 1999. Expression of the plague plasminogen activator in *Yersinia*
753 *pseudotuberculosis* and *Escherichia coli*. Infect Immun **67:**1359-1367.
754 32. **Kutyrev, V. V., A. A. Filippov, N. Shavina, and O. A. Protsenko.** 1989. Genetic
755 analysis and simulation of the virulence of *Yersinia pestis*. Mol Gen Mikrobiol
756 Virusol**:**42-47.
757 33. **Lawrenz, M. B., J. D. Lenz, and V. L. Miller.** 2009. A novel autotransporter adhesin
758 is required for efficient colonization during bubonic plague. Infect Immun **77:**317-
759 326.
760 34. **Leif, W. R., and A. P. Krueger.** 1950. Studies on the experimental epidemiology of
761 respiratory infections. I. An apparatus for the quantitative study of air-borne
762 respiratory pathogens. J Infect Dis **87:**103-116.
763 35. **Mollaret, H. H.** 1968. Conservation du bacille de la peste durant 28 mois en terrier
764 artificiel : démonstration expérimentale de la conservation interépizootique de la peste
765 dans ses foyers invétérés. Comptes rendus de l'Académie des Sciences de Paris
766 **267:**972-973.
767 36. **Parkhill, J., B. W. Wren, N. R. Thomson, R. W. Titball, M. T. Holden, M. B.**
768 **Prentice, M. Sebaihia, K. D. James, C. Churcher, K. L. Mungall, S. Baker, D.**
769 **Basham, S. D. Bentley, K. Brooks, A. M. Cerdeno-Tarraga, T. Chillingworth, A.**
770 **Cronin, R. M. Davies, P. Davis, G. Dougan, T. Feltwell, N. Hamlin, S. Holroyd,**
771 **K. Jagels, A. V. Karlyshev, S. Leather, S. Moule, P. C. Oyston, M. Quail, K.**
772 **Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead, and B. G.**

773 **Barrell.** 2001. Genome sequence of *Yersinia pestis*, the causative agent of plague.
774 Nature **413:**523-527.
775 37. **Perry, R. D., and J. D. Fetherston.** 1997. *Yersinia pestis*--etiologic agent of plague.
776 Clin Microbiol Rev **10:**35-66.
777 38. **Pouillot, F., A. Derbise, M. Kukkonen, J. Foulon, T. K. Korhonen, and E.**
778 **Carniel.** 2005. Evaluation of O-antigen inactivation on Pla activity and virulence of
779 *Yersinia pseudotuberculosis* harbouring the pPla plasmid. Microbiology **151:**3759-
780 3768.
781 39. **Pouillot, F., C. Fayolle, and E. Carniel.** 2007. A putative DNA adenine
782 methyltransferase is involved in *Yersinia pseudotuberculosis* pathogenicity.
783 Microbiology **153:**2426-2434.
784 40. **Pouillot, F., C. Fayolle, and E. Carniel.** 2008. Characterization of chromosomal
785 regions conserved in *Yersinia pseudotuberculosis* and lost by *Yersinia pestis*. Infect
786 Immun **76:**4592-4599.
787 41. **Pourcel, C., G. Salvignol, and G. Vergnaud.** 2005. CRISPR elements in *Yersinia*
788 *pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide
789 additional tools for evolutionary studies. Microbiology **151:**653-663.
790 42. **Radnedge, L., P. G. Agron, P. L. Worsham, and G. L. Andersen.** 2002. Genome
791 plasticity in *Yersinia pestis*. Microbiology **148:**1687-1698.
792 43. **Rakin, A., S. Schubert, I. Guilvout, E. Carniel, and J. Heesemann.** 2000. Local
793 hopping of IS*3* elements into the A+T-rich part of the high-pathogenicity island in
794 *Yersinia enterocolitica* 1B, O:8. FEMS Microbiol Lett **182:**225-229.
795 44. **Rojas, C. M., J. H. Ham, W. L. Deng, J. J. Doyle, and A. Collmer.** 2002. HecA, a
796 member of a class of adhesins produced by diverse pathogenic bacteria, contributes to
797 the attachment, aggregation, epidermal cell killing, and virulence phenotypes of

798 *Erwinia chrysanthemi* EC16 on *Nicotiana clevelandii* seedlings. Proc Natl Acad Sci U
799 S A **99:**13142-13147.
800 45. **Samoilova, S. V., L. V. Samoilova, I. N. Yezhov, I. G. Drozdov, and A. P.**
801 **Anisimov.** 1996. Virulence of pPst+ and pPst- strains of *Yersinia pestis* for guinea-
802 pigs. J Med Microbiol **45:**440-444.
803 46. **Sebbane, F., C. Jarrett, D. Gardner, D. Long, and B. J. Hinnebusch.** 2009. The
804 *Yersinia pestis* caf1M1A1 fimbrial capsule operon promotes transmission by flea bite
805 in a mouse model of bubonic plague. Infect Immun **77:**1222-1229.
806 47. **Sebbane, F., C. O. Jarrett, D. Gardner, D. Long, and B. J. Hinnebusch.** 2006.
807 Role of the *Yersinia pestis* plasminogen activator in the incidence of distinct
808 septicemic and bubonic forms of flea-borne plague. Proc Natl Acad Sci U S A
809 **103:**5526-5530.
810 48. **Smego, R. A., J. Frean, and H. J. Koornhof.** 1999. Yersiniosis I: microbiological
811 and clinicoepidemiological aspects of plague and non-plague Yersinia infections. Eur
812 J Clin Microbiol Infect Dis **18:**1-15.
813 49. **Sodeinde, O. A., Y. V. Subrahmanyam, K. Stark, T. Quan, Y. Bao, and J. D.**
814 **Goguen.** 1992. A surface protease and the invasive character of plague. Science
815 **258:**1004-1007.
816 50. **Steiner, W. W., and P. L. Kuempel.** 1998. Sister chromatid exchange frequencies in
817 *Escherichia coli* analyzed by recombination at the dif resolvase site. J Bacteriol
818 **180:**6269-6275.
819 51. **Sun, Y. C., B. J. Hinnebusch, and C. Darby.** 2008. Experimental evidence for
820 negative selection in the evolution of a *Yersinia pestis* pseudogene. Proc Natl Acad
821 Sci U S A **105:**8097-8101.

822 52. **Surgalla, M. J., and E. D. Beesley.** 1969. Congo red-agar plating medium for
823 detecting pigmentation in *Pasteurella pestis*. Appl Microbiol **18:**834-837.
824 53. **Thomson, N. R., S. Howard, B. W. Wren, M. T. Holden, L. Crossman, G. L.**
825 **Challis, C. Churcher, K. Mungall, K. Brooks, T. Chillingworth, T. Feltwell, Z.**
826 **Abdellah, H. Hauser, K. Jagels, M. Maddison, S. Moule, M. Sanders, S.**
827 **Whitehead, M. A. Quail, G. Dougan, J. Parkhill, and M. B. Prentice.** 2006. The
828 complete genome sequence and comparative genome analysis of the high
829 pathogenicity *Yersinia enterocolitica* strain 8081. PLoS Genet **2:**e206.
830 54. **Vallenet, D., S. Engelen, D. Mornico, S. Cruveiller, L. Fleury, A. Lajus, Z. Rouy,**
831 **D. Roche, G. Salvignol, C. Scarpelli, and C. Medigue.** 2009. MicroScope: a
832 platform for microbial genome annotation and comparative genomics. Database
833 (Oxford) **2009:**bap021.
834 55. **Welkos, S. L., A. M. Friedlander, and K. J. Davis.** 1997. Studies on the role of
835 plasminogen activator in systemic infection by virulent *Yersinia pestis* strain C092.
836 Microb Pathog **23:**211-223.
837 56. **Zhou, D., Y. Han, E. Dai, D. Pei, Y. Song, J. Zhai, Z. Du, J. Wang, Z. Guo, and R.**
838 **Yang.** 2004. Identification of signature genes for rapid and specific characterization of
839 *Yersinia pestis*. Microbiol Immunol **48:**263-269.
840 57. **Zhou, D., Y. Han, Y. Song, Z. Tong, J. Wang, Z. Guo, D. Pei, X. Pang, J. Zhai, M.**
841 **Li, B. Cui, Z. Qi, L. Jin, R. Dai, Z. Du, J. Bao, X. Zhang, J. Yu, P. Huang, and R.**
842 **Yang.** 2004. DNA microarray analysis of genome dynamics in *Yersinia pestis*:
843 insights into bacterial genome microevolution and niche adaptation. J Bacteriol
844 **186:**5138-5146.
845
846

847 **Legends**

848 **Figure 1. Genetic organization of the eight chromosomal loci acquired by *Y. pestis* and**
849 **comparison with the homologous regions in *Y. pseudotuberculosis*.**

850 White arrows indicate CDSs present in *Y. pestis* CO92 but not in *Y. pseudotuberculosis*
851 IP32953. Dark grey arrrows are CDSs harboring >98% nt identity betwen *Y. pestis* and *Y.*
852 *pseudotuberculosis*. Light grey areas represent the conserved regions on each side of the
853 acquired loci. Dashed lines with arrows below each region indicate the DNA fragments that
854 have been replaced by a kanamycin cassette, and the numbers at each extremity correspond to
855 the primer pairs used to generate the fragment to be deleted. In panel B, grey rectangles
856 corresponds to repeats identical in *Y. pseudotuberculosis* and *Y. pestis*. In panel F, the 250 nt-
857 long DNA repeats found at the borders of $R4_{pe}$ are indicated by black rectangles. In panel E
858 and H, black circles represent direct repeated sequences found at the extremities of each
859 region, and their size in nt are indicated below. In panel H, dashed arrows indicate ORFs that
860 are conserved in *Y. pestis* and *Y. pseudotuberculosis* although they might have been acquired
861 independently (see Fig. S1 for a more detailed comparison of this region).

862

863 **Figure 2. Position of the regions acquired and lost by *Y. pestis* on a schematic map of the**
864 **CO92 chromosome.**

865 The position on the *Y. pestis* CO92 chromosome of all loci acquired or lost by *Y. pestis* are
866 shown by black and grey arrows, respectively. The putative mechanisms of acquisition are
867 indicated next to each *Y. pestis*-specific regions.

868

869 **Figure 3**. ***Y. pestis* microevolution tree with the steps at which the specific regions are**
870 **predicted to have been acquired or to have undergone various mutations**.

871 The 21 *Y. pestis* strains for which a complete genome sequence was available were positioned
872 on the microevolutionary tree based on the presence or absence of a set of 16 SNPs (inside
873 white ovals) shown to characterize the various *Y. pestis* phylogenetic branches (1). To
874 position strain Pestoides A, additional SNPs defining branch 0.PE4 were examined. Among
875 them, some were present (sA: s95, s96, s97, s99, s102, s104, s105, s114, s116 and s120),
876 while others were absent (sB: s94, s98, s100, s101, s106, s107, s108, s109, s110, s111, s112,
877 s113, s115, s117, s118 and s119). The hatched rectangle corresponds to the position in the
878 tree where the *Y. pestis*-specific regions are predicted to have been acquired. The various
879 mutations in specific regions that characterize phylognentic branches are indicated by an
880 arrow and a grey rectangle which contains the type of mutation (SNP or partial deletion, ΔP)
881 and the locus involved. A more detailed description of these mutations is presented in Table
882 S6.
883
884 **Figure 4. Infection and blockage rates in fleas that fed on blood containing wild type or**
885 **mutant *Y. pestis* strains.** (A) Percentage of fleas that developed proventricular blockage
886 during a four-week period following the infectious blood meal. (B) Percentage of fleas
887 infected four weeks after the infectious blood meal. Vertical lines indicate the range of four
888 independent experiments with the wild-type strain.
889

Table 1: Virulence for mice of the *Y. pestis* recombinant strains after subcutaneous (sc) injection or aerosols exposure.

Strains	sc		Aerosols		
	CFU injected	#dead/total animal infected	suspension nebulized CFU/ml	#dead/total animal infected	mean time to death
CO92p	12	3/5	$0.95 \pm 0.3 \times 10^{8}$ [a]	12/12[a]	4 ± 1 days[a]
CO92⊗R1$_{pe}$	10	3/5	1.1×10^{8}	6/6	4.5 days
CO92⊗R2$_{pe}$	16	3/5	2×10^{8}	6/6	3 days
CO92⊗R3$_{pe}$	12	3/5	1.5×10^{8}	6/6	3 days
CO92 ⊗R4$_{pe}$	12	3/5	1.4×10^{8}	5/5	4 days
CO92 ⊗R5$_{pe}$	9	3/5	1.4×10^{8}	6/6	3 days
CO92 ⊗R6$_{pe}$	17	5/5	1.1×10^{8}	6/6	3 days
CO92 ⊗CDS1$_{pe}$	14	5/5	0.9×10^{8}	6/6	4 days
CO92 ⊗CDS2$_{pe}$	13	3/5	0.6×10^{8}	6/6	3 days

[a] Values are means ± standard deviations of two experiments with six mice each.

Fig. 1

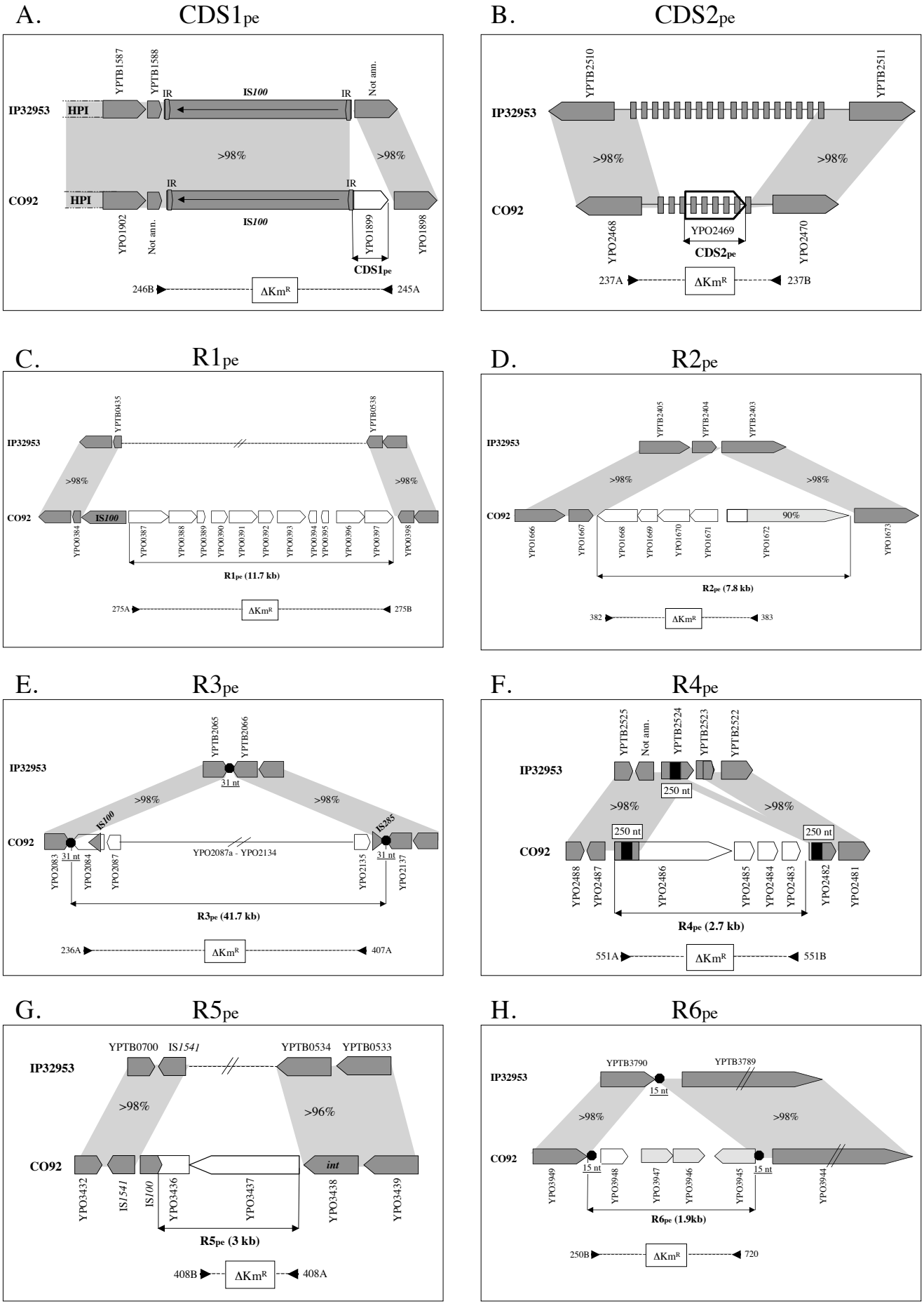

A. CDS1pe
IP32953
HPI
YPTB1587
YPTB1588
IR
IS100
IR
Not ann.
>98%
>98%
CO92
HPI
IR
IR
IS100
YPO1902
Not ann.
YPO1899
YPO1898
CDS1pe
246B
ΔKmR
245A
B. CDS2pe
IP32953
YPTB2510
YPTB2511
>98%
>98%
CO92
YPO2468
YPO2469
CDS2pe
YPO2470
237A
ΔKmR
237B
C. R1pe
IP32953
YPTB0435
YPTB0538
>98%
>98%
CO92
IS100
YPO0384
YPO0387
YPO0388
YPO0389
YPO0390
YPO0391
YPO0392
YPO0393
YPO0394
YPO0395
YPO0396
YPO0397
YPO0398
R1pe (11.7 kb)
275A
ΔKmR
275B
D. R2pe
IP32953
YPTB2405
YPTB2404
YPTB2403
>98%
>98%
CO92
90%
YPO1666
YPO1667
YPO1668
YPO1669
YPO1670
YPO1671
YPO1672
YPO1673
R2pe (7.8 kb)
382
ΔKmR
383
E. R3pe
IP32953
YPTB2065
YPTB2066
31 nt
>98%
>98%
CO92
IS100
IS285
31 nt
31 nt
YPO2083
YPO2084
YPO2087
YPO2087a - YPO2134
YPO2135
YPO2137
R3pe (41.7 kb)
236A
ΔKmR
407A
F. R4pe
IP32953
YPTB2525
Not ann.
YPTB2524
YPTB2523
YPTB2522
250 nt
>98%
>98%
CO92
250 nt
250 nt
YPO2488
YPO2487
YPO2486
YPO2485
YPO2484
YPO2483
YPO2482
YPO2481
R4pe (2.7 kb)
551A
ΔKmR
551B
G. R5pe
IP32953
YPTB0700
IS1541
YPTB0534
YPTB0533
>98%
>96%
CO92
int
YPO3432
IS1541
IS100
YPO3436
YPO3437
YPO3438
YPO3439
R5pe (3 kb)
408B
ΔKmR
408A
H. R6pe
IP32953
YPTB3790
YPTB3789
15 nt
>98%
>98%
CO92
15 nt
15 nt
YPO3949
YPO3948
YPO3947
YPO3946
YPO3945
YPO3944
R6pe (1.9kb)
250B
ΔKmR
720

Fig. 2

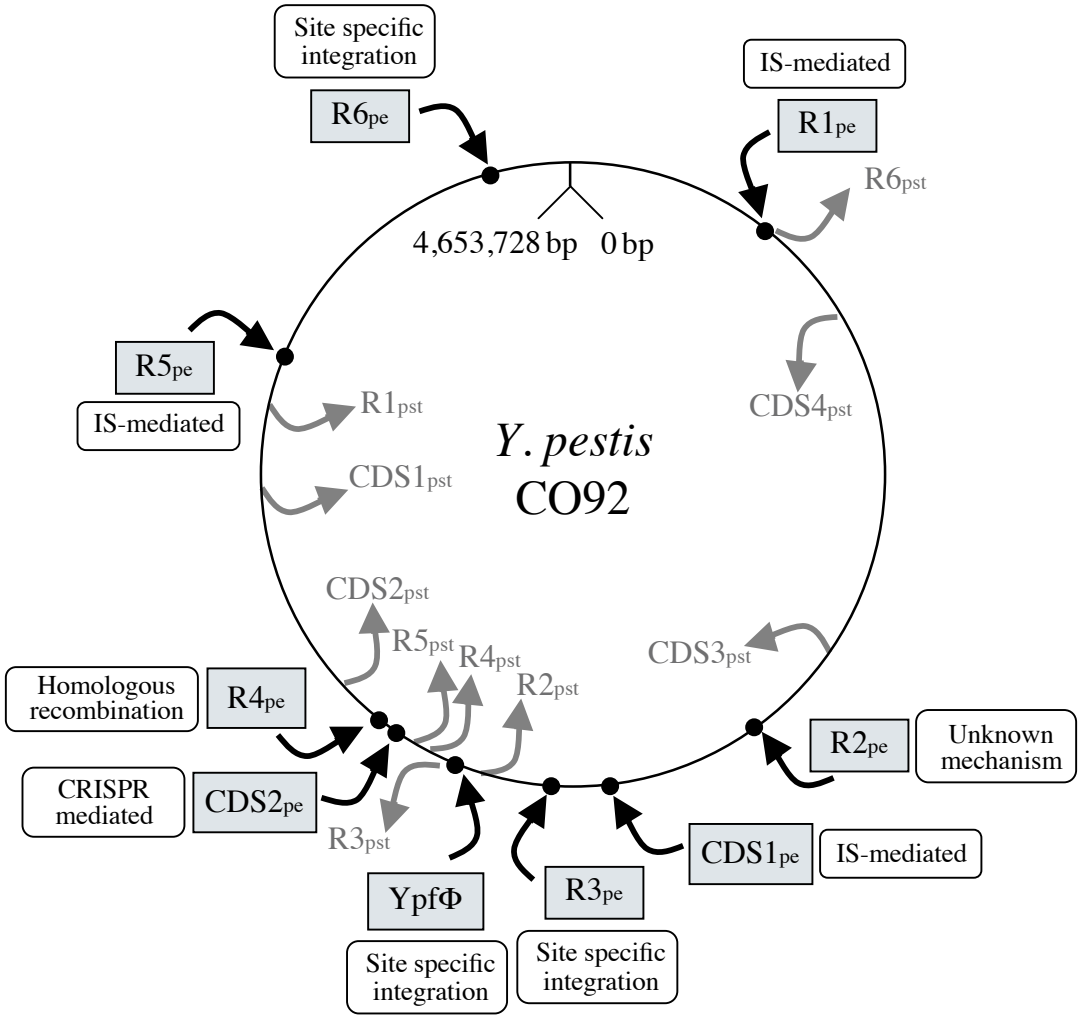

Site specific integration
R6pe
IS-mediated
R1pe
R6pst
4,653,728 bp
0 bp
R5pe
IS-mediated
R1pst
CDS4pst
Y. pestis
CO92
CDS1pst
CDS2pst
R5pst
R4pst
R2pst
CDS3pst
Homologous recombination
R4pe
R2pe
Unknown mechanism
CRISPR mediated
CDS2pe
R3pst
CDS1pe
IS-mediated
YpfΦ
R3pe
Site specific integration
Site specific integration

Fig. 3

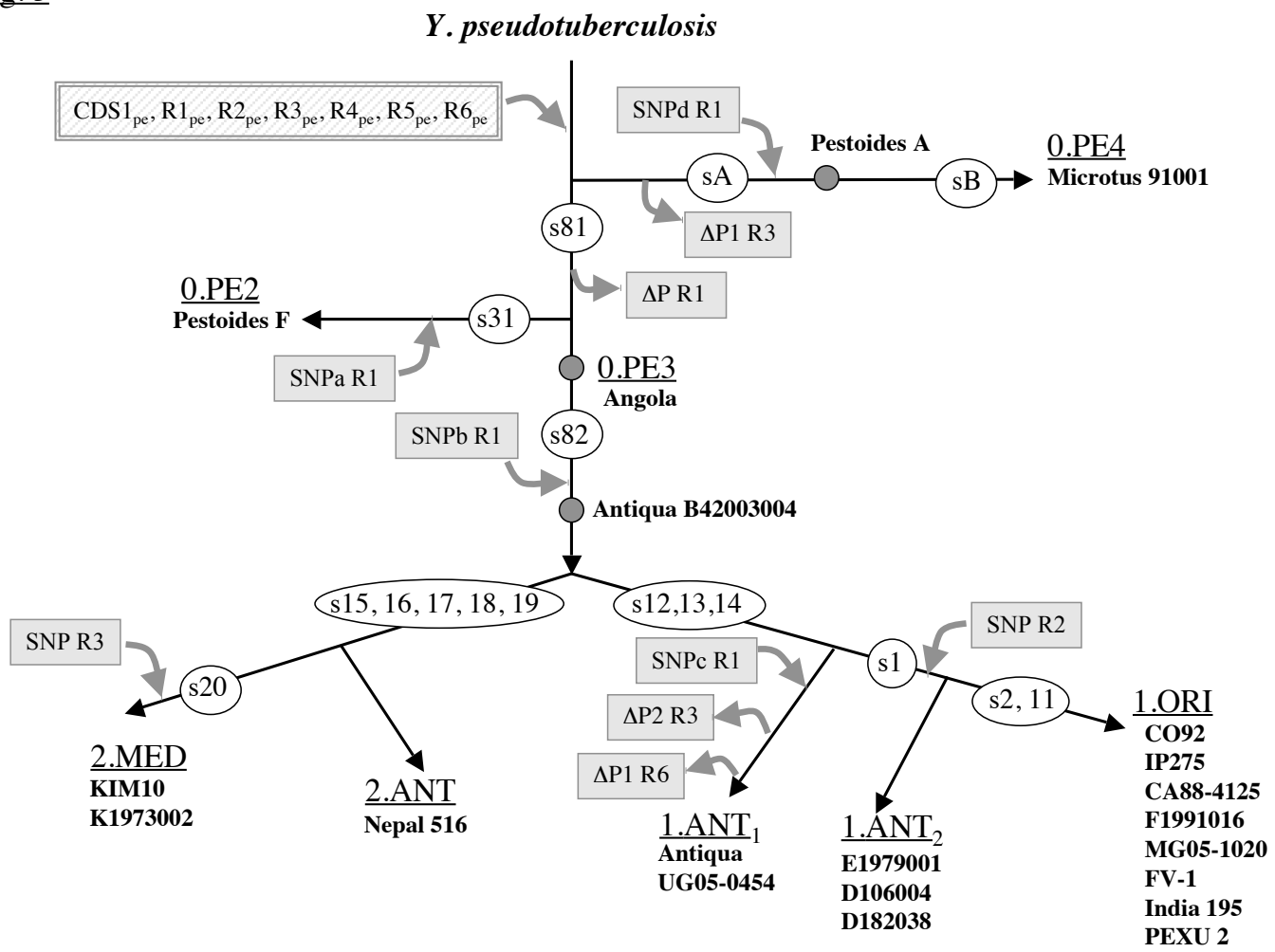
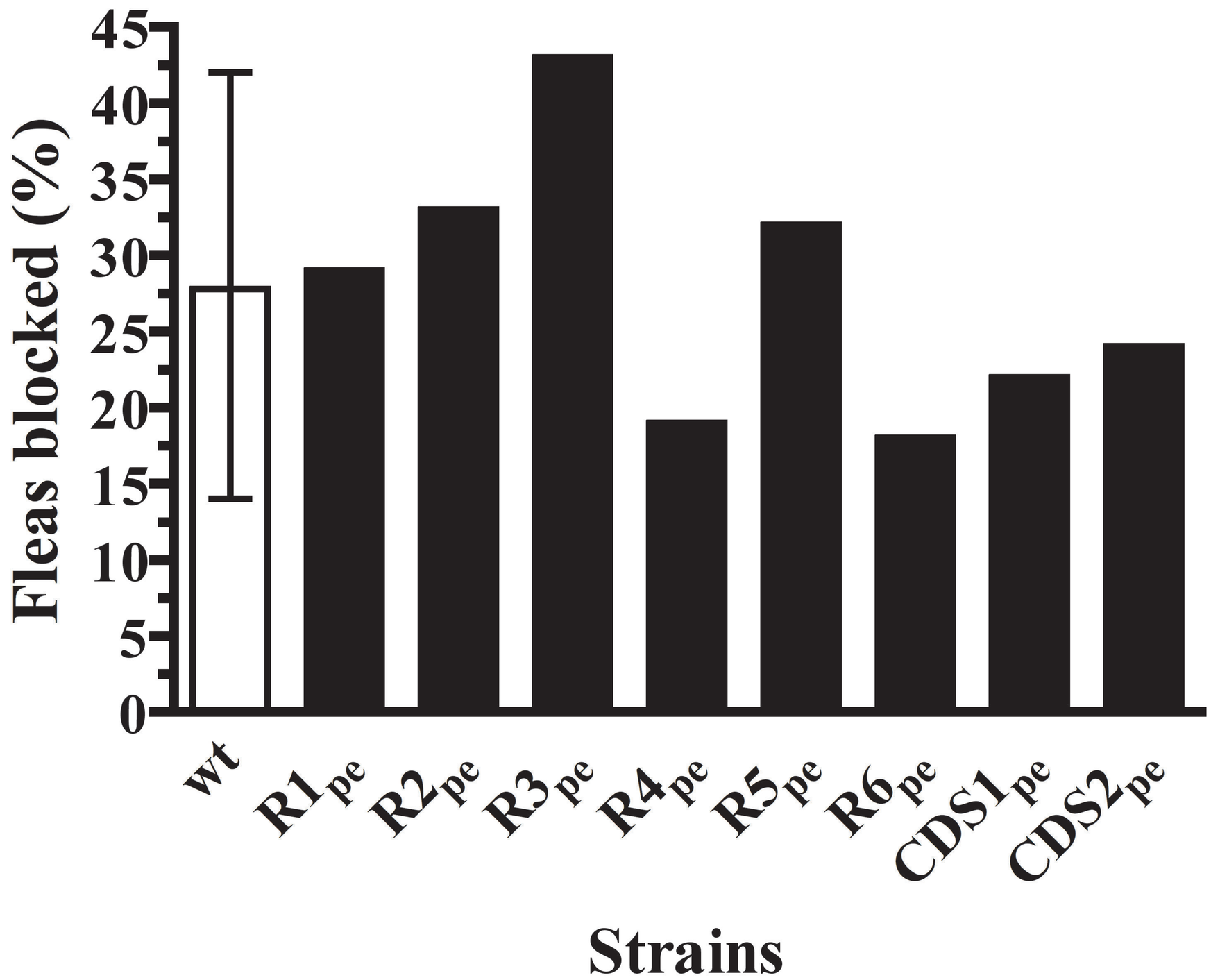
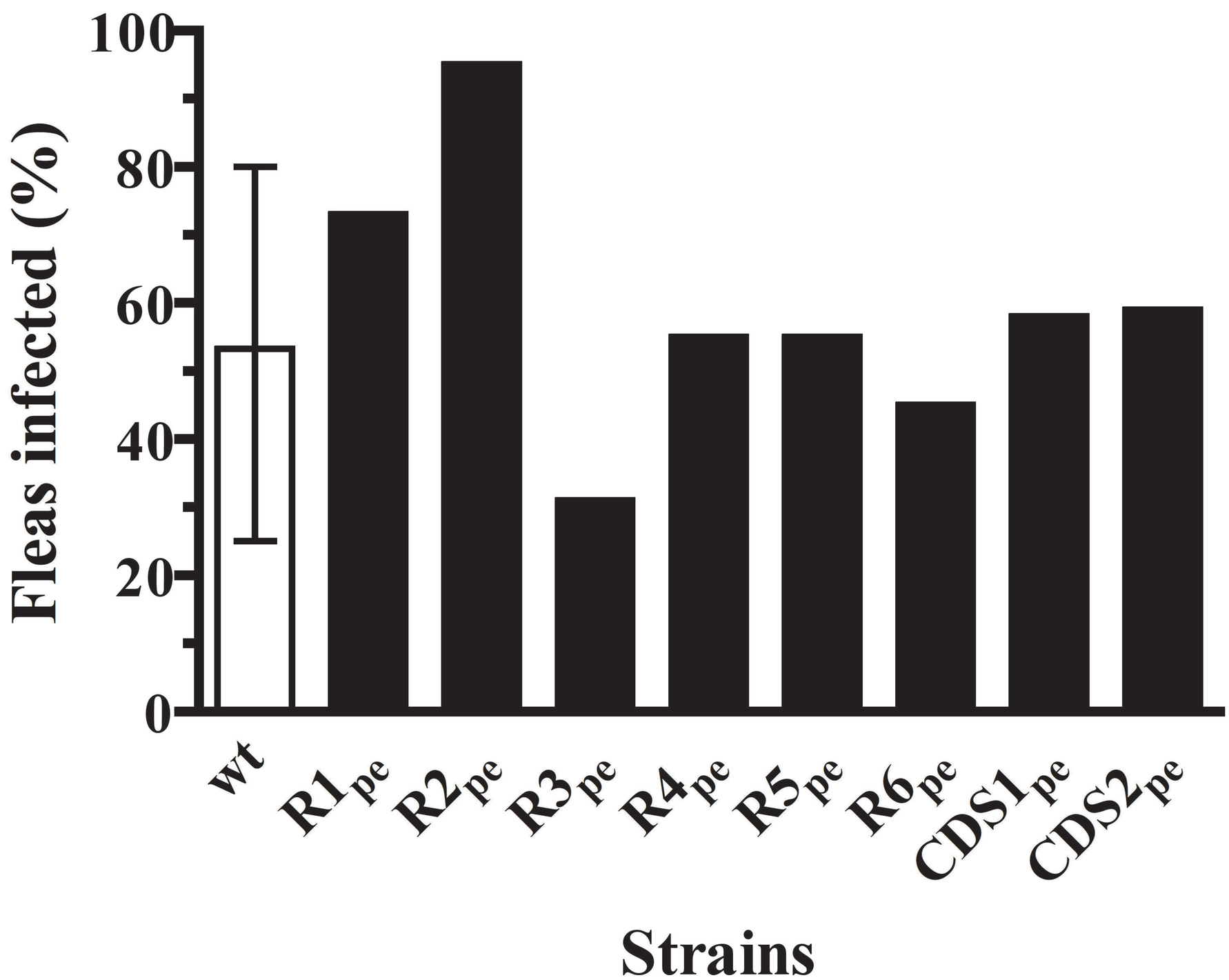

Y. pseudotuberculosis
CDS1pe, R1pe, R2pe, R3pe, R4pe, R5pe, R6pe
SNPd R1
Pestoides A
0.PE4
Microtus 91001
sA
sB
ΔP1 R3
s81
ΔP R1
0.PE2
Pestoides F
s31
SNPa R1
0.PE3
Angola
SNPb R1
s82
Antiqua B42003004
s15, 16, 17, 18, 19
s12,13,14
SNP R3
SNP R2
SNPc R1
s20
s1
s2, 11
ΔP2 R3
ΔP1 R6
1.ORI
CO92
IP275
CA88-4125
F1991016
MG05-1020
FV-1
India 195
PEXU 2
2.MED
KIM10
K1973002
2.ANT
Nepal 516
1.ANT1
Antiqua
UG05-0454
1.ANT2
E1979001
D106004
D182038

Fig. 4

A


Fleas blocked (%)
45
40
35
30
25
20
15
10
5
0
wt
R1pe
R2pe
R3pe
R4pe
R5pe
R6pe
CDS1pe
CDS2pe
Strains

B


Fleas infected (%)
100
80
60
40
20
0
wt
R1pe
R2pe
R3pe
R4pe
R5pe
R6pe
CDS1pe
CDS2pe
Strains