JOURNAL OF BACTERIOLOGY, Aug. 2002, p. 4489–4499 Vol. 184, No. 16
0021-9193/02/$04.00+0 DOI: 10.1128/JB.184.16.4489–4499.2002

The Superantigen Gene *ypm* Is Located in an Unstable Chromosomal Locus of *Yersinia pseudotuberculosis*

Christophe Carnoy,[1,2]* Stephanie Floquet,[1] Michael Marceau,[1] Florent Sebbane,[1] Stephanie Haentjens-Herwegh,[3] Annie Devalckenaere,[3] and Michel Simonet[1,3]

Equipe Mixte Inserm E9919—Université JE 2225—Institut Pasteur de Lille, Institut de Biologie de Lille,[1] *Faculté des Sciences Pharmaceutiques et Biologiques,*[2] *and Laboratoire de Bactériologie-Hygiène, Centre Hospitalier Régional et Universitaire de Lille,*[3] *Lille, France*

Received 15 February 2002/Accepted 24 May 2002

Yersinia pseudotuberculosis **produces YPM (*Y. pseudotuberculosis*-derived mitogen), a superantigenic toxin that exacerbates the virulence of the bacterium in vivo. To date, three alleles of the superantigen gene (*ypmA*, *ypmB*, and *ypmC*) have been described. These genes are not found in all *Y. pseudotuberculosis* strains and have a low GC content, suggesting their location on mobile genetic elements. To elucidate this question, the genetic environment of the superantigen-encoding genes was characterized and 11 open reading frames (ORFs) were defined. Sequence analysis revealed that the *ypm* genes were not associated with plasmids, phages, transposons, or pathogenicity islands and that the superantigen genes were always located in the chromosome between ORF3 and ORF4. Nonsuperantigenic strains exhibited the same genetic organization of the locus but lacked the *ypm* gene between ORF3 and ORF4. A new insertion sequence, designated IS*1398*, which displays features of the Tn*3* family, was characterized downstream of the *ypmA* and *ypmC* genes. A 13.3-kb region containing the *ypm* genes was not found in the genome of *Y. pestis* (CO92 and KIM 5 strains). We experimentally induced deletion of the *ypm* gene from a superantigen-expressing *Y. pseudotuberculosis*: using the association of *aph(3′)-IIIa* and *sacB* genes, we demonstrated that when these reporter genes were present in the *ypm* locus, deletion of these genes was about 250 times more frequent than when they were located in another region of the *Y. pseudotuberculosis* chromosome. These results indicate that unlike other superantigenic toxin genes, the *Yersinia ypm* genes are not associated with mobile genetic elements but are inserted in an unstable locus of the genome.**

Yersinia pseudotuberculosis, a microorganism causing gastrointestinal diseases and immunopathological complications such as reactive arthritis and erythema nodosum (10, 47, 59, 60), is at present the only gram-negative bacteria known to produce a superantigenic toxin. This molecule (designated YPM for *Y. pseudotuberculosis*-derived mitogen) is a 14.5-kDa protein able to induce proliferation of human T lymphocytes bearing Vβ3, Vβ9, Vβ13.1, and Vβ13.2 T-cell receptor variable regions (1, 61). In vivo, YPM induces lethal shock in mice (44) and exacerbates the virulence of *Y. pseudotuberculosis* in a murine experimental model of systemic infection (13). To date, three YPM variants (YPMa, YPMb, and YPMc) have been described. YPMa displays 83% identity with YPMb (11, 50) and differs from YPMc only by a single substitution at position 51 of the mature protein (12). A phylogenetic analysis based on the amino acid sequences of various bacterial superantigenic toxins indicated that the YPM variants belong to a new type of bacterial superantigen family (12, 46). The *ypm* genes encoding the YPM toxins have not been found in the genome of *Yersinia pestis*, a genetically related species (49). *Yersinia enterocolitica*, another pathogenic *Yersinia* species, has been described as a mitogen-producing microorganism (57, 58); however, the T-cell specificity of the bacterial mitogen and the absence of DNA hybridization with a *ypm*-specific probe clearly demonstrated that the mitogen activity of *Y. enterocolitica* cannot be attributed to a YPM variant (12, 63).

* Corresponding author. Mailing address: Département de Pathogenèse des Maladies Infectieuses, Institut de Biologie de Lille, 1 rue du Professeur Calmette, F-59021 Lille cedex, France. Phone: 33 (0)3 20 87 11 81. Fax: 33 (0)3 20 87 11 83. E-mail: christophe.carnoy@ibl.fr.

At present, little information on the genetic location of *ypm* genes is available. It is now established that *ypmA* is not located on the virulence plasmid pYV, since the gene is present in pYV-cured strains of *Y. pseudotuberculosis* (11, 43). Hence, as they are not present on pYV, *ypm* genes are likely to be chromosomal, although no experimental evidence confirming this hypothesis has been found (43). Nevertheless, some arguments suggest that *ypm* genes might be harbored by a mobile genetic element. First, *ypm* genes are not present in all *Y. pseudotuberculosis* strains (22, 63), raising the question of the transmission of the superantigen genes among the *Y. pseudotuberculosis* population. Secondly, analysis of the nucleotide sequences of *ypm* genes revealed a guanine and cytosine (GC) content of between 34.6 and 35.3%, whereas the genome of *Y. pseudotuberculosis* has a higher GC content (46.5%) (5), thus suggesting that *Y. pseudotuberculosis* obtained *ypm* genes from a microorganism with a low GC percentage.

In this study, we addressed the question of the genetic mobility of the superantigen-encoding genes. With this goal in mind, we characterized the genetic environment of the *Y. pseudotuberculosis ypm* genes and compared it with the genetic organization displayed by nonsuperantigenic strains. Sequence analysis ruled out the association of the superantigen genes with mobile genetic elements but indicated that the superantigen-producing strains represent a clonal population of *Y. pseudotuberculosis* that has evolved concomitantly with nonsuperantigenic *Y. pseudotuberculosis* clones. We also demonstrated that the locus containing the *ypm* gene was unstable and that DNA deletion in this region can occur with high frequency without any characteristic or sequence-specific excisions.

TABLE 1. *Y. pseudotuberculosis* strains used in this study

Y. pseudotuberculosis genotype	Strain (serotype)[a]	Source
ypmA$^+$	**AH (4b)**, KM (4b), MA (2b), NT (4b), ST (4b)	N. Takeda
	1402 (4), 1404 (6)	H. Müller-Alouf
	1553 (6), 1554 (6), 1833 (4)	E. Carniel
	1191/90 (4a)	S. Aleksić
ypmB$^+$	**487/90 (6)**, 1093 (11), 1096 (6), 1119 (5b), 1155 (5a)	S. Aleksić
ypmC$^+$	**YPT1 (3)**, YPT5 (3)	R. Van Noyen
	2887 (3), 32945 (3), 32975 (3), 32977 (3), 32992 (3)	E. Carniel
	200/90 (3), 298/89 (3), 304/89 (3), 1134/90 (3)	S. Aleksić
	1216/93 (3), 776/88 (1)	J. Sundar
	WE31/93 (3)	G. Wauters
Nonsuperantigenic	**9314/74 (1)**	J. Sundar
	1830 (4), 2821 (5), 2843 (5), 2926 (2)	E. Carniel
	199/90 (5a), 300/89 (2b), 367/89 (1a), 1432/94 (2c)	S. Aleksić
	1401 (3), 1403 (5)	H. Müller-Alouf
	YPT12 (2a)	R. Van Noyen

[a] Strains in bold were chosen as representative of each subgroup.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and plasmids. Forty-two *Y. pseudotuberculosis* strains were used in this study (Table 1). They were kindly provided by S. Aleksić (Hygiene Institut Hamburg, Hamburg, Germany), E. Carniel (Centre National de Référence des *Yersinia*, Institut Pasteur, Paris, France), H. Müller-Alouf (Institut Pasteur de Lille, Lille, France), J. Sundar (Statens Institutt for Folkehelse, Oslo, Norway), N. Takeda (Department of Pediatrics, Kurashiki Central Hospital, Okayama, Japan), R. Van Noyen (Imeldaziekenhuis, Bonheiden, Belgium), and G. Wauters (Université Catholique de Louvain, Louvain, Belgium). A spontaneous nalidixic acid-resistant mutant of *Y. pseudotuberculosis* AH (designated AH Nal^r) was obtained by plating a bacterial culture onto Luria-Bertani (LB) agar containing a nalidixic acid concentration gradient. The superantigen-deficient strains were not epidemiologically related, as indicated by a *Not*I genomic restriction profile analysis using pulsed-field gel electrophoresis (data not shown) (30, 48).

The plasmids used in this study are listed in Table 2. *Escherichia coli* DH5α (28) was the host for pUC recombinant plasmids, whereas *E. coli* SY327λ*pir* (41) and *E. coli* SM10λ*pir* (55) were hosts for the pMM7043 suicide plasmid. Bacteria were grown at 28°C (*Yersinia*) or 37°C (*E. coli*) in LB broth or agar. Kanamycin (50 μg/ml), chloramphenicol (50 μg/ml), ampicillin (100 μg/ml), and nalidixic acid (10 μg/ml) were added to media for bacterial selection when necessary.

DNA techniques. Genomic DNA was extracted from bacterial cells as previously described (39). Plasmid DNA and endonuclease-restricted fragments were purified on Qiagen columns according to the manufacturer's recommendation (Qiagen S. A., Courtaboeuf, France). Southern hybridizations were carried out as previously described (13). Standard PCR amplification was performed as described elsewhere (54), whereas conditions for recombinant PCR were as presented below. DNA fragments were ligated to endonuclease-restricted vectors using T4 DNA ligase, according to standard techniques (Life Technologies, Cergy Pontoise, France) (53). Recombinant plasmids were introduced into *E. coli* by transformation (53).

Oligonucleotide primers and probes. Intergenic regions were amplified with the following primers located within open reading frames (ORFs) (Fig. 1): no. 114 (5′GTGTTCCGTTTGATGAGGAGG3′), no. 93 (5′CCTGCCAATAAGCTAAGGCAG3′), no. 63 (5′GCTGTTCAGTGTTATGCCGCTG3′), no. 29 (5′GACCGCCAGCATCTACCTG3′), no. 16 (5′GCGGCAAGCTTTGAAGGGTTGTCACAATTGCACCT3′), no. 1 (5′ACACTTTTCTCTGGAGTAGCG3′), no. 2 (5′ACAGGACATTTCGTCA3′), no. 4 (5′TGTAGGAGGCAATGGATGGGG3′), no. 30 (5′GCTGCACCGTCTCTGTTATCAC3′), no. 13 (5′CCGATGCGATTAATACTGCC3′), no. 9 (5′CATGCTGGCACCTGCCTCGAA3′), no. 22 (5′GCGGATACATGCATCCGCAG3′), no. 44 (5′GGCAACTATGACACCTGCGC3′), no. 46 (5′GCCCGCAGGCACCCGGCAAAA3′), no. 56 (5′GAGTCGTTGATTGCACGCCGTAG3′), no. 101 (5′GGTCTGTGCTACGCATCGGGATC3′), no. 65 (5′GCAGTTGTGCACCCAGTTCGGC3′), no. 85 (5′CCGCGTCAATATTAACTCATTGG3′), no. 131 (5′CGACGCAGCATCGCACGGTAG3′), no. 139 (5′GTCGGTGCCGCACTCGGCATG3′), and no. 136 (5′GAAGGCTGCGGTGGGTGGAGGG3′). In nonsuperantigenic strains, the ORF3-ORF4 intergenic region was amplified with primers no. 16 and no. 30 and then sequenced with primers no. 30 and no. 58 (5′GCACCAAGGTGACGATAGGCG3′).

The probe specific for the three *ypm* alleles was generated by PCR with primers no. 1 and no. 2 (13). The IS*1398*-specific probe was PCR amplified with primers no. 14 (5′GTACCACCTGACCAAAGCTAG3′) and no. 20 (5′CCGGAAGCAACATCCAAGAG3′). Probes were generated by amplification in the presence of digoxigenin-11-dUTP (Roche Diagnostics, Meylan, France).

All primers used in this study were synthesized by Sigma-Genosys Ltd (Pampisford, United Kingdom) or Genset SA (Paris, France).

DNA sequencing and sequence analysis. The recombinant plasmids used to sequence the *ypm* loci are described in Table 2. Plasmid pCCY10 was sequenced by a walking strategy. Since high degrees of homology between the loci of the three superantigenic alleles exist, primers generated from the pCCY10 sequence were used for the sequencing of pSF11 (locus *ypmB*), pSC10 (locus *ypmC*), and pSC20 (locus deficient in *ypmA*). PCR amplification products were directly sequenced as necessary. Sequencing was performed on both DNA strands. ORFs over 300 bp long were identified with ORF Finder software on the National Center for Biotechnology Information (NCBI) server (http://www.ncbi.nlm.nih.gov). Sequences were searched against nucleotide and protein databases by using BLAST on the NCBI server. Sequences were also compared with the *Y. pestis* CO92 sequences presented by the *Y. pestis* Sequencing Group at the Sanger Center (http://www.sanger.ac.uk/Projects/Y_pestis and accession number AJ414151) (49) and with the *Y. pestis* KIM5 (strain P12) genomic sequence available on the University of Wisconsin Genome Project website (http://www.genome.wisc.edu).

TABLE 2. Plasmids used in this study

Plasmid	Description[a]	Reference or source
pUC18 and 19	Ap^r cloning vector	62
pACYC184	Source of the chloramphenicol resistance gene (*cat*)	14
pCVD442	Source of the *sacB* gene	19
pUC1318-KmII	Source of the kanamycin resistance gene *aph(3′)-IIIa*	Gift from P. Trieu-Cuot
pCCY10	pUC18Ω, 6.5-kb *Hin*dIII fragment of *Y. pseudotuberculosis* AH containing the *ypmA* gene	13
pFM10	pUC18Ω, 5-kb *Kpn*I fragment of *Y. pseudotuberculosis* AH containing ORF3 gene	This study
pSF11	pUC19Ω, 8-kb *Hin*dIII fragment of *Y. pseudotuberculosis* 487/90 containing the *ypmB* gene	This study
pFM30	pUC18Ω, 2.5-kb *Kpn*I fragment of *Y. pseudotuberculosis* 487/90 containing the *ypmB* gene	This study
pSC10	pUC18Ω, 12.7-kb *Pst*I fragment of *Y. pseudotuberculosis* YPT1 containing the *ypmC* gene	This study
pSC20	pUC18Ω, 9.7-kb *Pst*I fragment of *Y. pseudotuberculosis* 9314/74 hybridizing with an ORF5-specific probe	This study
pMM70413	Cm^r suicide vector; pGP704 (42) derivative in which the 675-bp *Pst*I fragment containing the *bla* gene was replaced by two copies in tandem of the *Pst*I PCR-generated *cat* gene	This study
pCCY41.2	pMM70413Ω, *sacB* and *aph(3′)-IIIa* genes inserted within *ypmA*	This study
pIL31.3	pMM70413Ω, *sacB* and *aph(3′)-IIIa* genes inserted upstream of the urease operon promoter	This study

[a] Ap^r and Cm^r, resistance to ampicillin and chloramphenicol, respectively; Ω, in vitro insertion.

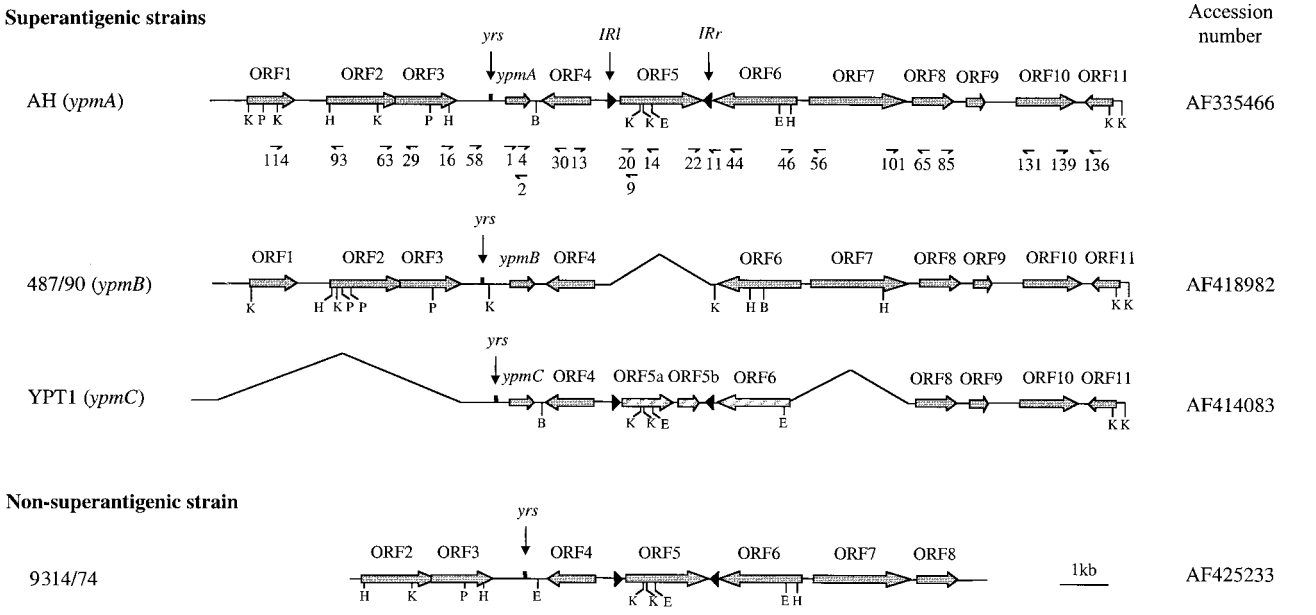


FIG. 1. Genetic organization of the *ypm* gene loci from superantigenic *Y. pseudotuberculosis* strains AH, 487/90, and YPT1 and the corresponding locus from the nonsuperantigenic strain 9314/74. Plain grey and hatched arrows represent complete and truncated ORFs, respectively. *yrs*, *Yersinia* recombination site; *IRl* and *IRr*, left and right inverted repeats, respectively. Relevant oligonucleotides were located on the genetic map of the *ypmA* locus of *Y. pseudotuberculosis* AH. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; P, *Pst*I.

Sequence comparisons were performed with Clustal W at http://www.infobiogen.fr.

Detection of *ypmA* spontaneous mutants. In order to evaluate the stability of the *ypmA* superantigen gene, *Y. pseudotuberculosis* AH was subcultured alternately at 28°C for 16 h and at 42°C for 8 h. Dilutions of bacterial cultures were plated, with isolated colonies then being tested for the presence of the *ypmA* gene by colony hybridization (53) using the *ypm*-specific probe labeled with digoxigenin.

Introduction of the *aph(3′)-IIIa* and *sacB* genes into the *Y. pseudotuberculosis* AH genome. To determine the deletion frequency within the superantigen gene locus, we inserted the kanamycin resistance gene *aph(3′)-IIIa* and the *sacB* gene into *ypmA* of *Y. pseudotuberculosis* AH. The *sacB* gene product is toxic for gram-negative bacteria grown in the presence of sucrose (8). Only the strains in which deletion (or inactivation) of the *sacB* gene has occurred will grow on sucrose. To insert the *aph(3′)-IIIa* and *sacB* genes into the *ypmA* gene, a *Spe*I restriction site was generated within *ypmA* (250 bp from the start codon) and two *Xba*I sites were also created (one 550 bp upstream of the start codon and the other 386 bp downstream of the *ypmA* stop codon) by overlap extension using PCR (20, 27, 29). PCR amplification with no. 98Xba (5′GCTCTAGACCTTGGGCTCCGATATTGATCCATTCC3′) and no. 97Spe (5′ATTGCGCCACTAGTCCTAGTAAAATTAACGTCATATCTGCATTTAC3′) primers yielded an 823-bp fragment encompassing the upstream and 5′ regions of *ypmA*, whereas no. 96Spe (5′TTACTAGGACTAGTGGCGCAATGTAGGAGGCAATGGATGGGGAG3′) and no. 99Xba (5′GCTCTAGAATGCAGTAAAGAATCAGGGTGGTGTTAC3′) primers produced a 611-bp fragment covering the 3′ half of *ypmA* and its downstream sequence. Internal primers no. 96Spe and no. 97Spe were generated to give a 22-bp overlapping region containing an *Spe*I restriction site. Amplimers produced with no. 98Xba plus no. 97Spe and no. 96Spe plus no. 99Xba were purified, combined, and annealed by their 22-bp overlapping sequence and were then 3′ extended following the complementary strand. Finally, the resulting PCR product was amplified with external primers no. 98Xba and no. 99Xba to give a 1.4-kb fragment. PCR conditions have been previously described (20). The amplimer was digested with *Xba*I and cloned into pUC19. The *sacB* gene was PCR amplified with sac1 (5′GGACTAGTGGAGATCTGGCCCGTAGTCTGCAAATCCTTTT3′) and sac2 (5′GGACTAGTCCGCTCGAGGGTTAGGAATACGGTTAGCCATTTGCC3′) primers by using pCVD442 as a template. Primer sac1 contains an *Xho*I restriction site downstream of a *Spe*I site, whereas primer sac2 was synthesized with a *Spe*I site. The 1,877-bp fragment generated with sac1 and sac2 was purified, digested with *Spe*I, and cloned into the *Spe*I site generated within *ypmA*. Next, the kanamycin resistance gene *aph(3′)-IIIa* was PCR amplified from plasmid pUC1318-KmII by using primers kan1 (5′GCCGCTCGAGGGATTTCAGGGGGCAAGGCATAG3′) and kan2 (5′GCCGCTCGAGCAGAGTATGGACAGTTGCGGATG3′), which both contain an *Xho*I site. The *aph(3′)-IIIa* gene was cloned into the *Xho*I site located downstream of the *sacB* gene. The orientation of the two reporter genes within *ypmA* was established by PCR amplification. Finally, the *Xba*I fragment containing the *aph(3′)-IIIa* and *sacB* genes inserted into *ypmA* was cloned into the suicide vector pMM70413, which contains the chloramphenicol resistance gene *cat*. The plasmid was designated pCCY41.2. Allelic exchange was carried out between *Y. pseudotuberculosis* AH Nal^r and *E. coli* SM10λ*pir* (pCCY41.2). The first recombination event was selected for on LB agar with nalidixic acid and kanamycin. To select for the second recombination event and eliminate the suicide plasmid, we used the chloramphenicol bacteriostatic properties. Bacteria were grown in LB containing chloramphenicol until the optical density at 600 nm reached a value of between 0.2 and 0.3. Next, cycloserine was added to a final concentration of 1.8 mg/ml. This antibiotic kills bacteria in the growth phase but has no effect on bacteria in the static phase. Hence, *Y. pseudotuberculosis* clones that still contained the suicide plasmid grew in the presence of chloramphenicol and were killed by cycloserine. After this treatment, the chloramphenicol-sensitive, kanamycin-resistant *Y. pseudotuberculosis* organisms were selected for on LB agar in the presence of kanamycin. Genetic analysis by PCR of the kanamycin-resistant clones confirmed the integration of the *aph(3′)-IIIa* and *sacB* genes into the *ypmA* gene.

As a control, *aph(3′)-IIIa* and *sacB* genes were inserted into another *Y. pseudotuberculosis* AH chromosomal region, upstream of the promoter controlling the expression of the urease operon (54). First, an *Spe*I restriction site was introduced into a noncoding region upstream of the urease operon promoter by overlap extension, as described above. External primers UreXbaF (5′GCTCTAGAGCGGTGGTAGCCCATGGTTGCAGTCAC3′) and UreXbaR (5′GCTCTAGAGCTAAAATCAAGACAAATTATCCACCACCC3′) contain an *Xba*I site, whereas internal primers UreSpeF (5′GCGATTGGACTAGTGGAAATTAAATATCACAGCTAGAATAATAACTAAT3′) and UreSpeR (5′TAATTTCCACTAGTCCAATCGCATTCCACGGTTCTTTTTAGTATTAACC3′) contain an *Spe*I site and share a 22-bp overlapping region. The resulting fragment was digested with *Xba*I and cloned into pUC19 and then into the pMM70413 suicide vector. The *Spe*I fragment of pCCY41.2 containing the *aph(3′-)IIIa* and *sacB* genes was introduced into the newly generated *Spe*I site, and the suicide plasmid was designated pIL31.3. Allelic exchange for introduction of the *aph(3′)-IIIa* and *sacB* genes upstream from the urease operon of *Y. pseudotuberculosis* AH Nal^r was carried out as above.

TABLE 3. ORFs identified upstream and downstream of the *ypmA* gene of *Y. pseudotuberculosis* AH

CDS[a]	Gene		Product				
				Protein similarity			
	Size (nt)[b]	G+C (%)	Size (aa)[c]	Relevant homologous protein	% Amino acid identity (ratio)	Putative function	Accession no.
ORF1	1,008	49.1	335	GalR of *E. coli*	41 (136/328)	Transcriptional regulator	AAA83860
ORF2	1,455	48.5	484	NhaC of *Haemophilus influenzae*	47 (219/458)	Na^+/H^+ antiporter	Q57007
ORF3	1,179	45.2	392	MalY of *E. coli*	36 (140/388)	Aminotransferase	P23256
ypmA	456	35.3	151	*Y. pseudotuberculosis*-derived mitogen	100	Superantigenic toxin	D38638 D38523
ORF4	924	36.5	307	PagO of *S. enterica* serovar Typhimurium YomA of *Y. enterocolitica*	77 (224/290) 59 (173/289)	Membrane protein	AAB82452 AAD16869
ORF5	1,599	55.6	532	No homology found			
ORF6	1,602	50.9	533	ExaC of *P. aeruginosa*	73 (372/506)	Aldehyde dehydrogenase	AAC79659
ORF7	1,845	51.1	614	AdhR of *Pseudomonas stutzeri*	35 (226/632)	Transcriptional regulator	AAG0924
ORF8	750	47.7	249	YdfG of *E. coli*	65 (163/247)	Oxidoreductase/dehydrogenase	F64908
ORF9	375	44.3	124	Hypothetical protein b1583 of *E. coli*	45 (43/95)	Unknown	A64914
ORF10	1,281	52.2	426	b1592 of *E. coli*	42 (179/424)	Chloride channel	AAC74664
ORF11	669	43.3	222	DTBS of *E. coli*	63 (141/223)	Dethiobiotin synthetase	BAA15317

[a] CDS, coding sequence.
[b] nt, nucleotides.
[c] aa, amino acids.

Measurement of the DNA deletion frequency. Dilutions of overnight cultures of *Y. pseudotuberculosis* AH containing the *aph(3′)-IIIa* and *sacB* genes were plated concomitantly on LB agar lacking sodium chloride in the presence of 10% sucrose on one hand and on LB agar alone on the other. The selection frequency for sucrose-resistant mutants was calculated as the ratio of the number of CFU growing on 10% sucrose agar to the number of CFU counted on LB agar without sucrose. To confirm that sucrose resistance was due to a deletion process rather than to a point mutation in *sacB*, deletions of the kanamycin resistance gene were sought. For this purpose, colonies growing on 10% sucrose were streaked onto LB agar in the presence of kanamycin, and the percentage of kanamycin-sensitive bacteria among the sucrose-resistant clones was established. Finally, the deletion frequency of the *aph(3′)-IIIa* and *sacB* genes was calculated as the frequency of appearance of the sucrose-resistant mutant multiplied by the percentage of kanamycin-sensitive bacteria among the sucrose-resistant clones. Frequency values were obtained from three separate experiments.

Nucleotide sequence accession numbers. The nucleotide sequences of the *ypm* loci have been submitted to GenBank and were given accession numbers AF335466 (*Y. pseudotuberculosis* strain AH), AF414083 (*Y. pseudotuberculosis* strain YPT1), AF418982 (*Y. pseudotuberculosis* strain 487/90), and AF425233 (*Y. pseudotuberculosis* strain 9314/74).

RESULTS

The *ypmA* gene is not associated with a mobile genetic element but is inserted in the chromosome in the vicinity of a DNA recombination site. In order to define the genetic environment of the superantigen gene of *Y. pseudotuberculosis* strain AH, a 17.1-kb DNA segment encompassing the *ypmA* gene was sequenced and 11 ORFs were identified (Fig. 1). The same genetic organization was found in 10 other *ypmA*$^+$ strains (KM, MA, NT, ST, 1191/90, 1402, 1404, 1553, 1554, and 1833) as shown by PCR amplification of the intergenic regions with primers located in ORFs 1 to 11 (Fig. 1). The *ypmA* locus of strain 1833 differs from strain AH only by the absence of ORF5. These ORFs encode proteins with homologies to regulatory proteins (ORF1 and ORF7), aminotransferases and synthetases (ORF3 and ORF11, respectively), dehydrogenases (ORF6 and ORF8), and membrane proteins (ORF2, ORF4, and ORF10) (Table 3). No function could be attributed to ORF5 or ORF9. It is noteworthy that the ORF4 product is highly homologous (77% identity) to PagO, a putative transmembrane protein regulated by the PhoP-PhoQ two-component system found in *Salmonella enterica* serovar Typhimurium (25) and, to a lesser extent (59% identity), to the *Y. enterocolitica* protein YomA, encoded by a gene adjacent to *yadA* on the virulence plasmid (56). Like *ypmA*, ORF4 has a low GC content (36.5%), contrasting with the higher GC content of the surrounding genes, which ranges from 43.3% for ORF11 to 55.6% for ORF5 (Table 3).

Another noticeable feature of the 17.1-kb locus is the presence (245 bp upstream of *ypmA*) of a 26-bp sequence that is homologous to a recombination sequence (*dif*) that promotes RecA-independent, site-specific recombination (33) and that might be responsible for hyperrecombination events in the *E. coli* chromosome terminus region (37) (Fig. 2). Another homologous sequence is present on the *E. coli* multidrug-resistant plasmid R1 and, like *dif*, is responsible for site-specific recombination (15, 16). Strong homology was also found with a phage integration locus found in *Xanthomonas campestris* (18). Thus, in light of the function attributed to these various sequences, the 26-bp sequence upstream of *ypmA* was designated *yrs* for *Yersinia* recombination site. Comparison of the nucleotide sequences revealed a 19-bp consensus sequence with two 6-bp inverted repeats (Fig. 2). Interestingly, this 19-bp common core sequence is widely distributed among various bacterial species, since it was found in the genomes of *Neisseria meningitidis*, *S. enterica* serovar Typhimurium, *Pseudomonas aeruginosa*, and *Pasteurella multocida* (Fig. 2).

A putative novel insertion sequence (IS), which we designated IS*1398*, was identified between ORF4 and ORF6 (Fig. 1). This 1.8-kb genetic element is composed of a single 1,599-kb ORF (ORF5) flanked by two inverted repeats (IRs) sharing 30 of 35 bp (Fig. 3). IS*1398* displays several features found in the Tn*1000* and Tn*3* transposon families. First, the two 35-bp IRs are homologous to the Tn*1000* and Tn*3* IRs (77.1 and 71.4% identity, respectively) (52) and to the terminal IRs of the *Pseudomonas* sp. strain BW13 mercury resistance Tn*3*-type transposon (42). Secondly, ORF5-specific Southern hybridization revealed that IS*1398*, like Tn*3* (35), was always present in a single copy in the *Y. pseudotuberculosis* genome

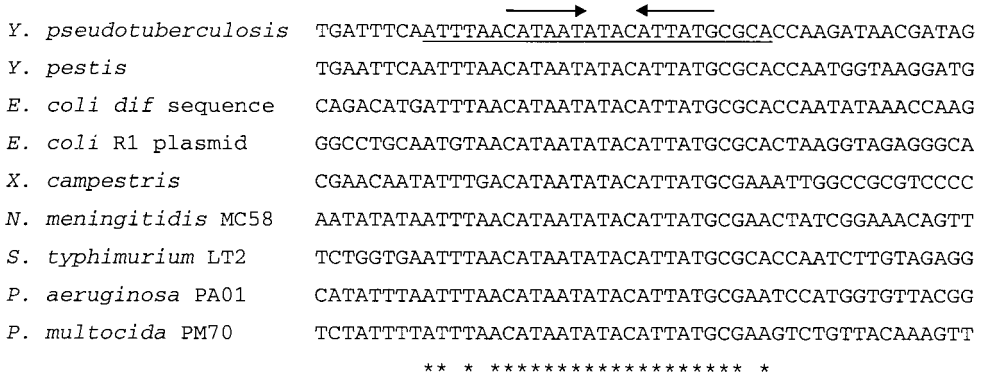


FIG. 2. Nucleotide sequence of the *Yersinia* recombination site (*yrs*) located upstream of *ypmA* aligned with the corresponding sequence found between ORFs YPO2283 and YPO2281 in *Y. pestis* CO92 (49), with the *E. coli dif* locus (33), with the recombination site found on plasmid R1 from *E. coli* (16), with the phage integration site in *X. campestris* (18), and with sequences from the genome of *N. meningitidis* strain MC58 (accession number AE002470), *S. enterica* serovar Typhimurium strain LT2 (accession number AE008767), *P. aeruginosa* strain PAO1 (accession number AE004648), and *P. multocida* strain PM70 (accession number AE006097). The underlined sequence corresponds to the *yrs* sequence. Asterisks indicate identical nucleotides, with arrows indicating the IR sequences.

(data not shown). Finally, a duplicate of a pentanucleotide (TTCAA) was present at the IS*1398* insertion site (52). However, apart from these features held in common with Tn*3*-type transposons, the 532-amino-acid ORF5 product located between the two IRs did not exhibit homology with any of the transposase genes known so far, thus suggesting a new type of IS or a Tn*3* transposon remnant that lost the transposase and resolvase genes. PCR amplification of ORF4-ORF5 and ORF5-ORF6 intergenic regions in 13 IS*1398*-positive *Y. pseudotuberculosis* strains (strains AH, KM, NT, ST, YPT1, YPT12, 199/90, 300/89, 367/89, 1191/90, 1402, 1432/94, and 9314/74) demonstrated that IS*1398* was always present at the same location in the genome, suggesting either an absence of mobility of this genetic element or the presence of a unique IS*1398* target site between ORF4 and ORF6.

A 918-bp fragment containing *ypm* is absent in most nonsuperantigenic *Y. pseudotuberculosis* strains. In the nonsuperantigenic strain 9314/74, sequence analysis of 12.3 kb of the *ypmA* corresponding locus showed the same genetic organization as in strain AH, except for the absence of the superantigen gene (Fig. 1). Comparative analysis of the ORF3-ORF4 intergenic region from strains 9314/74 and AH allowed us to define a 918-bp fragment containing *ypmA* in strain AH and missing in strain 9314/74. Interestingly, the *yrs* site was still present within the ORF3-ORF4 intergenic space. Particular sequences, such as direct or inverted repeats, were not found at the ends of this 918-bp sequence in strain AH. Sequencing of the PCR product given by primers no. 16 and no. 30 revealed the same ORF3-ORF4 intergenic space as 9314/74 in 10 nonepidemiologically related *ypm* mutant strains (strains YPT12, 199/90, 300/89, 367/89, 1401, 1403, 1432/94, 2821, 2843, and 2926). No amplification was produced from *Y. pseudotuberculosis* 1830: further investigation revealed that this strain lacked ORF4 to ORF6 and part of ORF7 and that a 1.6-kb fragment was inserted downstream of *yrs* (Fig. 4A). These results indicate that the absence of *ypm* in some nonsuperantigenic strains may be the consequence of a deletion process.

The genetic environment of the *ypmB* and *ypmC* alleles differs from that of the *ypmA* locus. To study the genetic environment of *ypmB* and *ypmC*, the two other alleles of the superantigenic gene, we cloned and sequenced the 15-kb *ypmB* locus of strain 487/90 and the 12.1-kb *ypmC* locus of *Y. pseudotuberculosis* YPT1. The main difference between the *ypmA* and *ypmB* loci lies in the absence of the IS*1398* genetic element in strain 487/90 (Fig. 1). Genes from the *ypmB* locus are between 97.5% (ORF4) and 88.7% (*ypm*) identical to the corresponding genes at the *ypmA* locus. Interestingly, *ypmB* has the lowest identity level (88.7%) among the 12 genes studied, suggesting a greater genetic drift for the superantigen gene. PCR analysis of the intergenic regions using primers located within the ORFs of other *ypmB*$^+$ strains from our collection (1093, 1096, 1155, and 1119) indicated a 1.1-kb insertion between ORF10 and ORF11 in strains 1096, 1119, and 1155.

When the superantigen gene loci of AH and YPT1 (*ypmC* locus) were compared, we found a 4,380-bp deletion (encompassing ORF1 to ORF3) upstream of *ypmC*, together with a 2,100-bp deletion downstream of ORF5 (including part of ORF6 and ORF7) (Fig. 1). Analysis of the sequence at the boundaries of the second deletion site showed that the excision had occurred at two heptanucleotide sequences (CCAATAC), suggesting a site-specific deletion. This heptanucleotide was not found flanking the first deletion. Furthermore, a frameshift due to deletion of a dinucleotide was observed within ORF5, giving rise to two smaller hypothetical proteins, ORF5a and ORF5b (Fig. 1). Besides these deletions, the genes of the *ypmC* locus of *Y. pseudotuberculosis* are over 99% identical to the corresponding genes of the *ypmA* locus. The genetic organization of the *ypmC* locus is not specific to the strain YPT1, since all other *ypmC*$^+$ strains tested (WE31/93, YPT5, 200/90, 298/89, 304/89, 776/88, 1134/90, 1216/93, 2887, 32945, 32975, 32977, and 32992) have the same genetic organization, as judged by PCR analysis of the intergenic regions with primers located within the ORFs. However, we found that the deletion upstream of *ypmC* was about 400 bp larger in strains 2887, 32945, 32977, and 32992.

The *Y. pseudotuberculosis ypmA*-containing locus is not

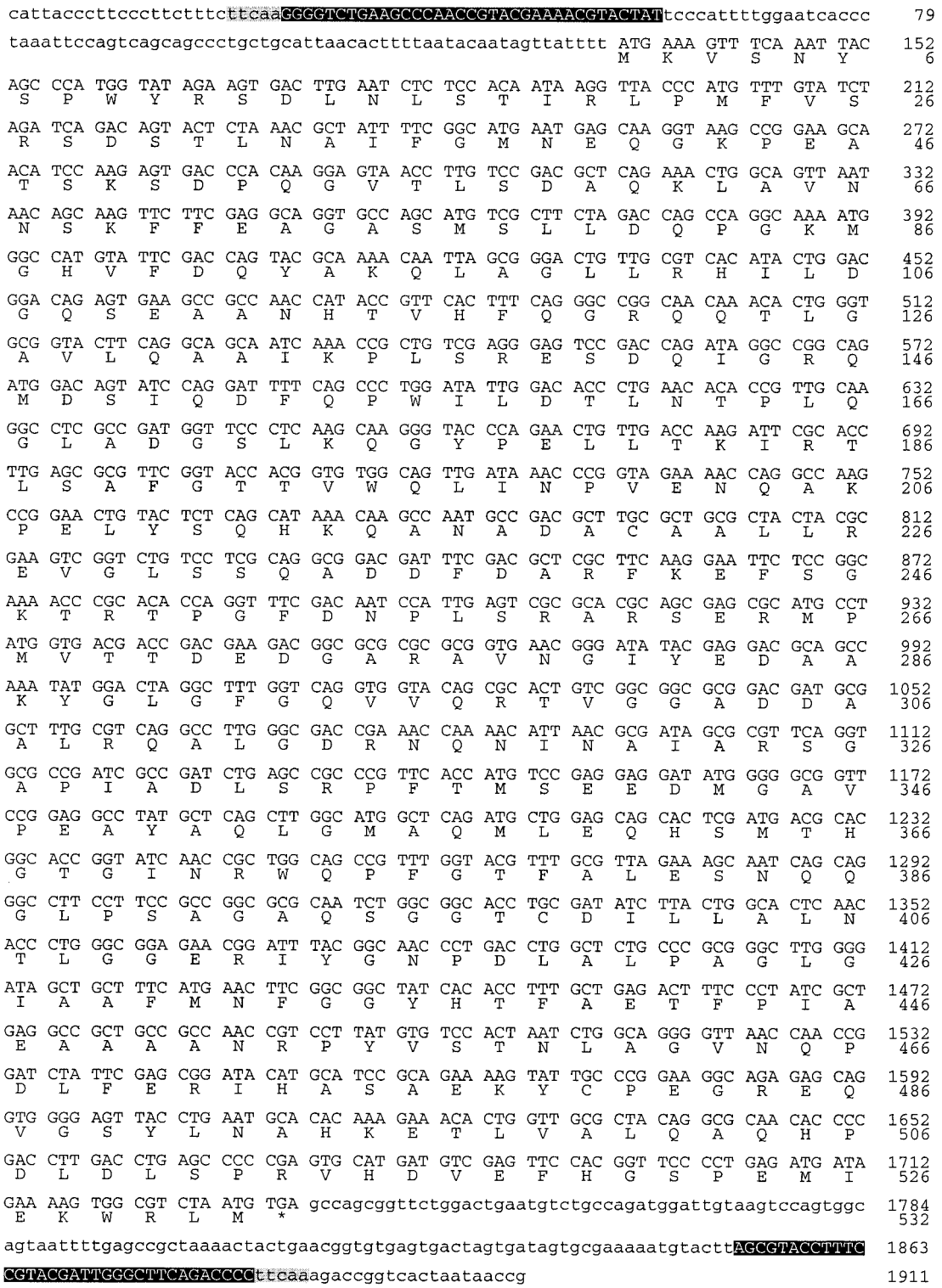


FIG. 3. Nucleotide sequence of IS*1398* and its immediate environment from *Y. pseudotuberculosis* AH. Black boxes represent the 35-bp IRs, and grey boxes indicate the 5-bp pair repeats at the target site. The stop codon is marked with an asterisk.

present in the *Y. pestis* genome. Since *Y. pseudotuberculosis* and *Y. pestis* are genetically related species (2, 45), we compared the genetic organizations of the *ypmA* loci of these two microorganisms (Fig. 5). Surprisingly, the 13.3-kb region of *Y. pseudotuberculosis* including ORF2 to ORF9 was not found in the genomes of *Y. pestis* CO92 and *Y. pestis* KIM5 when the *Y. pseudotuberculosis* AH sequence was compared with those presented by the *Y. pestis* sequencing group at the Sanger Center

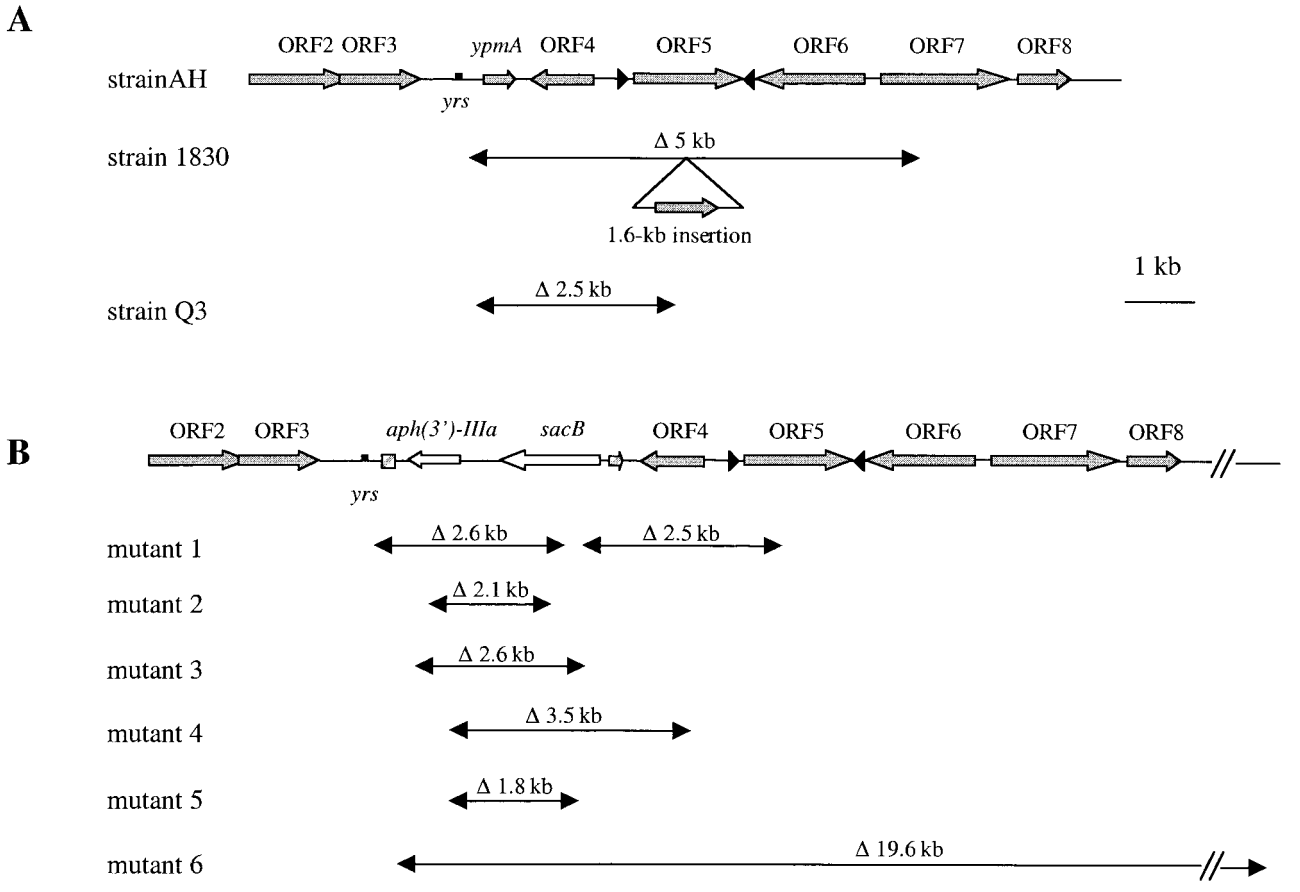


FIG. 4. Location of deletions (Δ) observed in the *ypm* locus of *Y. pseudotuberculosis.* (A) Positions of the deletions in strain Q3, a *Y. pseudotuberculosis* AH derivative spontaneously cured of *ypmA*, and in nonsuperantigenic clinical isolate 1830. (B) Deletions in mutants 1 to 6 generated after insertion of *aph(3')-IIIa* and *sacB* reporter genes into *ypmA* of *Y. pseudotuberculosis* strain AH with selection on sucrose LB agar and then on LB agar in the presence of kanamycin. The hatched arrow represents an incomplete *ypmA* gene. *yrs*, *Yersinia* recombination site.

(49) and by the University of Wisconsin Genome project, respectively. Twelve ORFs located in a 9.3-kb locus were present between ORF1 (YPO2283) and ORF10 (YPO2270) in *Y. pestis*, according to the recent genome annotation (49) (Fig. 5). The products of the YPO2274 and YPO2275 genes were similar to bacteriophage I2-2 proteins, whereas YPO2278 and YPO2280 gene products were similar to *Vibrio cholerae* filamentous bacteriophage proteins. Interestingly, a 581-bp region of *Y. pseudotuberculosis* upstream of *ypmA* and containing the *yrs* site was also found in *Y. pestis*. However, in *Y. pestis* this region was not found intact but was divided into two unequal parts flanking *Y. pestis* ORFs: a first 411-bp part containing the intact *yrs* site was located between ORF1 and YPO2281, whereas a second part (187 bp) containing half of the *yrs* site (TACATTATGCGCA) was found upstream of ORF10 (Fig. 5). This genetic organization demonstrates that *Y. pestis* phage genes were inserted within this 581-bp region and that this insertion generated a duplication of half of the *yrs* site. This strongly suggests that *yrs* represents a target site for phage insertion.

Deletion of the *ypmA* locus occurs with a high frequency. Several arguments from the above nucleotide sequence analysis suggested that the *ypm* locus is unstable: first, partial deletions within the *ypm* locus were detected in strain 1830 (Fig. 4A) and in *ypmC*$^+$ strains (Fig. 1), and second, the region is not present in *Y. pestis*, a species considered to be an emerging clone of *Y. pseudotuberculosis* (2) (Fig. 5). In order to address the question of the *ypm* region instability, we evaluated the frequency of appearance of spontaneous *ypmA* mutants. Hence, *Y. pseudotuberculosis* strain AH was subcultured alternately at 28 and 42°C in LB broth. Bacteria were tested for the presence of the superantigen gene by colony blot hybridization with a *ypmA*-specific probe. Among 12,000 *Y. pseudotuberculosis* clones examined, 1 clone (designated Q3) lost not only the *ypmA* gene but also ORF4 and part of IS*1398*, according to PCR amplification (Fig. 4A). Sequencing of the PCR products obtained with primers no. 16 and no. 11 revealed a 2,520-bp deletion in Q3. The size of the excision in Q3 contrasts with the 918-bp fragment missing in nonsuperantigenic strains such as 9314/74. Excepting the fact that the DNA fragment deletion occurred in an AT-rich region, Q3 and strain 1830 shared no obvious common genetic features at the deletion site.

In order to characterize the deletion process (frequency and site specificity) responsible for the *ypmA* loss observed in strain Q3, we inserted a selection gene [kanamycin resistance gene *aph(3')-IIIa*] and a counterselectable marker (the *sacB* levane sucrase-encoding gene) into the *ypmA* gene of *Y. pseudotuberculosis* AH. Since the product of the *sacB* gene is toxic for

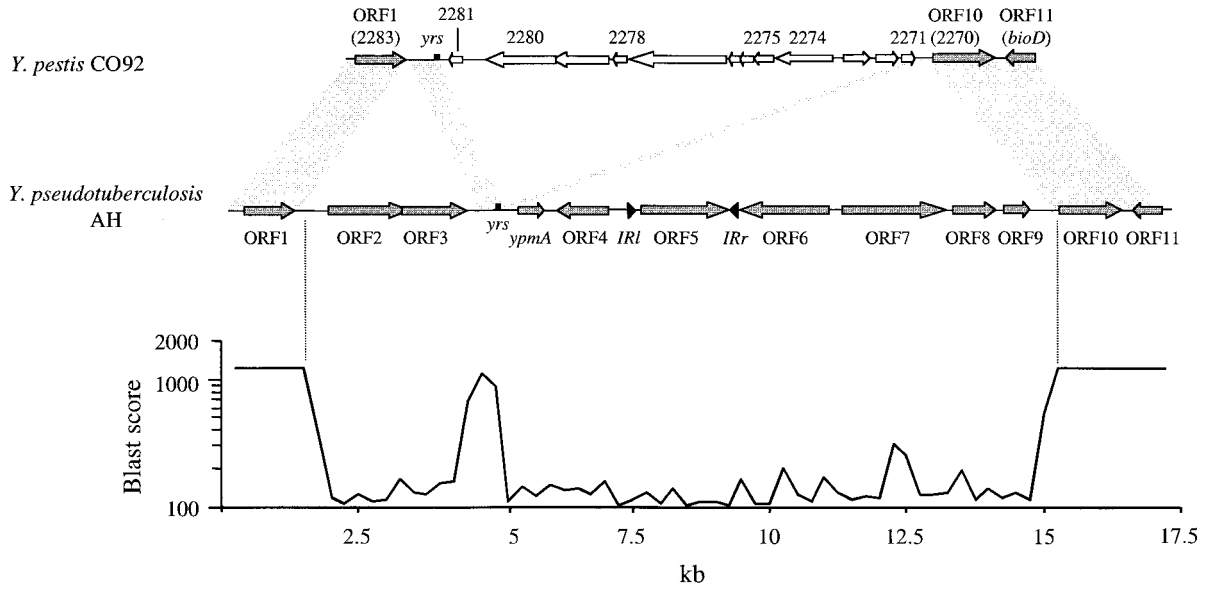
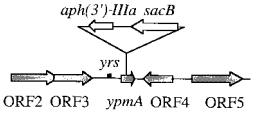
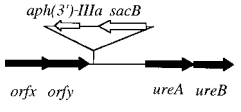


FIG. 5. Genetic organization of the *Y. pseudotuberculosis* AH locus containing the *ypmA* gene compared to the corresponding locus in *Y. pestis* CO92 biotype Orientalis (positions 285894 to 298332 of the genome sequence available under accession number AJ414151 [49]). Two hundred and fifty-base fragments were compared by using WU-Blast 2.0 software, and results were expressed as Blast scores. A score of 1250 represents 100% identity on 250 nucleotides. White arrows indicate *Y. pestis* ORFs that differ from *Y. pseudotuberculosis* ORFs. Shaded areas represent sequences common for *Y. pestis* and *Y. pseudotuberculosis*. Identical results were obtained when the *Y. pseudotuberculosis* AH sequence was compared with the genomic sequence of *Y. pestis* KIM5 strain P12 (http://www.genome.wisc.edu). Numbers on the *Y. pestis* sequence correspond to the annotation of the *Y. pestis* genome (49). *yrs*, *Yersinia* recombination site; *bioD*, dethiobiotine synthetase gene; *IRl* and *IRr*, left and right inverted repeats, respectively.

gram-negative bacteria in the presence of sucrose, only clones from which *sacB* is deleted will grow on agar containing this sugar. Sucrose-resistant *Yersinia* appeared with high frequency ($2.8 \times 10^{-5} \pm 7 \times 10^{-7}$), and among these sucrose-resistant clones, 21.2% ± 4.2% were sensitive to kanamycin (Table 4). Overall, deletion of the *sacB* and *aph(3′)-IIIa* genes occurred with a frequency of 5.9×10^{-6} (Table 4). Six of these sucrose-resistant, kanamycin-sensitive clones were further characterized (Fig. 4B). DNA deletions ranged from 1.8 to 19.6 kb, none of which were found twice. Comparison of 100 bp at the upstream and downstream deletion sites did not reveal any features common to the mutants, thus confirming that the deletion mechanism is not site specific. The *yrs* site was preserved in all mutants, suggesting that these 26-bp motifs might be important for the deletion process.

As a control, we placed the two reporter genes [*aph(3′)-IIIa* and *sacB*] upstream of the *Y. pseudotuberculosis* AH urease operon. This chromosomal region was chosen because as far as we know, no *Y. pseudotuberculosis* strain has been found to be deficient in urease activity due to deletion of its locus, indicating that the region is stable. Selection on 10% sucrose and then on kanamycin medium revealed that the occurrence of the sucrose-resistant, kanamycin-sensitive clones was much less frequent when the reporter genes were inserted upstream of

TABLE 4. Deletion frequencies of the *aph(3′)IIIa* and *sacB* genes as a function of their location in the *Y. pseudotuberculosis* AH genome

Location of the *aph(3′)-IIIa* and *sacB* genes	Frequency		
	Sucrose-resistant clones[a]	Kanamycin-sensitive clones[b]	Sucrose-resistant, kanamycin-sensitive clones[c]
	$2.8 \times 10^{-5} \pm 7 \times 10^{-7}$	21.2% ± 4.2%	5.9×10^{-6}
	$1.7 \times 10^{-6} \pm 2.7 \times 10^{-7}$	1.4% ± 0.4%	2.4×10^{-8}

[a] Mean value of three separate experiments ± standard deviation.
[b] Sucrose-resistant clones were plated on LB agar in the presence of kanamycin. The percentage represents the ratio of kanamycin-sensitive bacteria to the sucrose-resistant bacterial population (mean value of three separate experiments ± standard deviation).
[c] Frequency of sucrose-resistant bacteria multiplied by the proportion of kanamycin-sensitive bacteria to the sucrose-resistant bacterial population.

the urease locus than when they were located within *ypmA* (2.4×10^{-8} versus 5.9×10^{-6}, respectively) (Table 4). This means that deletions in the *ypmA* locus occur 245 times more frequently than those in the urease locus, demonstrating an instability of the chromosomal locus containing the superantigen gene *ypmA*.

DISCUSSION

Bacterial superantigenic toxin genes have been frequently associated with mobile genetic elements (21, 34, 40). Staphylococcal enterotoxin SEA and SEE genes and streptococcal exotoxin genes (*speA*, *speC*, *speH*, and *speI*) are found on phage DNA (7, 17, 21, 23, 32), whereas the enterotoxin SED gene and to a lesser extent the SEB gene are plasmid borne (3, 4, 31). Furthermore, the toxic shock syndrome toxin 1 gene and the recently described enterotoxin genes (*sei*, *seg*, *sek*, *sel*, *sem*, *seo*, and *sep*) appear to be genetically linked to four different staphylococcal pathogenicity islands (34, 36, 40). In this study, we demonstrate that *ypm* genes are chromosomally encoded and that although there is a putative phage integration site (*yrs*) 245 bp upstream of the *ypm* genes, bacteriophage proteins were not found in their vicinity. This rules out the association of a specific phage with the superantigen genes. Sequence analysis also showed that the *ypm* genes were associated neither with a transposon nor with a pathogenicity island as defined by Hacker and Kaper (26). For the 30 superantigen-producing strains that we tested, the *ypm* genes were always found at the same location in the *Y. pseudotuberculosis* genome, between ORF3 and ORF4. This feature is a strong argument that prompts us to question whether the *ypm* genes are mobile. Indeed, a number of different chromosomal locations of the *ypm* genes, as is the case for SEA (6), would have suggested mobility of the superantigen-encoding genes. Taken together, sequence analyses of the *ypm* locus demonstrated that the genetic organization of the superantigen gene of *Y. pseudotuberculosis* is different from that found for other superantigenic toxins.

Once the absence of an association of the *ypm* genes with mobile genetic elements was established, the following question remained: have nonsuperantigenic *Y. pseudotuberculosis* strains lost the superantigen gene or was *ypm* never taken up by these genomes? Nucleotide sequence analysis showed that the ORF3-ORF4 intergenic region was identical in 10 of 11 epidemiologically unrelated, nonsuperantigenic strains. Since *ypm* genes are not flanked by characteristic DNA sequences such as IRs, it is unlikely that a deletion occurred at the exact same location in these nonsuperantigenic strains. This suggests that the absence of the 918-bp, *ypm*-containing fragment between ORF3 and ORF4 is due to the nonincorporation of *ypm* rather than to a deletion of the gene. To test this hypothesis, we looked for spontaneous mutants and isolated mutant Q3, which did not display the ORF3-ORF4 intergenic region found in nonsuperantigenic strains but presented a larger deletion of 2.5 kb. Furthermore, when we selected deletion mutants with *aph(3')-IIIa* and *sacB* genes, none of the mutants resembled the wild-type nonsuperantigenic strains (Fig. 4). Hence, sequence analyses and experimental data all indicate that nonsuperantigenic *Y. pseudotuberculosis* probably never integrated a superantigen gene into their genome.

Although *Y. pestis* and *Y. pseudotuberculosis* are closely related genetically, the 13.3-kb region between ORF1 and ORF10 containing *ypmA* was not present in the *Y. pestis* genome (strains CO92 and KIM5). Instead, *Y. pestis* displays a 9.3-kb locus containing 13 ORFs, corresponding to phage remnants (Fig. 5) (49). Interestingly, the only common nucleotide sequence between *Y. pestis* and *Y. pseudotuberculosis* in the ORF1-ORF10 intergenic space was a 581-bp region containing the *yrs* site. Sequence analysis of the boundary regions in *Y. pestis* revealed that the phage genes were inserted within this 581-bp region (Fig. 5) and that a part of the *yrs* site was duplicated. This demonstrates that, like in *X. campestris* (18), the *yrs* site can function as a phage integration site. Genetic analysis indicates how *Y. pestis* acquired the phage genes but cannot explain the absence of ORF2 to ORF9. Since *Y. pestis* is considered to be a clone of *Y. pseudotuberculosis* that emerged 1,500 to 20,000 years ago (2), one can speculate that the 13.3-kb locus containing *ypmA* (except the *yrs* region) was deleted from the *Y. pestis* genome and that this event was followed by the insertion of the bacteriophage genes. However, we can also hypothesize that the *Y. pestis* genome never harbored the ORF2 to ORF9 genes, suggesting that the ancestral *Y. pseudotuberculosis* clone (from which *Y. pestis* derived) did not contain these genes either. Comparison of the genetic organization in the two *Yersinia* species also suggested that the ORF2-ORF9 region of *Y. pseudotuberculosis* might represent a pathogenicity islet as described in *Salmonella* (24). The absence of significant inverted or direct repeats flanking this element rules out the designation of the *Y. pseudotuberculosis* 13.3-kb sequence as a pathogenicity islet.

Other elements indicated the occurrence of genetic events in this genomic region of *Yersinia*. First, we found a wild-type strain of *Y. pseudotuberculosis* (strain 1830) with a 5-kb deletion within the *ypm* locus followed by a 1.6-kb insertion of an unknown ORF (Fig. 4). Second, the strains expressing the YPMc variant—which clearly represent a clonal population (12)—display two deletions in the *ypm* locus (Fig. 1). Third, we were able to demonstrate experimentally the occurrence in the locus of high-frequency deletion without precise excision (strain Q3 and mutants 1 to 6) (Fig. 4; Table 4). Taken as a whole, these arguments strongly indicate the genetic instability of the region containing the *ypm* genes.

In this work, we also described IS*1398*, a novel IS with no homology to any known mobile genetic element. IS*1398* displays some features of the Tn*3* family, that is to say, (i) a single copy per genome (35), (ii) 35-bp IRs ending the IS and homologous to the γδ sequences of Tn*1000* (52), and (iii) a pentanucleotide duplication at the genomic integration site (52) (Fig. 3). However, the functionality of this genetic element remains unproved. The small size of IS*1398* (1.8kb for IS*1398* versus 4.9 and 5.7 kb for Tn*3* and Tn*1000* [38], respectively) and the absence of homology of ORF5 with transposase (*tnpA)* and resolvase (*tnpR*) genes might suggest that IS*1398* represents a Tn*3*-like transposon remnant. Because of the proximity of *ypmA* and IS*1398*, it was initially tempting to associate the presence of IS*1398* with the heterogeneous distribution of the *ypm* genes among the *Y. pseudotuberculosis* strains. We clearly demonstrated that there was no close genetic relationship between IS*1398* and *ypm.* Indeed, some strains harboring IS*1398* do not contain a *ypm* gene and conversely some superantigen-

expressing strains were IS*1398*-free. Furthermore, IS*1398* is not associated with the ancestral gene *ypmB,* ruling out a role of IS*1398* in the initial integration of *ypmB* into the *Y. pseudotuberculosis* genome.

Nucleotide sequence analysis of the *ypm* genes and their flanking regions sheds some light on the *ypm* gene evolution. The GC contents of the *ypm* genes (34.6% for *ypmB* and 35.3% for *ypmA* and *ypmC*) with regard to that of the *Y. pseudotuberculosis* genome as a whole (46.5%) (5) suggest a horizontal gene transfer from a microorganism with low GC content. Superantigen-producing species such as *Streptococcus pyogenes*, *Staphylococcus aureus*, and *Mycoplasma arthritidis* could be good candidates, since their GC contents are low (38.5, 32.8, and 30 to 32.6%, respectively) (21, 34, 51). However, the codon usage of *ypmB,* which is supposed to be the ancestral *ypm* gene, according to a recent epidemiological study (12), is different from that found in group A streptococci, *S. aureus*, or *M. arthritidis* (data not shown). This indicates either another origin for the *ypm* genes or a drift of the *ypm* sequence towards the *Yersinia* GC content. In the near future, the various sequencing projects of bacterial genomes may well generate important information on the possible origin of the *ypm* genes. How the *ypmB* gene originally integrated into the genome of *Y. pseudotuberculosis* without the presence of flanking IRs or specific sites is still speculative. The most plausible hypothesis is that *ypmB* integration into the *Y. pseudotuberculosis* genome was the consequence of a recombination event, as was the case for the BRO β-lactamase gene in *Moraxella catarrhalis* (9). Nevertheless, since yersiniae are non-naturally competent, the *bro* gene entry mechanism in *M. catarrhalis* cannot be extended to *Y. pseudotuberculosis*. Hence, the uptake of *ypmB* by *Y. pseudotuberculosis* probably required the presence of a mobilizing structure, such as a phage or a plasmid. Considering the presence of a sequence with homology to a phage integration site (the *yrs* site) in the *ypm* locus and the association of this site with phage proteins in *Y. pestis*, we can speculate on the involvement of a phage in the uptake of the *ypm* genes. After the initial integration of *ypmB* into the *Y. pseudotuberculosis* genome, the superantigen gene evolved to give *ypmA* and *ypmC*. Interestingly, a recent epidemiological study bringing together more than 2,200 strains indicated that all *ypmB*$^+$ strains were found in nonhuman hosts and in the environment, whereas the *ypmA* and *ypmC* alleles were present in human pathogenic strains (22). This suggests that the genetic drift of the *Y. pseudotuberculosis* superantigen gene from *ypmB* towards the *ypmA* and *ypmC* alleles might be correlated with the increased virulence of *Y. pseudotuberculosis*.

ACKNOWLEDGMENTS

C. Carnoy was supported by a grant from the Centre Hospitalier Régional et Universitaire de Lille and by the Fondation pour la Recherche Médicale. This work was supported by the Region Nord-Pas de Calais and the European Regional Development Fund.

We thank J.-P. Bohin for critical reading of the manuscript and Ingrid Loison for technical assistance.

REFERENCES

1. **Abe, J., T. Takeda, Y. Watanabe, H. Nakao, N. Kobayashi, D. Y. M. Leung, and T. Kohsaka.** 1993. Evidence for superantigen production by *Yersinia pseudotuberculosis*. J. Immunol. **151:**4183–4188.
2. **Achtman, M., K. Zurth, G. Morelli, G. Torrea, A. Guiyoule, and E. Carniel.** 1999. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. Proc. Natl. Acad. Sci. USA **96:**14043–14048.
3. **Altboum, Z., I. Hertman, and S. Sarid.** 1985. Penicillinase plasmid-linked genetic determinants for enterotoxins B and C1 production in *Staphylococcus aureus*. Infect. Immun. **47:**514–521.
4. **Bayles, K. W., and J. J. Iandolo.** 1989. Genetic and molecular analyses of the gene encoding staphylococcal enterotoxin D. J. Bacteriol. **171:**4799–4806.
5. **Bercovier, H., and H. H. Mollaret.** 1984. Genus *Yersinia*, p. 498–506. *In* N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co., Baltimore, Md.
6. **Betley, M. J., S. Lofdahl, B. N. Kreiswirth, M. S. Bergdoll, and R. P. Novick.** 1984. Staphylococcal enterotoxin A gene is associated with a variable genetic element. Proc. Natl. Acad. Sci. USA **81:**5179–5183.
7. **Betley, M. J., and J. J. Mekalanos.** 1985. Staphylococcal enterotoxin A is encoded by a phage. Science **229:**185–187.
8. **Blomfield, I. C., V. Vaughn, R. F. Rest, and B. I. Eisenstein.** 1991. Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. Mol. Microbiol. **5:**1447–1457.
9. **Bootsma, H. J., H. van Dijk, P. Vauterin, J. Verhoef, and F. R. Mooi.** 2000. Genesis of BRO β-lactamase-producing *Moraxella catarrhalis*: evidence for transformation-mediated horizontal transfer. Mol. Microbiol. **36:**93–104.
10. **Butler, T.** 1983. Plague and other *Yersinia* infection. Plenum Press, New York, N.Y.
11. **Carnoy, C., H. Müller-Alouf, S. Haentjens, and M. Simonet.** 1998. Polymorphism of *ypm*, *Yersinia pseudotuberculosis* superantigen-encoding gene. Zentbl. Bakteriol. **29**(Suppl.)**:**397–398.
12. **Carnoy, C., and M. Simonet.** 1999. *Yersinia pseudotuberculosis* superantigenic toxins, p. 611–622. *In* J. E. Alouf and J. H. Freer (ed.), Bacterial protein toxins: a comprehensive sourcebook, 2nd ed. Academic Press, London, United Kingdom.
13. **Carnoy, C., C. Mullet, H. Müller-Alouf, E. Leteurtre, and M. Simonet.** 2000. Superantigen YPMa exacerbates the virulence of *Yersinia pseudotuberculosis* in mice. Infect. Immun. **68:**2553–2559.
14. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134:**1141–1156.
15. **Clerget, M.** 1984. A 140 base-pair DNA segment from the kanamycin resistance region of plasmid R1 acts as an origin of replication and promotes site-specific recombination. J. Mol. Biol. **178:**35–46.
16. **Clerget, M.** 1991. Site-specific recombination promoted by a short DNA segment of plasmid R1 and by a homologous segment in the terminus region of the *Escherichia coli* chromosome. New Biol. **3:**780–788.
17. **Couch, J. L., M. T. Soltis, and M. J. Betley.** 1988. Cloning and nucleotide sequence of the type E staphylococcal enterotoxin gene. J. Bacteriol. **170:** 2954–2960.
18. **Dai, H., T. Y. Chow, H. J. Liao, Z. Y. Chen, and K. S. Chiang.** 1988. Nucleotide sequences involved in the neolysogenic insertion of filamentous phage Cf16-v1 into the *Xanthomonas campestris* pv. citri chromosome. Virology **167:**613–620.
19. **Donnenberg, M. S., and J. B. Kaper.** 1991. Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. Infect. Immun. **59:**4310–4317.
20. **Donnenberg, M. S., J. Yu, and J. B. Kaper.** 1993. A second chromosomal gene necessary for intimate attachment of enteropathogenic *Escherichia coli* to epithelial cells. J. Bacteriol. **175:**4670–4680.
21. **Ferretti, J. J., W. M. McShan, D. Ajdic, D. J. Savic, G. Savic, K. Lyon, C. Primeaux, S. Sezate, A. N. Suvorov, S. Kenton, H. S. Lai, S. P. Lin, Y. Qian, H. G. Jia, F. Z. Najar, Q. Ren, H. Zhu, L. Song, J. White, X. Yuan, S. W. Clifton, B. A. Roe, and R. McLaughlin.** 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. Proc. Natl. Acad. Sci. USA **98:**4658–4663.
22. **Fukushima, H., Y. Matsuda, R. Seki, M. Tsubokura, N. Takeda, F. N. Shubin, I. K. Paik, and X. B. Zheng.** 2001. Geographical heterogeneity between Far Eastern and Western countries in prevalence of the virulence plasmid, the superantigen *Yersinia pseudotuberculosis*-derived mitogen, and the high-pathogenicity island among *Yersinia pseudotuberculosis* strains. J. Clin. Microbiol. **39:**3541–3547.
23. **Goshorn, S. C., and P. M. Schlievert.** 1989. Bacteriophage association of streptococcal pyrogenic exotoxin type C. J. Bacteriol. **171:**3068–3073.
24. **Groisman, E. A., and H. Ochman.** 1997. How *Salmonella* became a pathogen. Trends Microbiol. **5:**343–349.
25. **Gunn, J. S., W. J. Belden, and S. I. Miller.** 1998. Identification of PhoP-PhoQ activated genes within a duplicated region of the *Salmonella typhimurium* chromosome. Microb. Pathog. **25:**77–90.
26. **Hacker, J., and J. B. Kaper.** 2000. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. **54:**641–679.
27. **Haentjens-Herwegh, S., C. Carnoy, F. Wallet, and R. J. Courcol.** 2000. Development and use of internal positive controls for PCR detection of microorganisms, p. 547–556. *In* S. G. Pandalai (ed.), Recent research developments in microbiology, 4th ed., vol. 4. Transworld Research Network, Trivandrum, India.
28. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
29. **Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease.** 1989.

Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene **77:**51–59.
30. **Iteman, I., H. Najdenski, and E. Carniel.** 1995. High genomic polymorphism in *Yersinia pseudotuberculosis*. Contrib. Microbiol. Immunol. **13:**106–111.
31. **Johns, M. B., Jr., and S. A. Khan.** 1988. Staphylococcal enterotoxin B gene is associated with a discrete genetic element. J. Bacteriol. **170:**4033–4039.
32. **Johnson, L. P., M. A. Tomai, and P. M. Schlievert.** 1986. Bacteriophage involvement in group A streptococcal pyrogenic exotoxin A production. J. Bacteriol. **166:**623–627.
33. **Kuempel, P. L., J. M. Henson, L. Dircks, M. Tecklenburg, and D. F. Lim.** 1991. *dif*, a recA-independent recombination site in the terminus region of the chromosome of *Escherichia coli*. New Biol. **3:**799–811.
34. **Kuroda, M., T. Ohta, I. Uchiyama, T. Baba, H. Yuzawa, I. Kobayashi, L. Cui, A. Oguchi, K. Aoki, Y. Nagai, J. Lian, T. Ito, M. Kanamori, H. Matsumaru, A. Maruyama, H. Murakami, A. Hosoyama, Y. Mizutani-Ui, N. K. Takahashi, T. Sawano, R. Inoue, C. Kaito, K. Sekimizu, H. Hirakawa, S. Kuhara, S. Goto, J. Yabuzaki, M. Kanehisa, A. Yamashita, K. Oshima, K. Furuya, C. Yoshino, T. Shiba, M. Hattori, N. Ogasawara, H. Hayashi, and K. Hiramatsu.** 2001. Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. Lancet **357:**1225–1240.
35. **Lee, C. H., A. Bhagwat, and F. Heffron.** 1983. Identification of a transposon Tn*3* sequence required for transposition immunity. Proc. Natl. Acad. Sci. USA **80:**6765–6769.
36. **Lindsay, J. A., A. Ruzin, H. F. Ross, N. Kurepina, and R. P. Novick.** 1998. The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*. Mol. Microbiol. **29:**527–543.
37. **Louarn, J. M., J. Louarn, V. Francois, and J. Patte.** 1991. Analysis and possible role of hyperrecombination in the termination region of the *Escherichia coli* chromosome. J. Bacteriol. **173:**5097–5104.
38. **Maloy, S. R., J. E. Cronan, Jr., and D. Freifelder.** 1994. Microbial genetics, 2nd ed., p. 239–262. Jones and Bartlett Publishers, Inc., Sudbury, Mass.
39. **Marmur, J.** 1961. A procedure for the isolation of desoxyribonucleic acid from microorganisms. J. Mol. Biol. **3:**208–218.
40. **McCormick, J. K., J. M. Yarwood, and P. M. Schlievert.** 2001. Toxic shock syndrome and bacterial superantigens: an update. Annu. Rev. Microbiol. **55:**77–104.
41. **Miller, V. L., and J. J. Mekalanos.** 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. J. Bacteriol. **170:**2575–2583.
42. **Mindlin, S. Z., G. Y. Kholodii, Z. M. Gorlenko, S. V. Minakhina, L. S. Minakhin, E. S. Kalyaeva, A. V. Kopteva, M. A. Petrova, O. V. Yurieva, and V. G. Nikiforov.** 2001. Mercury resistance transposons of Gram-negative environmental bacteria and their classification. Res. Microbiol. **152:**811–822.
43. **Miyoshi-Akiyama, T., A. Abe, H. Kato, K. Kawahara, H. Narimatsu, and T. Uchiyama.** 1995. DNA sequencing of the gene encoding a bacterial superantigen, *Yersinia pseudotuberculosis*-derived mitogen (YPM), and characterization of the gene product, cloned YPM. J. Immunol. **154:**5228–5234.
44. **Miyoshi-Akiyama, T., W. Fujimaki, X. J. Yan, J. Yagi, K. Imanishi, H. Kato, K. Tomonari, and T. Uchiyama.** 1997. Identification of murine T cells reactive with the bacterial superantigen *Yersinia pseudotuberculosis*-derived mitogen (YPM) and factors involved in YPM-induced toxicity in mice. Microbiol. Immunol. **41:**345–352.
45. **Moore, R. L., and R. R. Brubaker.** 1975. Hybridization of deoxyribonucleotide sequences of *Yersinia enterocolitica* and other selected members of *Enterobacteriaceae*. Int. J. Syst. Bacteriol. **25:**336–339.
46. **Müller-Alouf, H., C. Carnoy, M. Simonet, and J. E. Alouf.** 2001. Superantigen bacterial toxins: state of the art. Toxicon **39:**1691–1701.
47. **Nakano, T., H. Kawaguchi, K. Nakao, T. Maruyama, H. Kamiya, and M. Sakurai.** 1989. Two outbreaks of *Yersinia pseudotuberculosis* 5a infection in Japan. Scand. J. Infect. Dis. **21:**175–179.
48. **Odaert, M., P. Berche, and M. Simonet.** 1996. Molecular typing of *Yersinia pseudotuberculosis* by using an IS*200*-like element. J. Clin. Microbiol. **34:**2231–2235.
49. **Parkhill, J., B. W. Wren, N. R. Thomson, R. W. Titball, M. T. Holden, M. B. Prentice, M. Sebaihia, K. D. James, C. Churcher, K. L. Mungall, S. Baker, D. Basham, S. D. Bentley, K. Brooks, A. M. Cerdeno-Tarraga, T. Chillingworth, A. Cronin, R. M. Davies, P. Davis, G. Dougan, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Leather, S. Moule, P. C. Oyston, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead, and B. G. Barrell.** 2001. Genome sequence of *Yersinia pestis*, the causative agent of plague. Nature **413:**523–527.
50. **Ramamurthy, T., K.-I. Yoshino, J. Abe, N. Ikeda, and T. Takeda.** 1997. Purification, characterization and cloning of a novel variant of the superantigen *Yersinia pseudotuberculosis*-derived mitogen. FEBS Lett. **413:**174–176.
51. **Razin, S., and E. A. Freundt.** 1984. The Mycoplasmas, p. 740–793. *In* N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co., Baltimore, Md.
52. **Reed, R. R., R. A. Young, J. A. Steitz, N. D. F. Grindley, and M. S. Guyer.** 1979. Transposition of the *Escherichia coli* insertion element γ generates a five-base-pair repeat. Proc. Natl. Acad. Sci. **76:**4882–4886.
53. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
54. **Sebbane, F., A. Devalckenaere, J. Foulon, E. Carniel, and M. Simonet.** 2001. Silencing and reactivation of urease in *Yersinia pestis* is determined by one G residue at a specific position in the *ureD* gene. Infect. Immun. **69:**170–176.
55. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram negative bacteria. Bio/Technology **1:**784–791.
56. **Snellings, N. J., M. Popek, and L. E. Lindler.** 2001. Complete DNA sequence of *Yersinia enterocolitica* serotype 0:8 low-calcium-response plasmid reveals a new virulence plasmid-associated replicon. Infect. Immun. **69:**4627–4638.
57. **Stuart, P. M., and J. G. Woodward.** 1992. *Yersinia enterocolitica* produces superantigenic activity. J. Immunol. **148:**225–233.
58. **Stuart, P. M., R. K. Munn, E. Demoll, and J. G. Woodward.** 1995. Characterization of human T-cell responses to *Yersinia enterocolitica* superantigen. Hum. Immunol. **43:**269–275.
59. **Tertti, R., K. Granfors, O.-P. Lehtonen, J. Mertsola, A.-L. Mäkelä, I. Välimäki, P. Hänninen, and A. Toivanen.** 1984. An outbreak of *Yersinia pseudotuberculosis* infection. J. Infect. Dis. **149:**245–250.
60. **Tertti, R., R. Vuento, P. Mikkola, K. Granfors, A.-L. Mäkelä, and A. Toivanen.** 1989. Clinical manifestations of *Yersinia pseudotuberculosis* infection in children. Eur. J. Clin. Microbiol. Infect. Dis. **8:**587–591.
61. **Uchiyama, T., T. Miyoshi-Akiyama, H. Kato, W. Fujimaki, K. Imanishi, and X.-J. Yan.** 1993. Superantigenic properties of a novel mitogenic substance produced by *Yersinia pseudotuberculosis* isolated from patients manifesting acute and systemic symptoms. J. Immunol. **151:**4407–4413.
62. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
63. **Yoshino, K.-I., T. Ramamurthy, G. B. Nair, H. Fukushima, Y. Ohtomo, N. Takeda, S. Kaneko, and T. Takeda.** 1995. Geographical heterogeneity between Far East and Europe in prevalence of *ypm* gene encoding the novel superantigen among *Yersinia pseudotuberculosis* strains. J. Clin. Microbiol. **33:**3356–3358.