JOURNAL OF CLINICAL MICROBIOLOGY, Nov. 2003, p. 5103–5112 Vol. 41, No. 11
0095-1137/03/$08.00+0 DOI: 10.1128/JCM.41.11.5103–5112.2003

Use of O-Antigen Gene Cluster-Specific PCRs for the Identification and O-Genotyping of *Yersinia pseudotuberculosis* and *Yersinia pestis*

Tatiana Bogdanovich,[1] Elisabeth Carniel,[2] Hiroshi Fukushima,[3] and Mikael Skurnik[1,4]*

Department of Medical Biochemistry, University of Turku, Turku,[1] Department of Bacteriology and Immunology, Haartman Institute, University of Helsinki and Helsinki University Central Hospital Laboratory Diagnostics, Helsinki,[4] Finland; Institut Pasteur, Paris, France[2]; and The Shimane Prefectural Institute of Public Health and Environmental Science, Matsue-shi, Shimane, Japan[3]

Received 6 June 2003/Returned for modification 29 July 2003/Accepted 12 August 2003

***Yersinia pestis* is a very recently evolved clone of *Yersinia pseudotuberculosis* serotype O:1b. This close relationship causes potential difficulties in DNA-based diagnostic methods. Analysis of the O-antigen gene clusters in these two organisms identified two regions that were used to specifically identify *Y. pestis-Y. pseudotuberculosis* as a group or *Y. pestis* alone. Both PCR assays were found to be 100% specific when tested on a large collection of *Yersinia* species and other *Enterobacteriaceae*. Furthermore, advantage was taken of the different setups of the O-antigen gene clusters of the 21 known *Y. pseudotuberculosis* serotypes to develop a multiplex PCR assay to replace the conventional serotyping method of *Y. pseudotuberculosis* by O-genotyping. The multiplex PCR assay contained nine sets of specific PCRs in a single tube and when used on *Y. pseudotuberculosis* reference strains allowed the distinction of 14 individual serotypes and two duplex serotypes (O:4a-O:8 and O:12-O:13). Serotype O:7, O:9, and O:10 strains required additional PCRs for O-genotyping. Once applied to *Y. pseudotuberculosis* strains of various origins, a very good correlation between classical serotypes and O-genotypes was observed, although some discrepancies were found. O-genotyping also proved useful to correct misidentification of some strains and to type *Y. pseudotuberculosis* isolates that had lost the expression of the O-antigen. The PCR-based O-genotyping can easily be applied in conventional laboratories, without the need for tedious preparation of a large set of specific antisera.**

Yersinia pseudotuberculosis and *Yersinia enterocolitica* are divided into serotypes based on differences in their lipopolysaccharide (LPS) O-side chain (O-antigen) antigenic determinants. *Yersinia pestis* does not have any serotypes because natural mutations in its O-antigen gene cluster prevent the synthesis of O-antigens (14). The genomes of *Y. pestis* and *Y. pseudotuberculosis* are almost identical; however, at least 149 genes in the *Y. pestis* genome are pseudogenes (10). Taxonomically, these two organisms should have been grouped into a single species (3); however, for safety reasons, they were kept as separate species. *Y. pestis* is regarded as a recently deviated clone of *Y. pseudotuberculosis* serotype O:1b (1, 14).

Despite high genetic relatedness, *Y. pestis* and *Y. pseudotuberculosis* have different epidemiological cycles and cause distinct types of diseases. *Y. pseudotuberculosis* is found either as a commensal or as a pathogen in a wide range of animals (birds, rodents, pigs, etc.) and is also recovered from food and water sources. Epizootics and human outbreaks may arise from these sources of contamination. *Y. pseudotuberculosis* causes a variety of intestinal (mesenteric adenitis and terminal ileitis) and extraintestinal (septicemia and liver diseases) infections in humans (5, 8), with high rates of postinfectious complications (erythema nodosum, arthritis, iritis, and nephritis).

Serotyping of *Y. pseudotuberculosis* strains has been the only means to routinely subtype different isolates for diagnosis and epidemiological purposes. The serotyping scheme, based on O-antigen, distinguishes 15 O-serotypes, of which serotypes O:1 and O:2 are divided into subtypes a, b, and c and serotypes O:4 and O:5 are divided into subtypes a and b, respectively. Serotypes O:1 to O:7 are characterized by the presence of serogroup a-specific 3,6-dideoxyhexoses (DDHs) in the O-antigen (e.g., paratose in serotypes O:1a, O:1b, and O:3; abequose in O:2; tyvelose in O:4; ascarylose in O:5a; and colitose in O:6 and O:7). Classically, serotyping is performed by slide agglutination of strains with serotype-specific antisera. To produce type-specific antisera, a set of *Y. pseudotuberculosis* reference strains are used to immunize rabbits (15). The rabbit immune sera are then absorbed on certain strains to remove cross-reacting antibodies and used for slide agglutination. Production of a reliable set of specific antisera is a tedious process done by only a few laboratories in the world. This conventional serotyping method is unable to type rough *Y. pseudotuberculosis* strains. In ordinary clinical microbiology laboratories, commercially available antisera are used to identify the most commonly isolated serotypes (e.g., O:1 and O:3). Therefore, a substantial number of strains remain untyped.

The genes responsible for the biosynthesis of the O-antigen are chromosomally located in *Yersinia* and are most often clustered between the *hemH* and *gsk* genes (13). The *Y. pestis* and *Y. pseudotuberculosis* O:1b O-antigen gene clusters were recently sequenced (14). Their genetic organizations were found to be highly conserved (Fig. 1), and their products were found to be of 99 to 100% identity, with the exception of the Wzx proteins, which were only 90.4% identical (14). The genetic organizations of the O-antigen clusters of the reference strains of *Y. pseudotuberculosis* representing all known serotypes were also determined by using specific primers (14). These clusters

* Corresponding author. Mailing address: University of Helsinki, Haartman Institute, Department of Bacteriology and Immunology, P.O. Box 63, 00014 University of Helsinki, Finland. Phone: 358-9-19125290. Fax: 358-9-19125302. E-mail: mikael.skurnik@helsinki.fi.

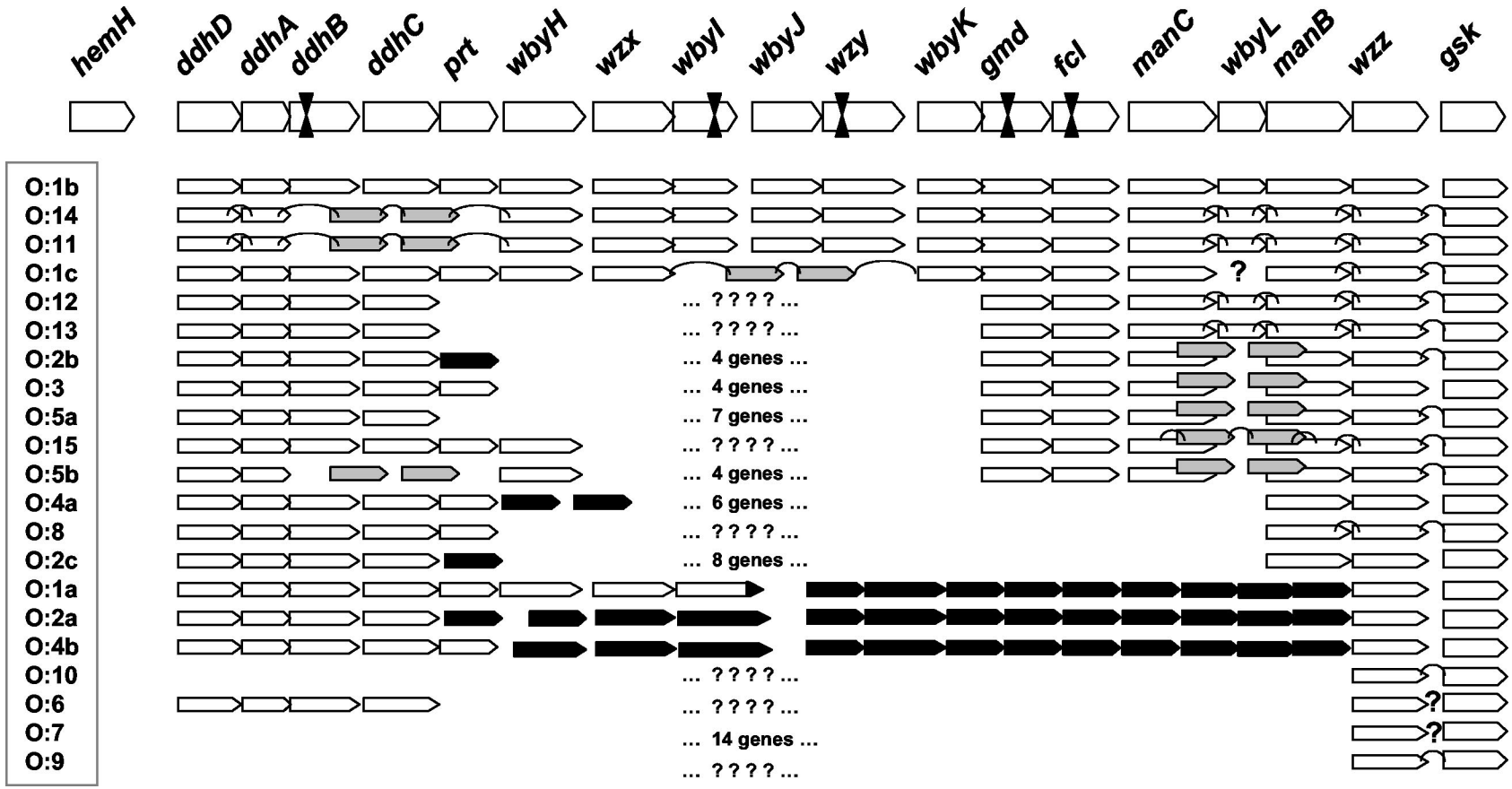

FIG. 1. Overview of the genetic composition of the O-antigen gene clusters of *Y. pestis* and *Y. pseudotuberculosis* serotypes O:1 to O:15. The O-antigen gene cluster of EV76 along with the inactivated genes indicated by inverted arrowheads is given at the top. Below it, the genetic setup of the gene clusters of different *Y. pseudotuberculosis* serotypes is presented. This figure is a compilation of published data (4, 9, 11, 14) and data obtained during this work. Open arrows indicate that a positive *Y. pestis* gene-specific PCR result was obtained, and gray arrows indicate that nonhomologous genes are present in these locations in these serotypes. The gray arrow data for serotypes O:1a, O:2a, O:4a and O:4b are from published sequences (9), and the data for serotypes O:1c, O:2b, O:3, O:5a O:5b, O:11, O:14, and O:15 are from the results of the present work. The number of genes present in the empty areas between the indicated genes in some serotypes is taken from (11). The empty areas with question marks indicate that the genetic setup has not been studied. The complete gene order is known for serotypes O:1a, O:1b, O:2a, O:2b, O:3, O:4a, O:4b, O:5a, O:5b, and O:7 (9, 11, 14). For the remaining serotypes, the verified gene order is indicated by an arc joining adjacent genes.

formed distinct groups reflecting the variations in the chemical structure of the known O-antigens (Fig. 1). All serotypes differed from each other by at least one PCR result, except serotypes O:12 and O:4a that were identical to serotypes O:13 and O:8, respectively (14). The recently reported complete DNA sequences of the O-antigen gene clusters of *Y. pseudotuberculosis* O:1a, O:2a, and O:4b (9) are in agreement with the genetic organization previously determined by PCR. Sequences of the O-antigen gene clusters of the serotypes O:2b, O:2c, O:3, O:4a, O:5, and O:7 will be published soon (11).

The aim of the present study was to take advantage of the polymorphism of the O-antigen gene clusters in various *Y. pseudotuberculosis* serotypes to develop PCR assays that could be used for the distinction of *Y. pestis* from *Y. pseudotuberculosis* and most importantly for the genetic typing of the various serotypes (O-genotyping) of *Y. pseudotuberculosis*, as an alternative to the conventional serotyping method.

MATERIALS AND METHODS

Bacterial strains. The characteristics of the 21 *Y. pseudotuberculosis* serotype O:1 to O:15 reference strains used to set up the conditions for the multiplex PCR are listed in Table 1. To evaluate the efficiency and specificity of the PCR-typing system 259 bacterial strains from the culture collection of Pasteur Institut were used (the number of strains of each species is given in parentheses): *Y. pestis* ($n = 16$), *Y. pseudotuberculosis* ($n = 85$), *Y. enterocolitica* ($n = 35$), *Y. kristensenii* ($n = 11$), *Y. intermedia* ($n = 12$), *Y. mollaretii* ($n = 13$), *Y. frederiksenii* ($n = 12$), *Y. bercovieri* ($n = 12$), *Y. aldovae* ($n = 1$), *Y. rohdei* ($n = 1$), *Y. ruckeri* ($n = 1$), *Escherichia* spp. ($n = 9$), *Enterobacter* spp. ($n = 10$), *Erwinia carotovora* ($n = 1$), *Hafnia alvei* ($n = 1$), *Serratia* spp. ($n = 10$), *Levinea malonatica* ($n = 1$), *Klebsiella* spp. ($n = 12$), *Salmonella* spp. ($n = 5$), *Shigella* spp. ($n = 5$), *Citrobacter freundii* ($n = 1$), *Edwardsiella* spp. ($n = 2$), *Proteus vulgaris* ($n = 1$), *Pseudomonas aeruginosa* ($n = 1$), and *Staphylococcus aureus* ($n = 1$).

In addition, *Y. pseudotuberculosis* strains from various collections were analyzed by using the multiplex PCR assay: (i) 76 strains from the Japanese collection, isolated mainly from Japan but also from other countries; (ii) 18 *Y. pseudotuberculosis* strains from the Finnish Collection, isolated in Finland; and (iii) 32 nonagglutinable *Y. pseudotuberculosis* strains from the Institut Pasteur collection, isolated mainly in France.

Extraction of bacterial DNA. The previously described cetyltrimethylammonium bromide (CTAB) technique was used to isolate bacterial DNA (2). Briefly, 1.5 ml of an overnight culture grown in Tryptone soya broth was centrifuged at 16,100 × *g* for 2 min (Eppendorf centrifuge 5415D). The pellet was resuspended in 567 μl of Tris-EDTA (TE) buffer; 30 μl of 10% sodium dodecyl sulfate, and 3 μl of 20-mg/ml proteinase K were added, and the tube was incubated for 1 h at 37°C. Then 100 μl of 5 M NaCl and 80 μl of 10% CTAB–0.7 M NaCl solution was mixed with the sample and incubated at 65°C for 10 min. The sample was extracted first with an equal volume of chloroform-isoamyl alcohol (24:1) solution. After centrifugation at 16,100 × *g* for 5 min, the aqueous phase was recovered and extracted with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) solution. After centrifugation at 16,100 × *g* for 5 min, the recovered aqueous phase was precipitated with isopropanol (0.7 vol), washed with 70% ethanol, briefly dried in lyophilizer (SpeedVac SVC100; Savant), and finally dissolved in 100 μl of TE buffer. The concentration of the DNA was measured with a spectrophotometer (DU 640 spectrophotometer; Beckman) at 260 nm, and DNA was diluted with water to a final concentration of 50 μg/ml. DNA samples were stored frozen at −20°C.

In some PCR assays, template DNA was prepared directly from bacterial colonies by the boiling method. Briefly, one colony was suspended in 100 μl of H_2O, vortexed, and boiled for 10 min. The tube was again vortexed and cooled on ice for a few minutes before being centrifuged at 13,000 × *g* for 5 min at 4°C and around 1 to 5 μl of supernatant was used as template for PCR.

PCRs. Primer sequences and amplification conditions are given in Table 2. All amplifications were performed with the MiniCycler (MJ Research, Inc.) according to the protocols presented in Table 2. For each amplification reaction, a negative control containing water instead of template DNA was included. Five-microliter aliquots of amplified products were separated by electrophoresis in 1.5% agarose gel and visualized by ethidium bromide staining.

For the multiplex PCR, nine sets of primers (Table 2) were designed (program PRIME of the Wisconsin Package version 10.0, Genetics Computer Group [GCG], Madison, Wis.) based on the *Y. pestis* EV76 O-antigen cluster sequence (accession no. AJ251713) and on the *Y. pseudotuberculosis* O:2a *abe* gene sequence (accession no. L01777). The 50-μl multiplex PCR mixture contained 0.4 mM each dATP, dCTP, dGTP, and dTTP (Amersham Pharmacia Biotech, Inc.); 8 μl of 10× DynaZyme II reaction buffer (1.6-fold more than the concentration recommended by the manufacturer, so that the final concentrations in the 50-μl reaction mixture were 16 mM Tris-HCl [pH 8.8], 2.4 mM $MgCl_2$, 80 mM KCl, 0.16% Triton X-100), 2 U of DynaZyme II (Finnzymes), and 100 ng of template DNA.

TABLE 1. Reference strains of *Y. pseudotuberculosis* used for setting up the multiplex PCR assay

Strain code	Serotype	Source	Source (reference)
St.1	O:1a	Human	Knapp (15)
pa3606	O:1b	Human	Fukushima (15)
Kuratani-2	O:1c	Wild rat	Tsubokura (15)
208	O:2a	Pig	Tsubokura (15)
1779	O:2b	Human	Knapp (12)
274	O:2c	Pig	Tsubokura (15)
83	O:3	Human	Mollaret (15)
51	O:4a	Pig	Tsubokura (15)
Pa3422	O:4b	Human	Fukushima (15)
204	O:5a	Human	Tsubokura (15)
197	O:5b	Pig	Tsubokura (15)
DD110	O:6	Dog	Fukushima (15)
257	O:7	Dog	Tsubokura (15)
151	O:8	Pig	Tsubokura (15)
R708	O:9	Wild rat	Fukushima (15)
6088	O:10	Raccoon dog	Inoue (15)
R80	O:11	Wild rat	Fukushima (15)
MW864-2	O:12	River water	Fukushima (15)
N916	O:13[a]	House rat	Zheng (15)
CN3	O:14	Wild rat	Zheng (15)
93422	O:15	Human	Tsubokura

[a] After reserotyping, this strain was found to be of serotype O:12.

Double (*wzz* and *hemH-ddhD*) and single (*wzx-wbyJ* or *wzz*) PCRs were performed in a 50-μl reaction mixture containing the primers (Table 2) and the following reagents: 0.3 mM each dATP, dCTP, dGTP, and dTTP (Amersham); 10 mM Tris-HCl (pH 8.8); 1.5 mM $MgCl_2$; 50 mM KCl; 0.1% Triton X-100; 1 U of DynaZyme II (Finnzymes); and 100 ng of template DNA.

To amplify the ca. 20-kb *hemH-gsk* fragments, the Expand Long Template PCR system (Roche Molecular Biochemicals) and primers yerfb1 and yerfb2 (Table 2) were used. Reactions were prepared on ice in two master mixes that were mixed 1:1 just before the beginning of the PCR. Master mix 1 included the following in addition to the primers: deoxynucleoside triphosphate (dNTP; final concentration, 500 μM), template (200 ng), additional $MgCl_2$ (4 mM), and water to 25 μl. Mastermix 2 contained PCR buffer 3 (10× concentrate), DNA polymerases (2.5 U), and water to 25 μl. The applied long-range PCR protocol is given in Table 2. Restriction digestions with *Eco*RV were performed according to the instructions provided by the supplier.

Nucleotide sequencing and sequence analysis. PCR products were purified with the Quantum Prep PCR Kleen spin columns (catalog no. 732-6300; Bio-Rad Laboratories, Hercules, Calif.). Sequencing reactions were performed with the Thermo Sequenase II dye terminator cycle sequencing kit (catalog no. US80970; Pharmacia Amersham Life Sciences, Cleveland, Ohio) and the ABI PRISM 377 DNA sequencer (Perkin-Elmer). Nucleotide sequence analysis was performed by using the computer programs of the Wisconsin Package.

Nucleotide sequence accession number. The partial *wzx* gene sequences of *Y. pseudotuberculosis* serotypes O:1a, O:1c, O:5a, O:11, O:14, and O:15 have been deposited in GenBank under accession no. AJ539147 to AJ539152. The serotype O:14 sequence over the region *wbyI* to *wbyJ* was deposited under accession no. AJ539153. The partial *wzz* gene sequences for serotypes O:1c, O:2c, O:5b, O:9, O:11, and O:12 were deposited under accession no. AJ539154 to AJ539159.

RESULTS

Testing of a *Y. pseudotuberculosis*-*Y. pestis*-specific PCR (*wzz*-PCR). The *wzz*-PCR (Table 2) was previously found to detect all serotypes of *Y. pseudotuberculosis* in the set of reference

TABLE 2. Sequences and concentrations of the primers and cycling conditions used in PCR assays

PCR	Cycling conditions	Target	Primer	Nucleotide sequence (5′ → 3′)	Amplimer size (bp)	Accession no. on which primer sequence is based	Amount (pmol/PCR)
wzz-PCR specific for *Y. pestis* and *Y. pseudotuberculosis*	1 cycle of initial denaturation at 94°C for 2 min; 40 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 30 s, extension at 72°C for 30 s; 1 cycle of final extension at 72°C for 6 min	*wzz*	Ypf-20120	GGTGATGAGCAAGTTCAAG	418	AJ251713	25
			Ypr-20538	GCTAAATCCACTGCTCGCTG		AJ251713	25
wzz-wbyJ-PCR specific for *Y. pestis*	1 cycle of initial denaturation at 94°C for 2 min; 35 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 90 s; 1 cycle of final extension at 72°C for 6 min	*wzx-wbyJ*	Ypf-8608	GAAATGGAGAAAAAATCCACGG	1,612	AJ251713	25
			Ypr-10219	GTAGTTATAATACGGTTTACTTAACAACTC		AJ251713	25
Multiplex PCR	1 cycle of initial denaturation at 94°C for 2 min; 35 cycles of denaturation at 94°C for 15 s, annealing at 53°C for 30 s, extension at 72°C for 90 s; 1 cycle of final extension at 72°C for 5 min	*gmd-fcl*	Ypf-14159	TCAAGATCGCCATGAGAC	1,370	AJ251713	10
			Ypr-15549	AGGTTCATTCGTTGGTTC		AJ251713	10
		ddhC-prt	Ypf-5270	CGCATAGAAGAGTTTGTTG	1,072	AJ251713	12.5
			Ypr-6342	CTTTCGCCTGAAATTAGAC		AJ251713	12.5
		manB	Ypf-18740	GCGAGCCATAACCCAATAGAC	963	AJ251713	3
			Ypr-19703	GCCACCCATCAAATTCCATAC		AJ251713	3
		abe	Abe1	AGAATAGTTCTGACTGGAGGAAG	775	L01777	7.5
			Abe2	TCAGGAGCCATTACCTCATC		L01777	7.5
		wbyL	Ypf-17770	TTGGAGAAACAAACCTATCTGG	644	AJ251713	7.5
			Ypr-18414	TTTGCATAAAAACGACATAGGC		AJ251713	7.5
		wbyH	Ypf-7170	CGTTATCCCAAAAAAGAGG	528	AJ251713	5
			Ypr-7698	ATGGGAGACGCTTGTGATG		AJ251713	5
		ddhA-B	Ypf-3057	TGTCGCCTAAAGTTATCG	407	AJ251713	12.5
			Ypr-3464	CGAATATCACCGATTTCC		AJ251713	12.5
		wbyK	Ypf-13231	CCGATTACCAGATTTTGAC	307	AJ251713	7.5
			Ypr-13538	CAAAATTCTTATAACCACCACG		AJ251713	7.5
		wzx	Ypf-8576	GAAATTCGCATGTAAAAGCTATTG	105	AJ251713	10
			Ypr-8681	GAACCTAGACTTACCACCCCCAAC		AJ251713	10
Additional double PCRs	1 cycle of initial denaturation at 94°C for 2 min; 35 cycles of denaturation at 94°C for 15 s, annealing at 55°C for 30 s, extension at 72°C for 60 s; 1 cycle of final extension at 72°C for 5 min	*wzz-gsk*	Ypf-20511	GAAAAATACAGCGAGCAG	742	AJ251713	15
			Yerfb2	GAYTTGCGYTTACCAGGAAATTTCATTG		Z47767, U13685, U46859	15

hemH-ddhD		Ypf-913	CAATCCAATGAAGAGTCAG		181	AJ251713		15
		Ypr-1094	CCCTATGACATAAAAACCC			AJ251713		15
Long-range PCR	1 cycle of initial denaturation at 94°C for 2 min; 10 cycles of denaturation at 92°C for 15 s, annealing at 62°C for 15 s, extension at 68°C for 18 min; 20 cycles of denaturation at 92°C for 15 s, annealing at 62°C for 15 s, extension at 68°C for 18 min + 20 s/cycle; 1 cycle of final extension at 68°C for 7 min	*hemH-gsk*	yerfb1	TGGAAGAAATMAAAGARCAAAATCGAGAG	20,000	Z47767, L01777, U46859		25
			yerfb2	GAYTTGCGYTTACCAGGAAATTTCATTG		Z47767, U13685, U46859		25

strains tested (14). To determine whether this test was generally applicable to all *Y. pseudotuberculosis* strains and to evaluate its specificity, 259 strains comprising 85 *Y. pseudotuberculosis* strains, 16 *Y. pestis* strains (15 of which were wild type *Y. pestis*), 98 other *Yersinia* species, and 60 strains belonging to 14 other species and genera were studied (see Materials and Methods). The *wzz*-PCR detected all *Y. pestis* and *Y. pseudotuberculosis* strains tested but remained negative for all other bacteria, including other *Yersinia* species. The sensitivity of this PCR was ca. 250 fg of DNA, corresponding roughly to 50 bacterial cells (data not shown). The *wzz*-PCR assay may thus be used to differentiate the *Y. pseudotuberculosis-Y. pestis* group from other bacterial species.

Identification of regions in the O-antigen gene cluster useful for the development of a *Y. pestis*-specific PCR *(wzx-wbyIJ*-PCR). It was previously observed that the major differences between the *Y. pestis* and *Y. pseudotuberculosis* O:1b sequences were located in the middle of the *wzx* gene, between nt 7650 and 8030 of the O:1b sequence (14). This region was therefore used as a possible target for the development of a *Y. pestis*-specific PCR. This PCR gave a positive signal for all *Y. pestis* strains analyzed, whatever their biotype was, and a negative signal for all *Y. pseudotuberculosis* O:1b strains tested (data not shown). However, positive PCR results were obtained with various non-O:1b *Y. pseudotuberculosis* serotypes (O:1a, O:1c, O:5a, O:5b, O:11, O:14, and O:15). To understand the reasons for these PCR-positive results, the variable regions of the *wzx* genes were sequenced in these serotypes (except for serotype O:5b, for which no sequence could be obtained). The polymorphic nucleotide sites within the *wzx* regions are presented in Fig. 2. The comparison revealed that the strains fall into three major phylogenetic groups (I, O:1a, *Y. pestis*; II, O:5a, O:14, and O:15; III, O:1b, O:1c, and O:11), with clear evidence of multiple recombination events having taken place between and within the strains of different groups. One clear example is the abrupt change in group II to resemble the group I sequences between positions 7771 and 7774. At about the same position, serotype O:11 appears to diverge from the O:1b and O:1c sequences. Of interest was the almost 100% identity between the O:1a and the EV76 sequences in this region. These sequence similarities indicated that it was not possible to develop a *Y. pestis*-specific PCR based on the *wzx* gene sequence.

Since it was known that *Y. pseudotuberculosis* O:1a is negative with the *wbyI*-PCR (14), this region could be a good candidate for the development of a *Y. pestis*-specific PCR. Therefore, further analysis of the *wbyI* region was undertaken. These analyses revealed that negative PCR results were due to the fact that, although the 5′ part of the *wbyI* gene was present in O:1a, the 3′ end of this gene was missing (between the last 27 to 76 codons of *wbyI*). When other serotypes were tested with this PCR, serotype O:14 strains appeared to give a positive signal. This serotype was also positive with the *wzx* and *wbyJ* PCRs. To determine whether degenerate sequences could be identified and used to design specific primers in this region, the O:14 *wbyI-wbyJ* region was sequenced. This sequence was found to be 97% identical to that of *Y. pestis*. However, just downstream of the *wbyI* gene, there was a 71-bp deletion in the O:14 sequence that shortened the *wbyI* gene by one codon but did not affect the beginning of the *wbyJ* gene.

Altogether, these results indicated that the only possibility to

develop a *Y. pestis*-specific PCR based on the differences in the O-antigen gene cluster was to use a primer pair with a forward primer in the *wzx* gene and the reverse primer in the *wbyI-wbyJ* intergenic region that was deleted in serotype O:14. The primers and amplification conditions chosen to test this PCR assay are described in Table 2. The specificity of the *wzx-wbyIJ*-PCR was tested on the set of 259 strains (see Materials and Methods). All *Y. pestis* strains were positive with the *wzx-wbyIJ*-PCR, while all other strains, including the 85 *Y. pseudotuberculosis* strains tested, remained negative. The sensitivity of this assay was 25 pg of DNA, corresponding to ca. 5,000 bacterial cells (data not shown).

Identification of regions in the O-antigen gene cluster usable for the development of *Y. pseudotuberculosis* serotype-specific PCRs. In order to design a multiplex PCR that could differentiate each serotype of *Y. pseudotuberculosis*, different regions of the O-antigen gene cluster previously found to display some degree of polymorphism among the various serotypes (14) were further analyzed.

Based on earlier results (14), the *wzz-gsk* intergenic PCR was considered as one of the potential targets in a multiplex PCR for O-genotyping. When several *wzz-gsk* primer combinations were screened, six serotypes (O:1c, 2c, O:5b, O:9, O:11, and O:12) of *Y. pseudotuberculosis* that were *wzz-gsk* negative with the previous set of primers (14) were positive with some other primer combinations. To identify optimal sets of primers, most of the *wzz* gene and the *wzz-gsk* intergenic region from these six serotypes were sequenced. The polymorphic nucleotide sites within these regions are presented in Fig. 3. The comparison revealed that the *wzz-gsk* region is quite conserved, showing over 93% identity between all serotypes of *Y. pseudotuberculosis* and of *Y. pestis* studied. There were mainly single-nucleotide changes in the *wzz* gene among all serotypes studied and a variable number of TATTTTAC repeats in the *wzz-gsk* intergenic region in *Y. pseudotuberculosis* serotypes O:1c, O:9, O:11, and O:12, while the other serotypes and *Y. pestis* had only one repeat. The alignment showed that the serotypes formed three distinct groups (I to III, Fig. 3) with some evidence of recombination events. For example, group II may have been formed after recombination between groups III and I between nt 19,940 and 19,961 within the *wzz* gene. In addition, the EV76 sequence appeared to be unique downstream of nt 20,114. Within each group, only a few single-nucleotide changes were observed. Group III was distinct in possessing variable numbers of the repeated intergenic motif. The polymorphism of the *wzz-gsk* intergenic region could thus be used to design primers specific for some serotypes.

Another portion of the O-antigen gene cluster that might be of interest to identify serotype-specific sequences was the *ddh* region. PCRs performed in this region showed that most serotypes were *ddhDABC* positive, but three (O:7, O:9, and O:10) were *ddhDABC* negative, and three (O:5b, O:11, and O:14) were *ddhDA* positive but *ddhBC* negative (Fig. 1). For the three latter serotypes, amplification of the region between the *ddhA* and *wbyH* genes yielded a product of a similar size, about 1 kb shorter than that obtained in serotype O:1b. Restriction digestions of the three fragments gave almost identical patterns (data not shown), and preliminary sequence analysis revealed that they had highly conserved sequences, clearly different from those of the *ddhB* and *ddhC* genes. These genes,

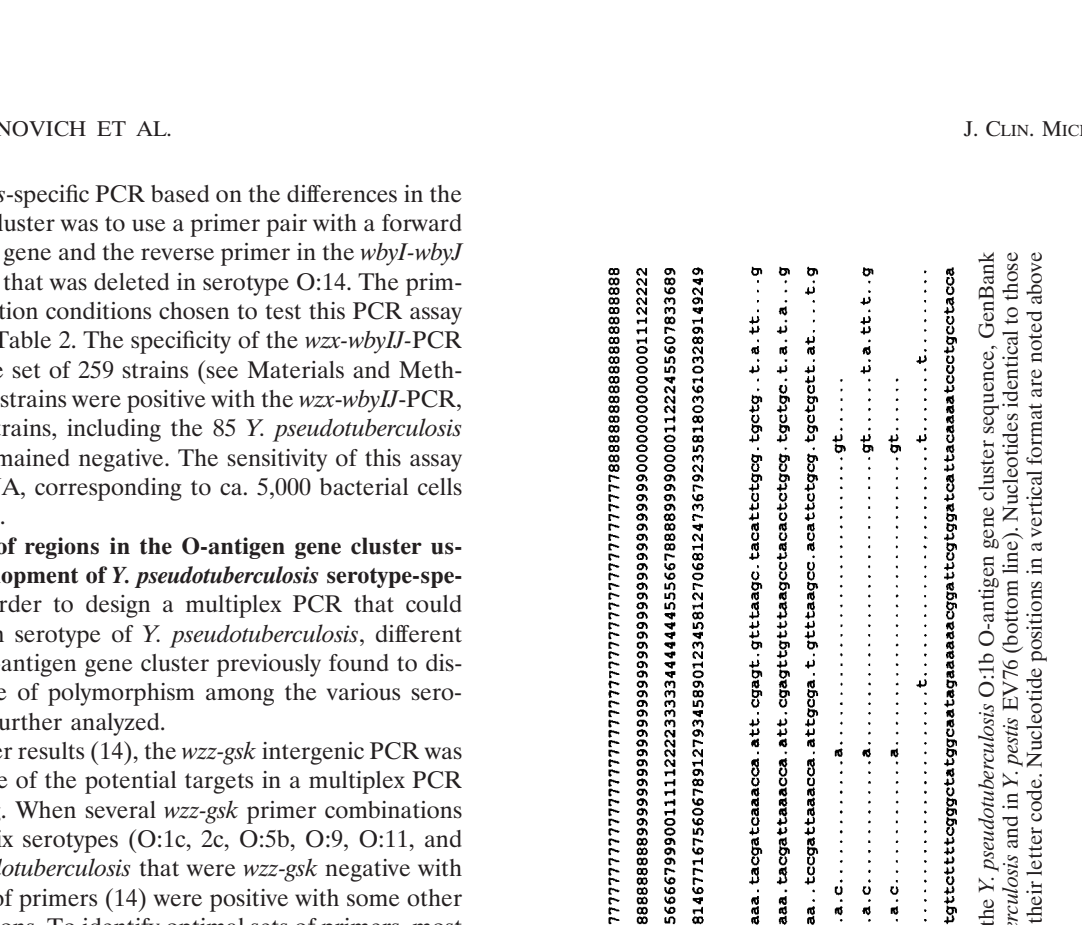

FIG. 2. Polymorphic nucleotide sites in the *wzx* gene between nt 7433 and 8383 (numbering is from the *Y. pseudotuberculosis* O:1b O-antigen gene cluster sequence, GenBank accession no. AJ251712; the *wzx* gene spans positions 7247 to 8590) in various serotypes of *Y. pseudotuberculosis* and in *Y. pestis* EV76 (bottom line). Nucleotides identical to those in the EV76 sequence are indicated by a period, and those differing from this sequence are indicated by their letter code. Nucleotide positions in a vertical format are noted above the polymorphisms.

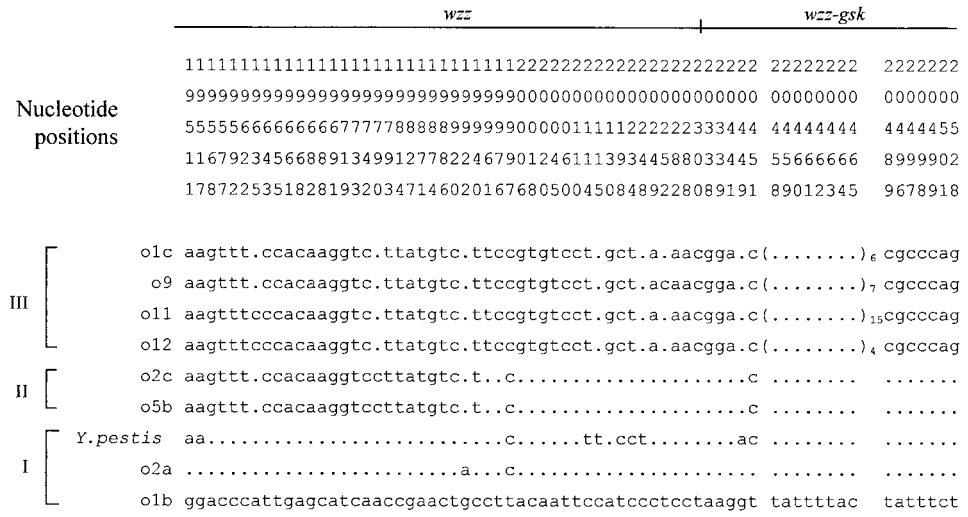

FIG. 3. Polymorphic nucleotide sites in the *wzz* gene (nt 19236 to 20387) and the *wzz-gsk* intergenic region in various serotypes of *Y. pseudotuberculosis* and in *Y. pestis* EV76 when compared to the serotype O:1b sequence (bottom line). Identical nucleotides to the O:1b sequence are indicated by a period and those differing from this sequence by their letter code. Nucleotide positions in vertical format are noted above the polymorphisms.

located between *ddhB* and *wbyH* genes in the O:5b, O:11, and O:14 serotypes, are homologous to NDP-sugar biosynthetic genes. Since the O:5b serotype has in its O-unit a 6-deoxy-L-altrofuranose residue instead of a DDH (7), these two genes may be involved in its biosynthesis, and the same sugar is most likely also present in the O-antigens of serotypes O:11 and O:14. It is noteworthy that the first and the fourth steps of the DDH biosynthetic pathway are retained in these serotypes. Future work is needed to elucidate the biosynthetic steps leading to the formation of NDP-6-deoxy-altrofuranose. This polymorphic *ddh* region could also be used to distinguish different sets of serotypes.

Several *Y. pseudotuberculosis* serotypes (O:1c, O:2b, O:3, O:5a, O:5b, and O:15) were previously found to be *manC* and *manB* positive but *wbyL* negative. In serotype O:1b, the *wbyL* gene is located between *manC* and *manB* (14); thus, this region was another potential target for a PCR assay. When the *manC-manB* region of these serotypes was amplified, the resulting PCR products were of the same size and ca. 1 kb larger than the corresponding PCR product of serotype O:1b. Restriction digestions of the fragments gave similar patterns (data not shown), suggesting that conserved sequences (probably two genes) are present in this region in serotypes O:1c, 2b, O:3, O:5a, O:5b, and O:15. Future work will elucidate the nature of these genes. Nonetheless, these results indicated that the *wbyL* gene or the unknown genes of this region may be used to distinguish different serotype subsets.

Development of multiplex double and single PCRs for O-genotyping of *Y. pseudotuberculosis*. Based on previous data and the results obtained in this study, the genetic organization of the O-antigen gene clusters of the different *Y. pseudotuberculosis* serotypes could be better understood (Fig. 1). This information allowed us to design combinations of primers to be used in a multiplex PCR assay. For this assay, several criteria had to be fulfilled: (i) the multiplex PCR should produce unique PCR product patterns from each reference strain to allow genotypic identification of the majority of the *Y. pseudotuberculosis* serotypes; (ii) the PCR product lengths should sufficiently differ from each other to allow unambiguous identification of the products based on their size; and (iii) the individual PCRs should not be affected by the other PCRs running in the same tube. These criteria were fulfilled when the nine selected sets of primers (Table 2), the optimized concentrations of primers, and a 1.6-fold concentration of the DynaZyme II reaction buffer were used (see Materials and Methods).

Sensitivity of the multiplex PCR assay was tested with the O:1b reference strain DNA (data not shown). Within a range of 5 to 658 ng of DNA per reaction, the multiplex PCR gave reliable results. For DNA concentrations of 1 ng per reaction or lower, the products of genes coding for the larger PCR products, *gmd-fcl* (1,390 bp) and *prt* (1,072 bp), started to disappear (data not shown). The use of bacterial colony boilates as templates in the multiplex PCR was also tested. This method did not give reproducible and satisfactory results. Some bands were not systematically amplified, and nonspecific smears were sometimes observed. Therefore, high-quality DNA appears to be crucial for the optimal performance of the multiplex PCR.

When tested on the reference strains of all *Y. pseudotuberculosis* serotypes, the multiplex PCR patterns (Fig. 4) allowed the clear distinction of 18 O-genotypes (which we designate with the abbreviation "gO:nn") specific for individual serotypes and one pattern that included serotypes O:12 and O:13 (however, see below). Three serotypes, O:7, O:9, and O:10, were negative with the multiplex PCR and therefore required the development of other PCR assays. This was partly achieved with the double *wzz-gsk* and *hemH-ddhD* PCR (Table 2). This double PCR differentiated the O:10 and O:7 serotypes. However, no product was observed for the O:9 serotype (Fig. 5). Since the only *Y. pseudotuberculosis*-specific positive PCR for serotype O:9 is the *wzz*-PCR (described above), a strain neg-

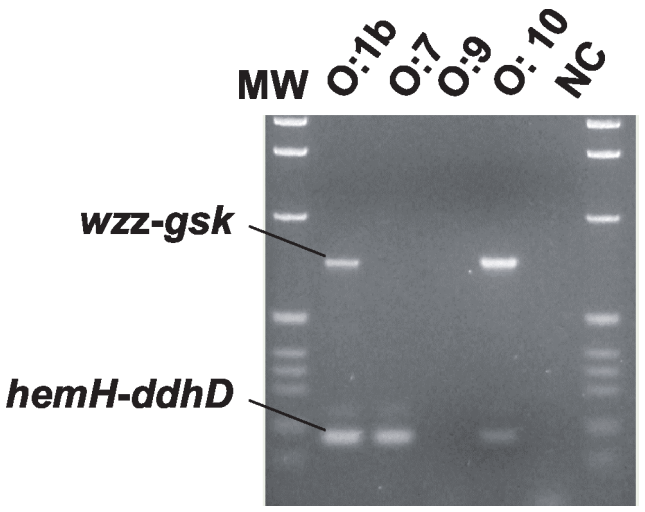

FIG. 4. Double PCR to differentiate *Y. pseudotuberculosis* strains of serotypes O:7, O:9, and O:10. MW, molecular weight standard (1-kb ladder; Gibco BRL); NC, negative control.

ative in multiplex and double PCRs but positive in *wzz*-PCR will be assigned to the genotype gO:9.

In the multiplex PCR, *Y. pestis* pattern differed from that of *Y. pseudotuberculosis* O:1b by being *wzx* positive (data not shown). That is why the *Y. pestis*-specific PCR (*wzx-wbyJ*) should be used for strains giving the gO:1b pattern. It was also noted that in some multiplex PCRs, the *wbyL*- and *wbyH*-amplified fragments were either weak (*wbyH*) or missing (*wbyL*). These, however, did not prevent the proper identification of the O-genotypes.

Application of the O-genotyping method to various *Y. pseudotuberculosis* strains. In order to validate the developed O-genotyping assay, we analyzed clinical, veterinary, and environmental isolates of *Y. pseudotuberculosis* originating from different parts of the world.

Strains from the Japanese collection. The multiplex and double-PCR assays were evaluated on a Japanese collection of *Y. pseudotuberculosis* strains. This included 56 strains of known serotypes (most of the known serotypes were represented by 3 strains each, some of which were already in the reference strain list but were reanalyzed here). O-genotyping and conventional serotyping gave concordant results in 44 of 56 strains. One strain was completely negative in the PCRs, and reidentification showed that it was not *Y. pseudotuberculosis*.

Upon reserotyping, O-serotypes of 3 of the 11 discrepant strains were found indeed to match their O-genotypes, one strain was rough (thus nontypeable) and three still gave discrepant results (four of the strains were not reserotyped). The following discrepancies were noticed. One O:2b strain produced the gO:2a pattern, and one O:4a strain gave the gO:2c pattern. One O:6 strain showed the gO:12 pattern, and two serotype O:10 strains had the gO:9 and the gO:15a patterns. (The latter differed from the gO:15 pattern in only being *wzx* negative.) Interestingly, one serotype O:1c strain also produced the gO:15a pattern, but in conventional serotyping, it was repeatedly designated to the O:1c serotype.

Most discrepancies were seen with serotype O:12 and O:13 strains. Only one of the three serotype O:12 strains had the

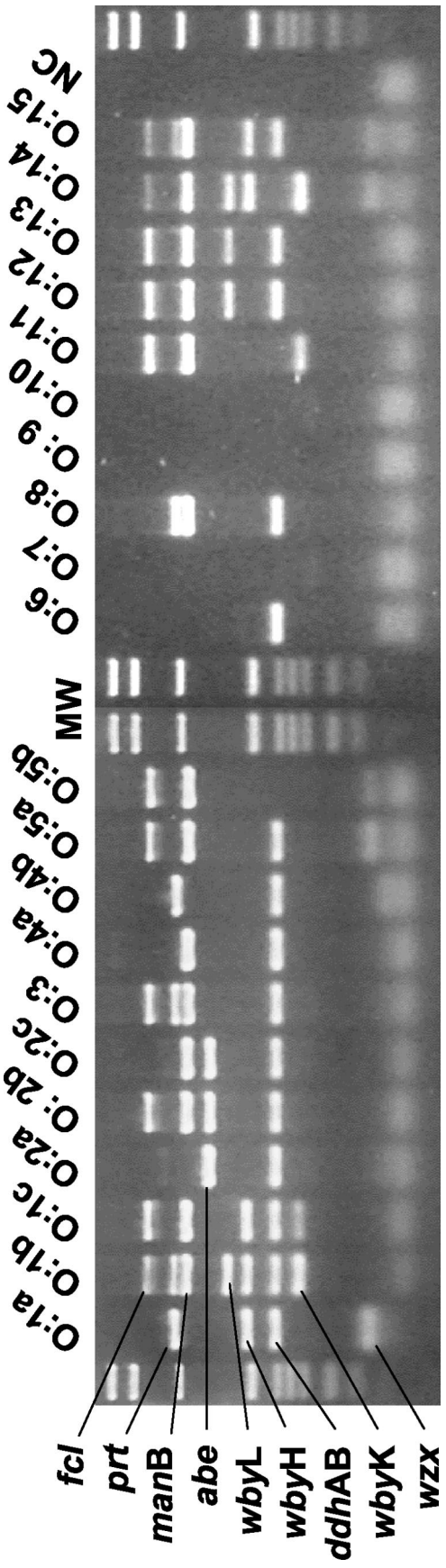

FIG. 5. Multiplex PCR assay on *Y. pseudotuberculosis* strains representing all serotypes. Serotypes are indicated above the lanes. MW, molecular weight standard (1-kb ladder; Gibco BRL) with the 2-kb band at the top of the gel; NC, negative control.

gO:12/13 pattern. Another O:12 strain gave a gO:5a-like PCR-pattern designated as gO:5c, which differed from the gO:5a pattern in being *wzx* negative. However, on reserotyping, the strain was shown to be rough. The third serotype O:12 strain tested produced a completely new pattern, designated gO:12a. Supporting the heterogeneity within serotype O:12, the 20-kb *hemH-gsk* regions, amplified by long-range PCR, produced different *Eco*RV restriction digestion patterns, although they shared several common fragments (data not shown).

None of the serotype O:13 strains fell into the gO:12/13 genotype; instead, they gave gO:1a, gO:1b, and gO:3 genotypes. However, their reserotyping confirmed they were of the O:13 serotype. These results indicate that O-genotyping will need further development for this group of strains. Since *Eco*RV restriction digestion analysis of the amplified ca. 20-kb *hemH-gsk* fragments revealed different restriction fragment patterns (data not shown), a certain genetic heterogeneity probably exists within the O:13 serotype.

Another set of 14 rough strains and 6 nontypeable strains from the Japanese collection were analyzed. An O-genotype could be assigned to 13 of them. Reserotyping of these strains confirmed the concordance between serotyping and O-genotyping in four instances and gave a discrepant result (O:1a versus gO:5c) in one case, and seven strains were still rough and nontypeable. One of the rough strains gave a unique multiplex PCR pattern designated as gO:16. Three of the rough strains and four of the nontypeable strains gave no identifiable O-genotype, and they were subsequently found to be non-*Y. pseudotuberculosis* strains.

O-genotyping of Finnish *Y. pseudotuberculosis* strains. O-genotyping of 18 *Y. pseudotuberculosis* strains isolated from human stool samples in Finland resulted in 1 gO:1a strain, 6 gO:1b strains, and 11 gO:3 strains and indicated a very strong correlation between this method and classical serotyping. We noticed, however, that in many cases, gO:1b strains had been (mistakenly) serotyped as O:1 or O:1a serotypes.

O-genotyping of French nonagglutinable *Y. pseudotuberculosis* strains. Altogether 32 *Y. pseudotuberculosis* strains from the collection of the Institut Pasteur (Paris) that were nonagglutinable with the set of five conventional antisera (O:1 to O:5) were tested with the multiplex PCR. Three O-genotypes were identified. One strain had the gO:1a type and was rough upon reserotyping. Another strain had the gO:1b genotype and was assigned to serotype O:12 upon reserotyping. All of the other strains had the gO:5c genotype. Of the five strains from this group that were reserotyped, one was assigned to serotype O:13 or O:14, and the four others were assigned to serotype O:12. When the *Eco*RV restriction patterns of the amplified 20-kb *hemH-gsk* fragments of these strains were compared, those of the three gO:5c strains were identical, while those of the gO:1a and gO:1b strains were completely different. The serotype O:12/gO:5c strain had an *Eco*RV pattern very similar to that of the O:12 reference strain.

DISCUSSION

Three PCR assays based on the genetic polymorphism of the O-antigen gene cluster were developed and tested in this study. One assay (*wzz*) allowed us to distinguish the *Y. pseudotuberculosis-Y. pestis* group from other species of *Yersinia* and *Enterobacteriaceae*. Another assay (*wzx-wbcIJ*) could readily differentiate *Y. pestis* from other *Yersinia* species and other *Enterobacteriaceae*. The main aim of this study was to design a PCR assay capable of genotyping the different *Y. pseudotuberculosis* serotypes. Here we have presented a multiplex PCR that can identify and O-genotype 18 of the 21 known serotypes right away, including the most commonly circulating ones. With the use of a second, double PCR, two additional serotypes (O:7 and O:10) could be identified. The last serotype (O:9) may be confirmed with a third, single PCR. A very good correlation between classical serotyping and our O-genotyping was observed in most instances; only a few strains of some serotypes were assigned to a different O-genotype. Most discrepancies concerned the two very rare serotypes O:12 and O:13. To the best of our knowledge, only six strains of these serotypes have been isolated in the Far East from water and rodents (6); thus, the probability of encountering them in practical work is low. Furthermore, the O-genotyping method detected a certain degree of heterogeneity within some conventionally defined serogroups that could be used for more accurate classification of the isolated strains. This PCR-based O-genotyping method is applicable in conventional laboratories, without the need for the preparation of a large set of cross-absorbed antisera. It is also useful to correct misidentification of some strains as *Y. pseudotuberculosis* and to type *Y. pseudotuberculosis* isolates that have lost the expression of the O-antigen and thus cannot be serotyped.

ACKNOWLEDGMENTS

This work was supported by the Academy of Finland grant (project no. 50441), the Finnish Scientific Advisory Board for Defense, and the Center for International Mobility (CIMO).

Anne Peippo is thanked for excellent technical help.

REFERENCES

1. **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.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, O. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, New York, N.Y.
3. **Bercovier, H., H. H. Mollaret, J. M. Alonso, J. Brault, G. R. Fanning, A. G. Steigerwalt, and D. J. Brenner.** 1980. Intra- and interspecies relatedness of *Yersinia pestis* by DNA hybridization and its relationship to *Yersinia pseudotuberculosis*. Curr. Microbiol. **4:**225–229.
4. **Bogdanovich, T. M., E. Carniel, H. Fukushima, and M. Skurnik.** 2003. Genetic (sero)typing of *Yersinia pseudotuberculosis*, p. 337–340. *In* M. Skurnik, K. Granfors, and J. A. Bengoechea (ed.), The genus *Yersinia*: entering the functional genomic era. Kluwer Academic/Plenum Publishers, New York, N.Y.
5. **Butler, T.** 1994. *Yersinia* infections: centennial of the discovery of the plague bacillus. Clin. Infect. Dis. **19:**655–661.
6. **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.
7. **Korchagina, N. I., R. P. Gorshkova, and Y. S. Ovodov.** 1982. Studies on O-specific polysaccharide from *Yersinia pseudotuberculosis* VB serovar. Bioorg. Khim. **8:**1666–1669.
8. **Ljungberg, P., M. Valtonen, V. P. Harjola, S. S. Kaukorantatolvanen, and M. Vaara.** 1995. Report of four cases of *Yersinia pseudotuberculosis* septicemia and a literature review. Eur. J. Clin. Microbiol. Infect. Dis. **14:**804–810.
9. **Pacinelli, E., L. Wang, and P. R. Reeves.** 2002. Relationship of *Yersinia pseudotuberculosis* O antigens IA, IIA, and IVB: the IIA gene cluster was derived from that of IVB. Infect. Immun. **70:**3271–3276.

10. **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.
11. **Reeves, P. R., E. Pacinelli, and L. Wang.** 2003. O antigen gene clusters of *Yersinia pseudotuberculosis*, p. 199–206. *In* M. Skurnik, K. Granfors, and J. A. Bengoechea (ed.), The genus *Yersinia*: entering the functional genomic era. Kluwer Academic/Plenum Publishers, New York, N.Y.
12. **Samuelsson, K., B. Lindberg, and R. R. Brubaker.** 1974. Structure of O-specific side chains of lipopolysaccharides from *Yersinia pseudotuberculosis*. J. Bacteriol. **117:**1010–1016.
13. **Skurnik, M.** 1999. Molecular genetics of *Yersinia* lipopolysaccharide, p. 23–51. *In* J. Goldberg (ed.), Genetics of bacterial polysaccharides. CRC Press, Boca Raton, Fla.
14. **Skurnik, M., A. Peippo, and E. Ervelä.** 2000. Characterization of the O-antigen gene clusters of *Yersinia pseudotuberculosis* and the cryptic O-antigen gene cluster of *Yersinia pestis* shows that the plague bacillus is most closely related to and has evolved from *Y. pseudotuberculosis* serotype O:1b. Mol. Microbiol. **37:**316–330.
15. **Tsubokura, M., and S. Aleksic.** 1995. A simplified antigenic scheme for serotyping of *Yersinia pseudotuberculosis*: phenotypic characterization of reference strains and preparation of O and H factor sera. Contrib. Microbiol. Immunol. **13:**99–105.