# Typing and Clustering of *Yersinia pseudotuberculosis* Isolates by Restriction Fragment Length Polymorphism Analysis Using Insertion Sequences

E. Voskresenskaya,[a,b] C. Savin,[a] A. Leclercq,[a] G. Tseneva,[b] E. Carniel[a]

*Yersinia* Research Unit, National Reference Laboratory, WHO Collaborating Center for *Yersinia*, Institut Pasteur, Paris, France[a]; Laboratory of Bacterial Respiratory Infections, Institut Pasteur, Saint Petersburg, Russia[b]

**Yersinia pseudotuberculosis is an enteropathogen that has an animal reservoir and causes human infections, mostly in temperate and cold countries. Most of the methods previously used to subdivide *Y. pseudotuberculosis* were performed on small numbers of isolates from a specific geographical area. One aim of this study was to evaluate the typing efficiency of restriction fragment length polymorphism of insertion sequence hybridization patterns (IS-RFLP) compared to other typing methods, such as serotyping, ribotyping, and multilocus sequence typing (MLST), on the same set of 80 strains of *Y. pseudotuberculosis* of global origin. We found that IS100 was not adequate for IS-RFLP but that both IS285 and IS1541 efficiently subtyped *Y. pseudotuberculosis*. The discriminatory index (DI) of IS1541-RFLP (0.980) was superior to those of IS285-RFLP (0.939), ribotyping (0.944), MLST (0.861), and serotyping (0.857). The combination of the two IS (2IS-RFLP) further increased the DI to 0.998. Thus, IS-RFLP is a powerful tool for the molecular typing of *Y. pseudotuberculosis* and has the advantage of exhibiting well-resolved banding patterns that allow for a reliable comparison of strains of worldwide origin. The other aim of this study was to assess the clustering power of IS-RFLP. We found that 2IS-RFLP had a remarkable capacity to group strains with similar genotypic and phenotypic markers, thus identifying robust populations within *Y. pseudotuberculosis*. Our study thus demonstrates that 2IS- and even IS1541-RFLP alone might be valuable tools for the molecular typing of global isolates of *Y. pseudotuberculosis* and for the analysis of the population structure of this species.**

*Y*ersinia pseudotuberculosis is an enteropathogenic species belonging to the genus *Yersinia*. This species is transmitted by the fecal-oral route and has a wide range of animal reservoirs (1). In humans, the main clinical manifestations of pseudotuberculosis are abdominal pain (mimicking appendicitis), fever, and sometimes diarrhea (2). Dissemination to deeper tissues and the bloodstream are frequent in older patients with underlying conditions (3). In some specific areas (e.g., Russia and Japan), *Y. pseudotuberculosis* may also cause a specific disease known as Far East scarlet-like fever (4–6), characterized by a strong inflammatory syndrome accompanying intestinal disorders. Although human *Y. pseudotuberculosis* infections are less frequent than those caused by *Yersinia enterocolitica* worldwide, human pseudotuberculosis outbreaks are reported in various parts of the world, such as Japan (7, 8), Russia (9, 10), France (3, 11), and Scandinavia (12–15).

Phenotypically, *Y. pseudotuberculosis* is subdivided into 15 serotypes and several subserotypes (16). The most common serotypes are O:1 to O:5, while the others are restricted to geographical areas in Asia. Because most strains isolated from human cases are of the major serotypes O:1 and O:3, the discriminatory power of serotyping is limited. However, since *Y. pseudotuberculosis* exhibits a certain degree of genetic polymorphisms (17), molecular techniques represent valuable alternatives for subtyping this species. One of the most commonly used molecular typing methods is the analysis of the genomic restriction fragment length polymorphism (RFLP) obtained by pulsed-field gel electrophoresis (PFGE) (18–21). This technique is valuable for investigating an outbreak to find a common source of contamination. Typing of *Y. pseudotuberculosis* by multilocus enzyme electrophoresis (22–24), RFLP of the pYV virulence plasmid (25), or multilocus variable-number tandem-repeat analysis (MLVA) (26) have also been de-

scribed. These techniques have been applied to a small number of *Y. pseudotuberculosis* strains or to isolates from a given geographical area during outbreaks. Recently, larger sets of *Y. pseudotuberculosis* strains of global origin have been subjected to other typing methods. Ribotyping of 80 *Y. pseudotuberculosis* strains isolated from various countries and different hosts demonstrated that this technique allowed the subdivision of the strains into 27 ribotypes (27). However, the small number of hybridizing bands and the poor resolution of some profiles led to the conclusion that this technique has some intrinsic limitations (27). The analysis of multilocus sequence typing (MLST) was also recently used to study the diversity of *Y. pseudotuberculosis*. In one work, 79 strains from various countries were analyzed (28), of which 11 were subsequently found to be *Yersinia similis* and 68 were true *Y. pseudotuberculosis*. Among these 68 *Y. pseudotuberculosis* isolates, 54 sequence types (STs) were identified. In another work involving MLST, 417 strains belonging to the *Y. pseudotuberculosis* complex and isolated from all continents were typed (29). The 386 true *Y. pseudotuberculosis* strains were divided into 76 STs. The performance of IS-RFLP for typing *Y. pseudotuberculosis* strains has also been investigated in two studies. In one of them, IS1541-RFLP

typing of 20 *Y. pseudotuberculosis* isolates indicated that its discriminatory power was similar to that of PFGE (21). In another study, IS*100*- and IS*285*-RFLP typing of 27 isolates showed that IS-RFLP distinguished 24 IS types (30). Therefore, different discriminatory powers were observed for different methods and different sets of strains.

An analysis of the genetic polymorphisms of a species may also be used to study subpopulations within this species and sometimes to identify epidemiological links between geographical origins or sources and specific genetic clusters. MLST analysis of a large set of *Y. pseudotuberculosis* strains of global origins did not identify strong patterns of specificity for geographical origin, host type, or serotype (29). IS-RFLP typing using two or three different insertion sequences (IS*100*, IS*285*, and IS*1541*) has been successfully applied to the molecular analysis of a collection of diverse strains of *Yersinia pestis* (31) and recently to the epidemiological study of a plague outbreak (32). The 3IS-RFLP dendrogram demonstrated a clear clustering of *Y. pestis* strains, not only according to their biovars but also according to their geographical origin. Whether IS-RFLP can efficiently identify genetic subgroups of *Y. pseudotuberculosis* has never been determined.

The aims of this study were to evaluate the discriminatory power of IS-RFLP compared to those of serotyping, ribotyping, and MLST on the same set of strains of global origin and to estimate the potential for IS-RFLP to delineate populations within the species *Y. pseudotuberculosis*.

## MATERIALS AND METHODS

**Bacterial strains and growth conditions.** The 80 *Y. pseudotuberculosis* strains used in this study were taken from the strain collection of the *Yersinia* Research Unit and the National Reference Laboratory at the Institut Pasteur. Their main characteristics are described in Table 1. The serotypes of the strains were determined by genoserotyping according to Bogdanovich et al. (33), with one modification: the O:8 genoserotype was redefined as O:4c since the strains agglutinated with the anti-O:4 antiserum and their PCR profile were a combination of the O:4a and O:4b banding patterns. The ribotypes were determined in a previous study (27). The STs were already established for 66 strains (29). We determined the STs of the remaining strains, except for strain IP32921, which had lost its viability. Bacterial suspensions were prepared from stock cultures kept at −80°C and streaked onto Luria-Bertani agar plates or grown in peptone broth with shaking for 24 h to 48 h at 28°C.

**DNA extraction, restriction, and transfer to nylon membranes.** The extraction of genomic DNA from each *Y. pseudotuberculosis* strain was performed as described previously (34). Five micrograms of each sample was digested overnight at 37°C with EcoRI or EcoRV before being loaded onto 0.8% agarose gels and subjected to electrophoresis (50 V in 1× Tris-borate-EDTA buffer) for 24 h. The DNA of *Y. pestis* strain IP304 was systematically loaded on each gel to serve in intergel normalization. DNA bands were stained with ethidium bromide. Alkaline denaturation, neutralization, and transfer of total DNA onto nylon filters (Hybond-N$^+$; Amersham, England) with a VacuGene apparatus (Pharmacia LKB Biotechnology, Uppsala, Sweden) were performed as previously described (35).

**Preparation of the IS probes and hybridization.** The primer pairs used to amplify a portion of the insertions sequences were IS*100*-F (5′-AAAACGTTCGAAGAGTATGA-3′) and IS*100*-R (5′-GATGAGCAGGCGGGGGGCCA-3′) (255 bp), IS*1541*-F (5′-AAAGCTTTCAGCTTTGGGTC-3′) and IS*1541*-R (5′-TCTTTCCCTTCAGGTACCCC-3′) (319 bp), and IS*285*-F (5′-AGCTTACCGAACACCTCGGG-3′) and IS*285*-R (5′-GTTGATGCCCAGCGCTAGGA-3′) (406 bp) (31). PCR amplification reactions were performed on the genomic DNA of *Y. pestis* strain CO92 as a template, as previously described (31). The IS probes were peroxidase

labeled with the enhanced chemiluminescence (ECL) direct nucleic acid labeling and detection system (Amersham). Hybridization was performed overnight at 42°C.

**Bioinformatic analysis of the IS-RFLP patterns.** The hybridization patterns obtained with each IS were scanned, and the computerized data were analyzed using the BioNumerics software version 6.6 (Applied Maths, Kortrijk, Belgium). Bands automatically assigned by the computer were checked by eyes and corrected manually when necessary. A position tolerance of 1.8 was selected for each IS. Cluster analysis of the individual or combined IS-RFLP patterns was done by the unweighted-pair group method using average linkages (UPGMA), using the Dice coefficient to analyze the similarities of the banding patterns. The discriminatory power of each IS-RFLP was determined by calculating the discrimination index (DI) based on the Simpson's index of diversity (36). The DI depends on the number of types defined by the test method and the relative frequencies of these types, using the equation DI $= 1 - \{[1/(N \times [N-1])] \times (\Sigma nj \, [nj-1])\}$, where $N$ is the total number of unrelated isolates and $nj$ is the number of strains that belong to the $j$th type. An algorithm minimal spanning tree (MST) based on serotype, ribotype, ST, and 2IS type of each strain was constructed using BioNumerics version 6.6.

## RESULTS

**Setup of the IS-RFLP conditions.** The 80 *Y. pseudotuberculosis* strains used to evaluate the IS-RFLP method were chosen based on their variety of geographical origins, sources, serotypes (O:1 to O:6), and on the fact that their ribotypes (27) and STs (29) were already determined for most of them, thus allowing for a comparison of the three molecular typing methods on the same set of diverse strains.

The IS*100* element was previously reported to be either missing or in low copy numbers (≤5) in various *Y. pseudotuberculosis* strains (30, 37–39). Using a set of 11 strains of various serotypes, no IS*100* hybridizing fragments were detected in seven of them, and only one or two fragments were detected in three other strains (data not shown). Therefore, because of the low discriminatory power of IS*100*-RFLP, we decided not to use this IS further for IS-RFLP typing of *Y. pseudotuberculosis*. In contrast, IS*1541* and IS*285* were reported to be systematically present and in several copies in the genome of *Y. pseudotuberculosis* isolates previously tested (4, 21, 30, 40). We therefore decided to estimate the performance of a typing method for *Y. pseudotuberculosis* based on the polymorphisms of individual and combined IS*1541* and IS*285* fingerprints.

To optimize the resolution of the banding patterns, the genomic DNA of a few strains of *Y. pseudotuberculosis* was digested with different restriction enzymes and hybridized with the IS*1541* and IS*285* probes. The hybridization profiles that gave the best resolution were obtained after EcoRI digestion for IS*285* and after EcoRV digestion for IS*1541*. An example of the banding patterns is provided in Fig. 1. These two enzymes were therefore selected for further use.

To ensure that the hybridization profiles were correct, we also compared the sizes and numbers of fragments expected from the genome analysis of the sequenced strain IP32953 with those visualized after digestion of the DNA of this strain and IS*1541* or IS*285* hybridization. Bands with the expected sizes were obtained. We also noted the absence of IS*1541* or IS*285* copies on the IP32953 pYV virulence plasmid, indicating that the loss of this replicon would not change the hybridization profile of the strain.

**IS*285*-RFLP patterns.** Upon hybridization with the IS*285* probe of the DNA of the 80 *Y. pseudotuberculosis* strains, the number of bands detected varied from 1 to 11 (Fig. 2), with 60% of the

TABLE 1 Characteristics of the 80 strains of *Y. pseudotuberculosis* used for IS-RFLP typing

| Strain | Serotype | Country of isolation | Isolation source | yr of isolation | Ribotype | ST | IS*285* pattern | IS*1541* pattern | 2IS type |
|--------|----------|---------------------|------------------|-----------------|----------|-----|-----------------|------------------|----------|
| IP32953 | O:1b | France | Human | 1990 | 2 | 42 | 26 | 1 | 1 |
| IP32790 | O:1a | Italy | Pig | 1986 | 4 | 42 | 40 | 4 | 2 |
| IP32954 | O:1a | France | Human | 1990 | 9 | 9 | 18 | 18 | 3 |
| IP32637 | O:1b | France | Unknown | 1983 | 11 | 43 | 17 | 26 | 4 |
| IP32950 | O:1b | France | Human | 1990 | 19 | 14 | 24 | 6 | 5 |
| IP33109 | O:1b | France | Human | 2000 | 19 | 14 | 24 | 6 | 5 |
| IP32949 | O:1b | France | Human | 1990 | 11 | 43 | 3 | 27 | 6 |
| IP32777 | O:1b | France | Human | 1986 | 11 | 43 | 3 | 27 | 6 |
| IP30284 | O:1a | Italy | Pigeon | 1961 | 6 | 42 | 33 | 4 | 7 |
| IP30437 | O:1b | Canada | Beaver | 1961 | 8 | 42 | 4 | 61 | 8 |
| IP30642 | O:1a | Tunisia | Mouse | 1963 | 9 | 9 | 20 | 56 | 9 |
| IP31878 | O:1a | Tunisia | Rodent | 1970 | 8 | 42 | 19 | 4 | 10 |
| IP32533 | O:1b | New Zealand | Deer | 1978 | 11 | 43 | 12 | 60 | 11 |
| IP32665 | O:1a | Yugoslavia | Hare | 1983 | 8 | 42 | 6 | 4 | 12 |
| IP32939 | O:1a | Romania | Soil | 1989 | 12 | 42 | 23 | 4 | 13 |
| IP33005 | O:1a | Germany | Monkey | 1993 | 6 | 42 | 34 | 4 | 14 |
| IP33038 | O:1b | Australia | Marsupial | 1995 | 11 | 43 | 43 | 23 | 15 |
| IP33291 | O:1a | France | Hare | 2001 | 8 | 42 | 13 | 4 | 16 |
| IP32709 | O:1b | England | Bird | 1984 | 8 | 10 | 16 | 17 | 17 |
| IP32670 | O:1b | England | Pig | 1983 | 11 | 43 | 27 | 24 | 18 |
| IP33285 | O:1b | France | Human | 2001 | 11 | 43 | 28 | 3 | 19 |
| IP31524 | O:1a | Czechoslovakia | Human | 1968 | 9 | 9 | 20 | 22 | 20 |
| IP32651 | O:1b | Yugoslavia | Hare | 1983 | 6 | 42 | 29 | 21 | 21 |
| IP32745 | O:1a | Italy | Human | 1985 | 4 | 42 | 37 | 4 | 22 |
| IP32590 | O:1a | Switzerland | Human | 1981 | 4 | 42 | 15 | 4 | 23 |
| IP32614 | O:1a | Yugoslavia | Hare | 1982 | 21 | 42 | 9 | 4 | 24 |
| IP32879 | O:1b | Switzerland | Bird | 1988 | 1 | 14 | 1 | 25 | 25 |
| IP32524 | O:1b | Holland | Human | 1978 | 11 | 43 | 14 | 26 | 26 |
| IP32929 | O:2b | France | Hare | 1989 | 17 | 16 | 30 | 46 | 27 |
| IP32934 | O:2b | France | Monkey | 1989 | 17 | 16 | 30 | 44 | 28 |
| IP32951 | O:2a | France | Human | 1990 | 24 | 14 | 1 | 31 | 29 |
| IP32921 | O:2b | France | Hare | 1989 | 20 | ND[a] | 30 | 45 | 30 |
| IP32881 | O:2b | Switzerland | Monkey | 1988 | 20 | 16 | 30 | 45 | 30 |
| IP30215 | O:2b | Denmark | Guinea pig | 1960 | 16 | 16 | 7 | 20 | 31 |
| IP30911 | O:2b | Holland | Hare | 1967 | 13 | 16 | 8 | 19 | 32 |
| IP32323 | O:2[b] | Norway | Water | 1973 | 8 | 42 | 21 | 4 | 33 |
| IP32581 | O:2a | Belgium | Human | 1981 | 14 | 14 | 1 | 2 | 34 |
| IP32721 | O:2a | Italy | Hare | 1984 | 10 | 14 | 1 | 10 | 35 |
| IP33293 | O:2a | France | Human | 2001 | 14 | 14 | 1 | 48 | 36 |
| IP32584 | O:2a | Spain | Pig | 1981 | 14 | 14 | 1 | 49 | 37 |
| IP32585 | O:2a | France | Antelope | 1981 | 14 | 14 | 1 | 50 | 38 |
| IP33012 | O:2a | Germany | Monkey | 1994 | 14 | 14 | 1 | 33 | 39 |
| IP33023 | O:2a | Switzerland | Monkey | 1994 | 14 | 14 | 1 | 32 | 40 |
| IP33054 | O:2a | Spain | Human | 1993 | 14 | 14 | 1 | 34 | 41 |
| IP33098 | O:2b | France | Hare | 2000 | 17 | 16 | 10 | 44 | 42 |
| IP33088 | O:2a | France | Human | 2000 | 15 | 14 | 1 | 43 | 43 |
| IP32589 | O:2a | New Zealand | Human | 1981 | 14 | 14 | 1 | 47 | 44 |
| IP32887 | O:3 | Argentina | Cattle | 1988 | 18 | 19 | 22 | 5 | 45 |
| IP32564 | O:3 | Belgium | Human | 1980 | 18 | 19 | 22 | 5 | 45 |
| IP32666 | O:3 | Spain | Human | 1983 | 25 | 14 | 35 | 8 | 46 |
| IP32544 | O:3 | South Africa | Pig | 1979 | 18 | 19 | 22 | 55 | 47 |
| IP31829 | O:3 | England | Ovine fetus | 1969 | 22 | 50 | 22 | 41 | 48 |
| IP32938 | O:3 | Argentina | Cattle | 1989 | 18 | 19 | 22 | 7 | 49 |
| IP32889 | O:3 | Spain | Unknown | 1988 | 25 | 14 | 2 | 35 | 50 |
| IP32802 | O:3 | Italy | Pig | 1986 | 18 | 19 | 22 | 28 | 51 |
| IP32984 | O:3 | Spain | Human | 1992 | 25 | 14 | 25 | 36 | 52 |
| IP32992 | O:3 | Australia | Cattle | 1992 | 18 | 19 | 22 | 37 | 53 |
| IP33097 | O:3 | Argentina | Deer | 1999 | 18 | 19 | 22 | 38 | 54 |
| IP33051 | O:3 | France | Goat | 1996 | 18 | 19 | 22 | 38 | 54 |
| IP33105 | O:3 | Argentina | Cattle | 2000 | 18 | 19 | 22 | 39 | 55 |

(Continued on following page)

TABLE 1 (Continued)

| Strain | Serotype | Country of isolation | Isolation source | yr of isolation | Ribotype | ST | IS*285* pattern | IS*1541* pattern | 2IS type |
|--------|----------|----------------------|------------------|-----------------|----------|-----|-----------------|------------------|----------|
| IP33108 | O:3 | Bulgaria | Human | 1999 | 18 | 19 | 22 | 40 | 56 |
| IP31411 | O:4c | Denmark | Hare | 1961 | 7 | 48 | 36 | 9 | 57 |
| IP31833 | O:4c | England | Sheep | 1969 | 3 | 48 | 11 | 29 | 58 |
| IP30151 | O:4a | Sweden | Otter | 1960 | 23 | 23 | 39 | 42 | 59 |
| IP31830 | O:4c | England | Human | 1969 | 25 | 14 | 42 | 54 | 60 |
| IP32687 | O:4c | France | Wild species | 1983 | 26 | 14 | 5 | 11 | 61 |
| IP32817 | O:5b | Japan | Hare | 1986 | 19 | 73 | 38 | 52 | 62 |
| IP32821 | O:5a | France | Human | 1986 | 17 | 16 | 30 | 14 | 63 |
| IP32816 | O:5b | Japan | Hare | 1986 | 19 | 7 | 32 | 51 | 64 |
| IP32952 | O:5a | France | Human | 1990 | 17 | 16 | 10 | 30 | 65 |
| IP32463 | O:5a | Switzerland | Guinea pig | 1977 | 17 | 16 | 30 | 13 | 66 |
| IP33061 | O:5a | Germany | Monkey | 1997 | 17 | 16 | 30 | 12 | 67 |
| IP32699 | O:5a | France | Wild species | 1983 | 17 | 16 | 30 | 53 | 68 |
| IP31553 | O:6 | Japan | Guinea pig | 1969 | 5 | 27 | 31 | 16 | 69 |
| IP31554 | O:6 | Japan | Guinea pig | 1969 | 5 | 27 | 31 | 15 | 70 |
| IP33156 | O:1b | Russia | Human | 2000 | 27 | 2 | 1 | 58 | 71 |
| IP33157 | O:1b | Russia | Human | 2000 | 27 | 2 | 1 | 59 | 72 |
| IP33158 | O:1b | Russia | Human | 2000 | 27 | 2 | 1 | 59 | 72 |
| IP33162 | O:1b | Ukraine | Human | 1999 | 9 | 42 | 41 | 57 | 73 |
| IP33161 | O:1b | Ukraine | Rodent | 1999 | 9 | 42 | 41 | 57 | 73 |

[a] ND, not determined because the strain was no longer viable.
[b] O:2: this strain agglutinated with the O:2 antiserum but had an unclassifiable genoserotype, different from O:2.
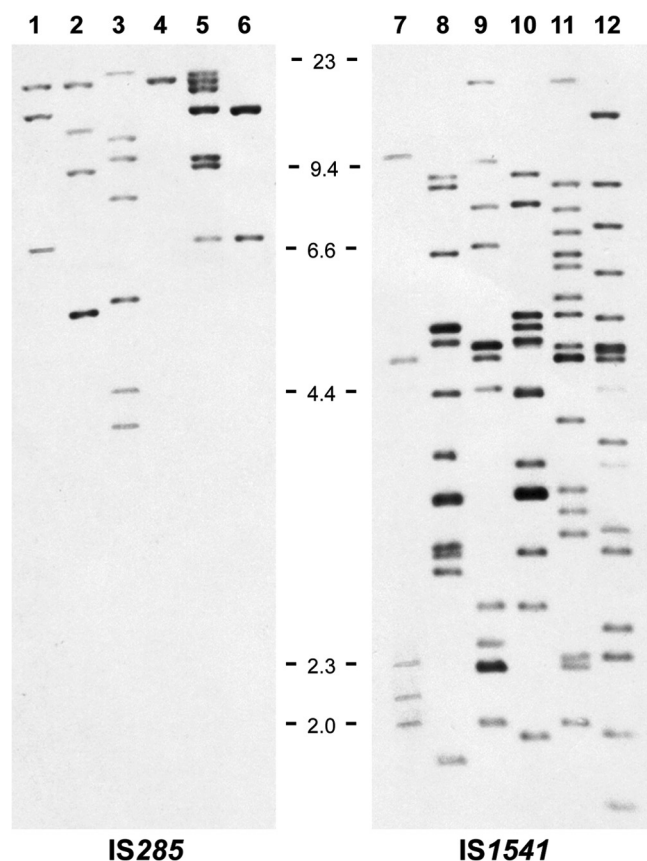


**FIG 1** Examples of IS-RFLP profiles obtained after EcoRI or EcoRV digestion of the genomic DNA of various *Y. pseudotuberculosis* strains and hybridization with the IS*285* and IS*1541* probes, respectively. Lane 1, IP32879; lane 2, IP30911; lane 3, IP32614; lane 4, IP32938; lane 5, IP32984; lane 6, IP31554; lane 7, IP32953; lane 8, IP32581; lane 9, IP31411; lane 10, IP32721; lane 11, IP33051; lane 12, IP32817. The tick marks between the two panels indicate the size of the molecular mass markers (lambda DNA-HindIII digest) in kb.

strains displaying one to four copies of IS*285*. No hybridizing fragments common to all patterns were observed, but a band of ≈20.5 kb was found in 64 strains (80%), and another one of ≈7 kb was found in 41 strains (51%) (Fig. 2).

The 80 strains were separated into 43 distinct IS*285* profiles, while serotyping subdivided them into 10 groups, ribotyping into 27 groups, and MLST into 14 groups (Table 1). The dominant IS*285* profiles were: IS*285* profile/pattern 1 (285#1) (15 strains), 285#22 (11 strains), and 285#30 (8 strains). Together they represented almost half of the strains (42.5%). Six profiles (285#3, 285#10, 285#20, 285#24, 285#31, and 285#41) were shared by two strains, and 34 profiles were unique. The discrimination index (DI) for IS*285*-RFLP was 0.939, compared to 0.857 for serotyping, 0.944 for ribotyping, and 0.861 for MLST. Therefore, IS*285*-RFLP had a discriminatory power higher than those of serotyping and MLST but lower than that of ribotyping.

We then examined whether a correlation could be observed between some epidemiological, phenotypic, or genotypic characteristics of the strains and their IS*285* types. The IS*285* profiles were not associated with specific geographical origins or sources, but it is worth noting that all four strains from Argentina belonged to 285#22, and all three strains from Russia belonged to 285#1 (Table 1). Furthermore, the majority of the 285#22 strains (9/11) were isolated from animal sources.

No strict association between serotypes and IS*285* profiles was observed; nevertheless, the major IS*285* profiles were associated with a limited number of serotypes: 285#1 with serotypes O:1b and O:2a, 285#22 with serotype O:3, and 285#30 with serotypes O:2b and O:5a (Table 1). Similarly, some ribotypes correlated with specific IS*285* profiles: all eight ribotype 14 (R14) and all three R27 strains were 285#1, the eight R17 strains were 285#10 or 285#30, and all 10 R18 strains were 285#22. Associations between the ST and IS*285* profiles were also noted in some instances: all 15 285#1 strains were from ST14 or ST2, 10 of the 11 285#22 strains

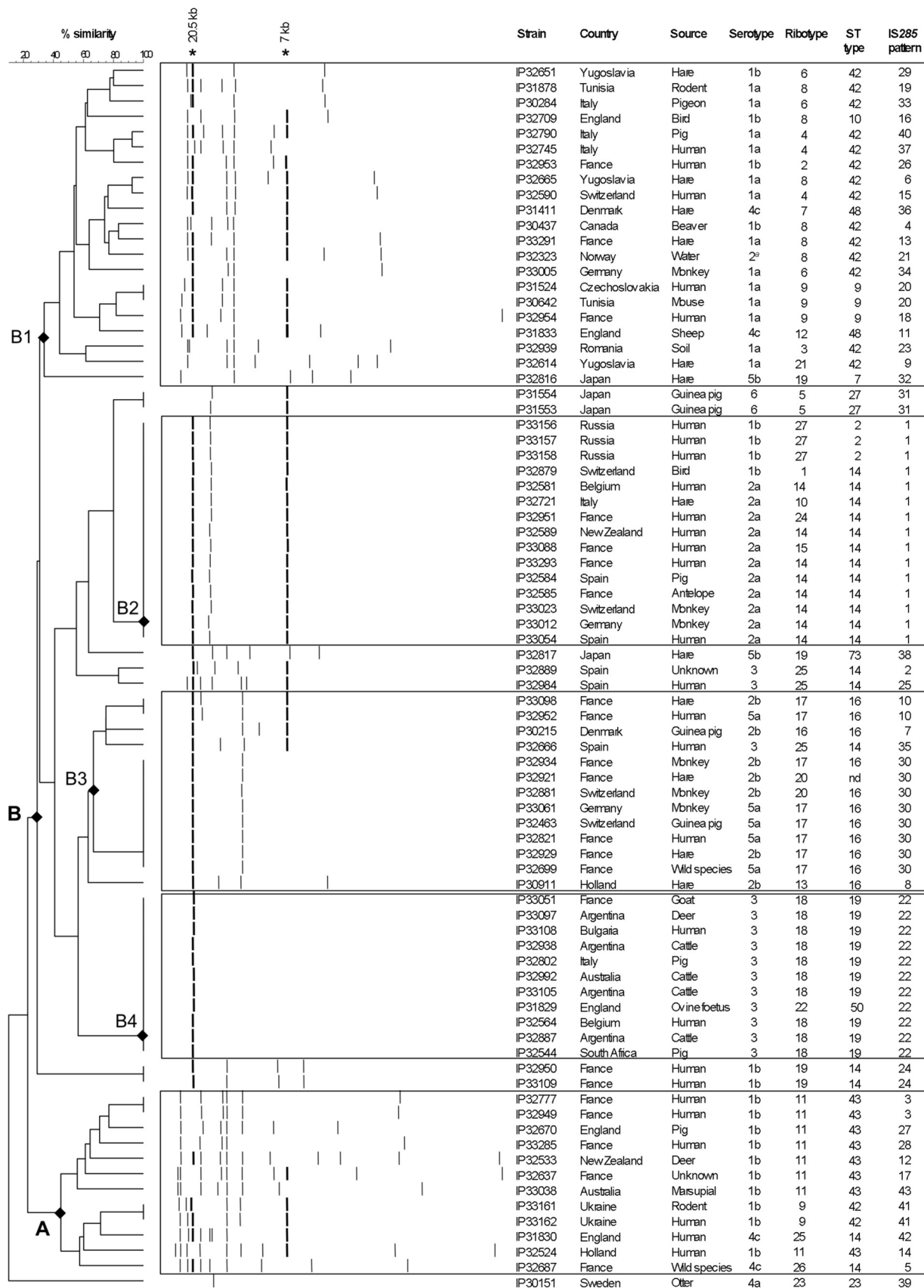| Strain | Country | Source | Serotype | Ribotype | ST type | IS285 pattern |
|---|---|---|---|---|---|---|
| IP32651 | Yugoslavia | Hare | 1b | 6 | 42 | 29 |
| IP31878 | Tunisia | Rodent | 1a | 8 | 42 | 19 |
| IP30284 | Italy | Pigeon | 1a | 6 | 42 | 33 |
| IP32709 | England | Bird | 1b | 8 | 10 | 16 |
| IP32790 | Italy | Pig | 1a | 4 | 42 | 40 |
| IP32745 | Italy | Human | 1a | 4 | 42 | 37 |
| IP32953 | France | Human | 1b | 2 | 42 | 26 |
| IP32665 | Yugoslavia | Hare | 1a | 8 | 42 | 6 |
| IP32590 | Switzerland | Human | 1a | 4 | 42 | 15 |
| IP31411 | Denmark | Hare | 4c | 7 | 48 | 36 |
| IP30437 | Canada | Beaver | 1b | 8 | 42 | 4 |
| IP33291 | France | Hare | 1a | 8 | 42 | 13 |
| IP32323 | Norway | Water | 2a | 8 | 42 | 21 |
| IP33005 | Germany | Monkey | 1a | 6 | 42 | 34 |
| IP31524 | Czechoslovakia | Human | 1a | 9 | 9 | 20 |
| IP30642 | Tunisia | Mouse | 1a | 9 | 9 | 20 |
| IP32954 | France | Human | 1a | 9 | 9 | 18 |
| IP31833 | England | Sheep | 4c | 12 | 48 | 11 |
| IP32939 | Romania | Soil | 1a | 3 | 42 | 23 |
| IP32614 | Yugoslavia | Hare | 1a | 21 | 42 | 9 |
| IP32816 | Japan | Hare | 5b | 19 | 7 | 32 |
| IP31554 | Japan | Guinea pig | 6 | 5 | 27 | 31 |
| IP31553 | Japan | Guinea pig | 6 | 5 | 27 | 31 |
| IP33156 | Russia | Human | 1b | 27 | 2 | 1 |
| IP33157 | Russia | Human | 1b | 27 | 2 | 1 |
| IP33158 | Russia | Human | 1b | 27 | 2 | 1 |
| IP32879 | Switzerland | Bird | 1b | 1 | 14 | 1 |
| IP32581 | Belgium | Human | 2a | 14 | 14 | 1 |
| IP32721 | Italy | Hare | 2a | 10 | 14 | 1 |
| IP32951 | France | Human | 2a | 24 | 14 | 1 |
| IP32589 | New Zealand | Human | 2a | 14 | 14 | 1 |
| IP33088 | France | Human | 2a | 15 | 14 | 1 |
| IP33293 | France | Human | 2a | 14 | 14 | 1 |
| IP32584 | Spain | Pig | 2a | 14 | 14 | 1 |
| IP32585 | France | Antelope | 2a | 14 | 14 | 1 |
| IP33023 | Switzerland | Monkey | 2a | 14 | 14 | 1 |
| IP33012 | Germany | Monkey | 2a | 14 | 14 | 1 |
| IP33054 | Spain | Human | 2a | 14 | 14 | 1 |
| IP32817 | Japan | Hare | 5b | 19 | 73 | 38 |
| IP32889 | Spain | Unknown | 3 | 25 | 14 | 2 |
| IP32984 | Spain | Human | 3 | 25 | 14 | 25 |
| IP33098 | France | Hare | 2b | 17 | 16 | 10 |
| IP32952 | France | Human | 5a | 17 | 16 | 10 |
| IP30215 | Denmark | Guinea pig | 2b | 16 | 16 | 7 |
| IP32666 | Spain | Human | 3 | 25 | 14 | 35 |
| IP32934 | France | Monkey | 2b | 17 | 16 | 30 |
| IP32921 | France | Hare | 2b | 20 | nd | 30 |
| IP32881 | Switzerland | Monkey | 2b | 20 | 16 | 30 |
| IP33061 | Germany | Monkey | 5a | 17 | 16 | 30 |
| IP32463 | Switzerland | Guinea pig | 5a | 17 | 16 | 30 |
| IP32821 | France | Human | 5a | 17 | 16 | 30 |
| IP32929 | France | Hare | 2b | 17 | 16 | 30 |
| IP32699 | France | Wild species | 5a | 17 | 16 | 30 |
| IP30911 | Holland | Hare | 2b | 13 | 16 | 8 |
| IP33051 | France | Goat | 3 | 18 | 19 | 22 |
| IP33097 | Argentina | Deer | 3 | 18 | 19 | 22 |
| IP33108 | Bulgaria | Human | 3 | 18 | 19 | 22 |
| IP32938 | Argentina | Cattle | 3 | 18 | 19 | 22 |
| IP32802 | Italy | Pig | 3 | 18 | 19 | 22 |
| IP32992 | Australia | Cattle | 3 | 18 | 19 | 22 |
| IP33105 | Argentina | Cattle | 3 | 18 | 19 | 22 |
| IP31829 | England | Ovine foetus | 3 | 22 | 50 | 22 |
| IP32564 | Belgium | Human | 3 | 18 | 19 | 22 |
| IP32887 | Argentina | Cattle | 3 | 18 | 19 | 22 |
| IP32544 | South Africa | Pig | 3 | 18 | 19 | 22 |
| IP32950 | France | Human | 1b | 19 | 14 | 24 |
| IP33109 | France | Human | 1b | 19 | 14 | 24 |
| IP32777 | France | Human | 1b | 11 | 43 | 3 |
| IP32949 | France | Human | 1b | 11 | 43 | 3 |
| IP32670 | England | Pig | 1b | 11 | 43 | 27 |
| IP33285 | France | Human | 1b | 11 | 43 | 28 |
| IP32533 | New Zealand | Deer | 1b | 11 | 43 | 12 |
| IP32637 | France | Unknown | 1b | 11 | 43 | 17 |
| IP33038 | Australia | Marsupial | 1b | 11 | 43 | 43 |
| IP33161 | Ukraine | Rodent | 1b | 9 | 42 | 41 |
| IP33162 | Ukraine | Human | 1b | 9 | 42 | 41 |
| IP31830 | England | Human | 4c | 25 | 14 | 42 |
| IP32524 | Holland | Human | 1b | 11 | 43 | 14 |
| IP32687 | France | Wild species | 4c | 26 | 14 | 5 |
| IP30151 | Sweden | Otter | 4a | 23 | 23 | 39 |

**FIG 2** Dendrogram generated from the IS285-RFLP patterns of the 80 *Y. pseudotuberculosis* strains studied, using the UPGMA clustering analysis with the BioNumerics software. A position tolerance of 1.8% was chosen. Stars indicate hybridizing fragments of ca. 20.5 kb and 7 kb that were present in the majority of strains. Rectangles indicate clusters of strains of interest.

were from ST19, and seven 285#30 strains were from ST16 (Table 1). Therefore, a certain degree of association between some genotypes and IS*285* profiles was observed.

To further delineate the IS*285* genetic groups and their association with other phenotypic and genetic traits, a clustering analysis of the IS*285* profiles was performed. The resulting UPGMA dendrogram displayed two main branches (A and B) and one more distant branch corresponding to a single strain (IP30151, Fig. 2). The strains in branch A exhibited multiple IS*285*-hybridizing fragments (Fig. 2). Most of these strains had a unique IS*285* profile (10 among 12 strains), but they were all of serotype O:1b or O:4c, and they also belonged to a small number of MLST types (ST14, ST42, and ST43) and ribotypes (R9, R11, R25, and R26). Within branch B, at least four subclusters (B1 to B4) were delineated (Fig. 2). Subbranch B1, like branch A, contained strains with high copy numbers of IS*285* and with mostly unique IS*285* profiles. There was, however, a strong association with ST42 (14/21 strains), and to a lesser extent with R9, R8, and R4 (12/21 strains). Most of them were of serotype O:1a or O:1b (17/21). Subbranches B2, B3, and B4 had lower IS*285* copy numbers (Fig. 2). All strains in B2 had the same 285#1 profile, which was associated with serotypes O:1b and O:2a, ST14 (12/15 strains), and to a lesser extent with R14 (8/15 strains). B3 comprised strains exhibiting five different IS*285* profiles with a predominance of 285#30 (8/13 strains). Most strains in this cluster were of serotypes O:2b and O:5a (12/13 strains), ST16 (11/13 strains), and R17 (8/13 strains). B4 was composed of strains having the 285#22 profile, which was characterized by a single hybridizing fragment corresponding to the most conserved ≈20.5-kb band (Fig. 2). All strains in this cluster were of serotype O:3, and all but one were ST19 and R18. The clustering of *Y. pseudotuberculosis* strains based on their IS*285* profiles thus showed a good, although not perfect, association with other phenotypic and genotypic traits.

**IS*1541*-RFLP patterns.** The polymorphisms of the IS*1541* profiles were greater than that of IS*285*, with 61 different RFLP patterns among 80 *Y. pseudotuberculosis* strains (Table 1). This higher number of profiles might be linked to an overall higher copy number of IS*1541*, which varied from one to 33 copies per strain (Fig. 3). The most frequent IS*1541* copy number was two (13 strains), followed by seven (10 strains). No band common to all profiles was observed; however, a ≈8.5-kb hybridizing fragment was present in 40 strains (50%), and a ≈4.7-kb band was present in 44 strains (55%) (Fig. 3).

The most frequent IS*1541* profile was 1541#4, which was found in 11 strains (Table 1). The strains with this pattern were not those that had the most common IS*285* profiles. Of note, all but one strain with the 1541#4 pattern were of serotype O:1a, and all were of ST42. However, they had five different ribotypes, indicating that in that case, ribotyping was more discriminatory than IS*1541*-RFLP. Except for 1541#4, no more than two strains shared a similar IS*1541* profile, and 51 of them had a unique IS*1541* profile. The discrimination index for IS*1541*-RFLP was 0.98 and therefore higher than those of all other compared methods (serotyping, MLST, ribotyping, and IS*285*-RFLP).

Because of this high degree of diversity, it was hardly possible to establish a link between individual IS*1541* profiles and specific countries of origin, serotypes, ribotypes, or STs. However, this relationship was examined in the UPGMA dendrogram generated by the clustering analysis. Similarly to IS*285*, the IS*1541* dendrograms displayed two main branches (A and B) and an additional

branch that, like in the IS*285* dendrogram, corresponded to strain IP30151 (Fig. 3). Branch A was characterized by a low number of IS*1541* copies and branch B by intermediate to high numbers of IS*1541* copies (Fig. 3). Branch A contained all strains with the 1541#4 pattern (10/11 of serotype O:1a) plus four other strains with distinct patterns, but all from serotype O:1b. There was a strong association with ST42 (14/15 strains) and to a lesser extent with R8, R6, and R4 (12/15 strains). Branch B was divided into six subbranches (B1 to B6) in which a robust association with specific serotypes, STs, and ribotypes was observed (Fig. 3). All strains in B1 were from O:2a and ST14, and most of them (4/6 strains) were R14. All 10 strains in B2 were O:3, ST19, and R18. This subbranch contained a large number of strains (6/10) of non-European origin (from Argentina, Australia, and South Africa) that were isolated from animals (8/10 strains). Subbranch B3 was slightly more heterogeneous, with two serotypes (O:3 and O:4c), but all strains were ST14, and most of them (4/5) were R25. Most strains (5/6) in B4 were O:2a, all were ST14, and a majority (4/6) were R14. All 8 strains in B5 were O:1b, ST43, and R11. Finally, subbranch B6 was composed of strains of serotype O:2b or O:5a that all belonged to ST16 and most of which (8/12) belonged to R17. This subbranch was composed of European strains only, and most of them (10/12) had an animal source.

Additional strains that were not included in the above clusters were nevertheless grouped according to their epidemiological and genotypic characteristics (Fig. 3). Group 1 (G1) contained two strains isolated from hares in Japan that were both serotype O:5b and R19. G2 contained three strains that were O:1a, R9, and ST9. In G3, the two strains were from O:4c and ST48. Finally, within groups G4 to G7, the strains had identical serotypes, ribotypes, and STs, as well as the same geographical origin and sometimes the same source (guinea pigs in Japan, humans in France, and humans in Russia). Therefore, almost all strains that were grouped by IS*1541*-RFLP shared several common phenotypic and genotypic traits and sometimes even geographical origins or sources.

**2IS-RFLP typing.** We then wondered whether combining the IS*1541* and IS*285* profiles would increase the discriminatory power and clustering potential of IS-RFLP. This combination yielded 73 different patterns (designated 2IS types), with no more than two strains sharing the same 2IS type (Table 1). The discriminatory power of 2IS typing (DI = 0.998) was thus higher than that of IS*1541*-RFLP alone.

Remarkably, the seven pairs of strains that had the same 2IS type also had the same serotype, ribotype, and ST (Table 1), suggestive of a clonal origin of the strains composing each pair. Four of these pairs contained isolates from the same country: 2IS#5 and 2IS#6 from France, 2IS#72 from Russia, and 2IS#73 from Ukraine. This might indicate the local spread of a specific strain. However, strains from France that had the same 2IS type were isolated from distant regions, with a time interval of 4 years (2IS#6) or 10 years (2IS#5), indicating a geographical and temporal spread. Three other pairs were composed of strains isolated from different countries (2IS#30 from France and Switzerland), and even from different continents (2IS#45 from Belgium and Argentina and 2IS#54 from France and Argentina), further demonstrating the geographical spread of *Y. pseudotuberculosis* clones, not only locally but also globally.

The 2IS-RFLP clustering analysis also identified two main branches (A and B) and, not unexpectedly, an outgroup branch corresponding to strain IP30151 (Fig. 4). Five clusters (B1 to B5)
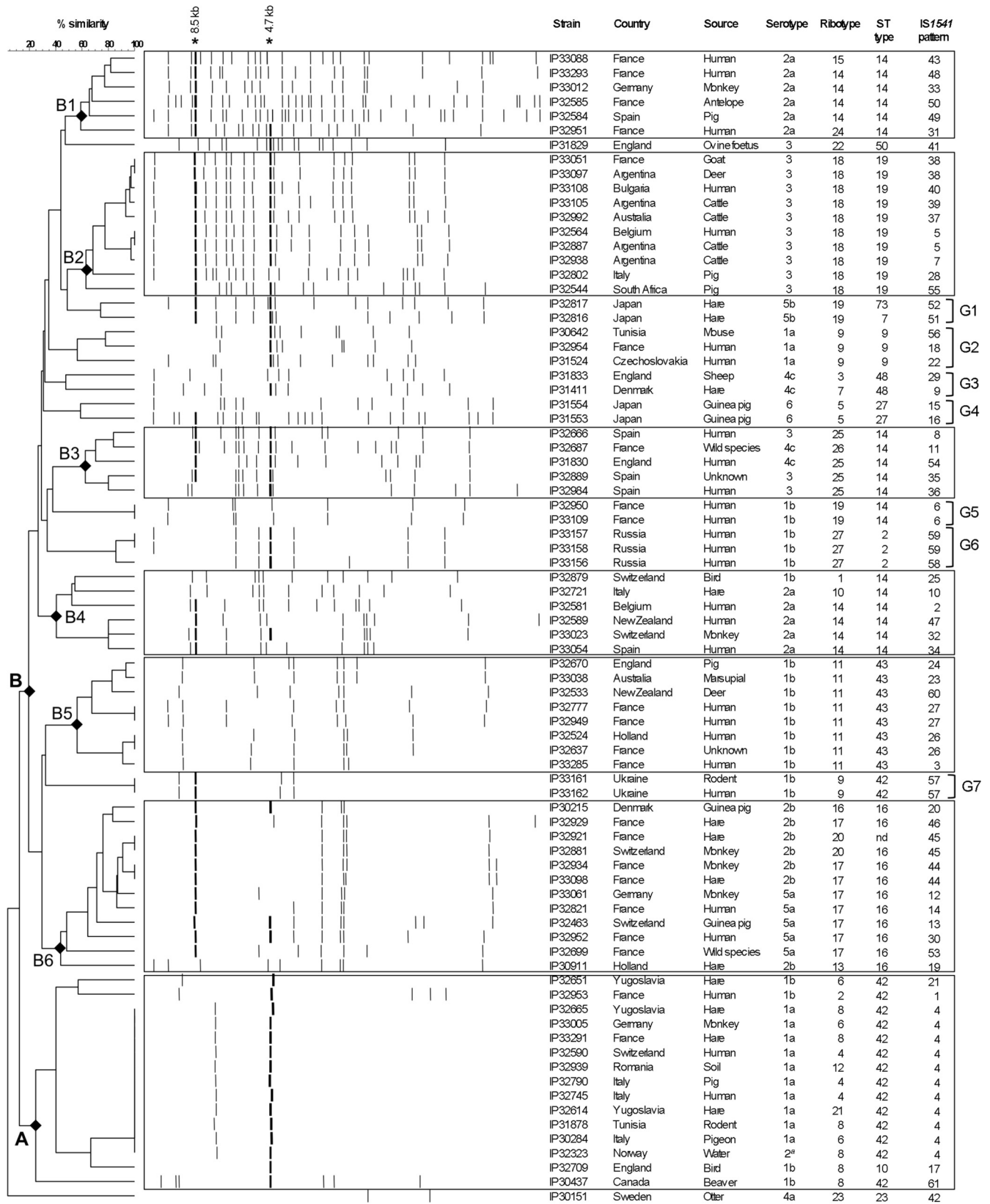
**FIG 3** Dendrogram generated from the IS1541-RFLP patterns of the 80 *Y. pseudotuberculosis* strains studied, using the UPGMA clustering analysis with the BioNumerics software. A position tolerance of 1.8% was chosen. Stars indicate hybridizing fragments of ca. 8.5 kb and 4.7 kb that were present in the majority of strains. Rectangles indicate clusters of strains of interest, and braces indicate groups of two or three strains with common characteristics.

were identified in branch B. Some similarities between the clustering obtained with 2IS- and IS285-RFLP were noted, but they were moderate. In contrast, despite having a different organization, the 2IS and IS1541 dendrograms displayed almost identical clusters. Branch A1 in the 2IS dendrogram corresponded to subbranch B5 in the IS1541 dendrogram. The 2IS subbranch B1 was equivalent to IS1541 branch A, 2IS B2 was equivalent to IS1541 branch B6, 2IS B4 grouped with IS1541 B1 and B4, and 2IS B5 corresponded to IS1541 B3 (Fig. 3 and 4). The only minor difference was 2IS B3, which was very close to IS1541 B2 but contained an additional strain (IP31829) that differed from all other strains in the cluster by a different ST (ST50) and ribotype (R22). Furthermore, six groups of strains outside these clusters that shared identical characteristics in the IS1541 dendrogram (G2 to G7) were also identified in the 2IS dendrogram (Fig. 3 and 4).

The different branches, subbranches, and groups delineated in the 2IS dendrogram were defined by grouping strains that appeared to share some common genotypes or serotypes, but this grouping may have been somewhat arbitrary. To evaluate the solidity of this clustering, we constructed a minimal spanning tree (MST) based on the combination of the characteristics of each strain (serotype, ribotype, ST, and 2IS type). Remarkably, the MST formed clusters (Fig. 5) that corresponded to the subgroups identified in the 2IS dendrogram, demonstrating the robustness of the 2IS-RFLP clustering. These MST populations also corresponded to the IS1541 clusters. However, the IS1541 clusters B1 and B4, which formed distinct and distant branches in the dendrogram (Fig. 3), formed a single branch in the 2IS dendrogram (Fig. 4) and also a single population in the MST (Fig. 5), indicating that 2IS-RFLP is slightly more potent than IS1541-RFLP for defining *Y. pseudotuberculosis* populations. Overall, our data indicate that IS-RFLP might be a useful and valuable tool for studying the population structure of the species *Y. pseudotuberculosis*.

## DISCUSSION

Different methods have been used to subtype the species *Y. pseudotuberculosis*. One of the most classical and simple techniques is serotyping. However, this method has a poor discriminatory power, especially since most of the strains isolated from clinical cases belong to the major serotypes O:1 and O:3. Another commonly used and more modern typing method is PFGE (18–21). This technique has a high discriminatory power and is therefore valuable for investigating the source of an outbreak. However, because of the complexity of its banding patterns, this technique is hardly applicable to the study of global isolates. A more recently developed method that may facilitate interstrain comparison is MLVA. This technique was recently and successfully applied to the typing of *Y. pseudotuberculosis* isolates from Finland (26). However, out of 63 strains studied from serotype O:1, 52 had the same MLVA profile, suggesting that the power of this method to differentiate strains within a given geographical area may be limited. Ribotyping and MLST have also been used to subtype *Y. pseudotuberculosis* (27–29). Both methods exhibited a higher discriminatory power than serotyping and a lower complexity than PFGE, but which of these two methods has the higher capacity to differentiate isolates is unknown. In this study, the comparison of ribotypes and STs on the same set of *Y. pseudotuberculosis* strains showed that ribotyping had a stronger discriminatory power (DI = 0.944) than MLST (DI = 0.861).

Previous works suggested that IS-RFLP might be a valuable

tool for subtyping the species *Y. pseudotuberculosis* (21, 30). We thus decided to evaluate the subtyping efficiencies of individual and combined IS-RFLP on a larger panel of strains of global origin and to compare them with those of ribotyping and MLST. The species *Y. pestis* has previously been successfully subtyped by IS-RFLP using a combination of three IS (IS100, IS285, and IS1541) (31). However, we found that IS100 was frequently in low copy number or absent from the *Y. pseudotuberculosis* genome and therefore that this insertion element was not usable for the identification of subgroups within this species. In contrast, IS285 and IS1541 proved to be systematically present, and often in several copy numbers on the *Y. pseudotuberculosis* chromosome; therefore, they were utilizable for IS-RFLP. IS285-RFLP was slightly less discriminatory (DI = 0.939) than ribotyping (DI = 0.944) despite a higher number of profiles, because almost half of the strains belonged to three main IS285 profiles. IS1541 had a much stronger discriminatory power (DI = 0.98) and yielded 61 different IS1541 patterns among the 80 *Y. pseudotuberculosis* strains studied. The higher subtyping efficiency of IS1541-RFLP over IS285-RFLP is most likely due to the higher copy number of IS1541 in the genome of *Y. pseudotuberculosis*. A combination of the two IS profiles further increased the discriminatory power, allowing the subdivision of the isolates into 73 different types and yielding a discrimination index as high as 0.998. 2IS-RFLP is thus an extremely powerful tool for the molecular typing of *Y. pseudotuberculosis*. It has the advantage over PFGE of exhibiting a better resolution of the banding patterns, thus permitting an easier and more reliable comparison of strains of worldwide origin.

Despite the high number of profiles obtained with IS285- and IS1541-RFLP, strains were nonetheless grouped based on the similarity of their banding patterns. The IS285 and IS1541 dendrograms generated were different, but several branches or subbranches were relatively similar (although not identical) in the two dendrograms, suggesting that the location and copy number of these two IS are quite stable and that they reflect the existence of defined genetic groups within *Y. pseudotuberculosis*.

The clustering efficiency of IS1541-RFLP was more robust than that of IS285-RFLP. This was illustrated, for instance, by the fact that the two strains of serotype O:5 and ribotype R19 isolated from hares in Japan clustered together in the IS1541 dendrogram (group G1), while they were in different subbranches in the IS285 dendrogram. Similarly, the two strains that formed group G3 in the IS1541 dendrogram and were from both O:4c and ST48 were more distantly related (although both in branch A) in the IS285 dendrogram. IS1541-RFLP also allowed a better distinction of genetic clusters, as the three strains from Russia that had identical sources, serotypes, ribotypes, and STs formed a specific cluster (G6) in the IS1541 dendrogram, while they had an IS285 pattern identical to that of strains with completely different epidemiological and genotypic characteristics (branch B2 of the IS285 dendrogram). We then wondered whether combining the two IS would further enhance the clustering performance of IS-RFLP. When the IS1541 and 2IS dendrograms were compared, it appeared that their shapes were different, but they generated almost identical clusters and groups of strains. We noted that IS1541-RFLP was more efficient at grouping the two strains isolated from hares in Japan (G1) that were at distant positions in the 2IS dendrogram, but on the other hand, 2IS-RFLP grouped in cluster B4 strains that had several similar characteristics and that were in two distinct clusters (B1 and B4) in the IS1541 dendrogram. Overall, our data
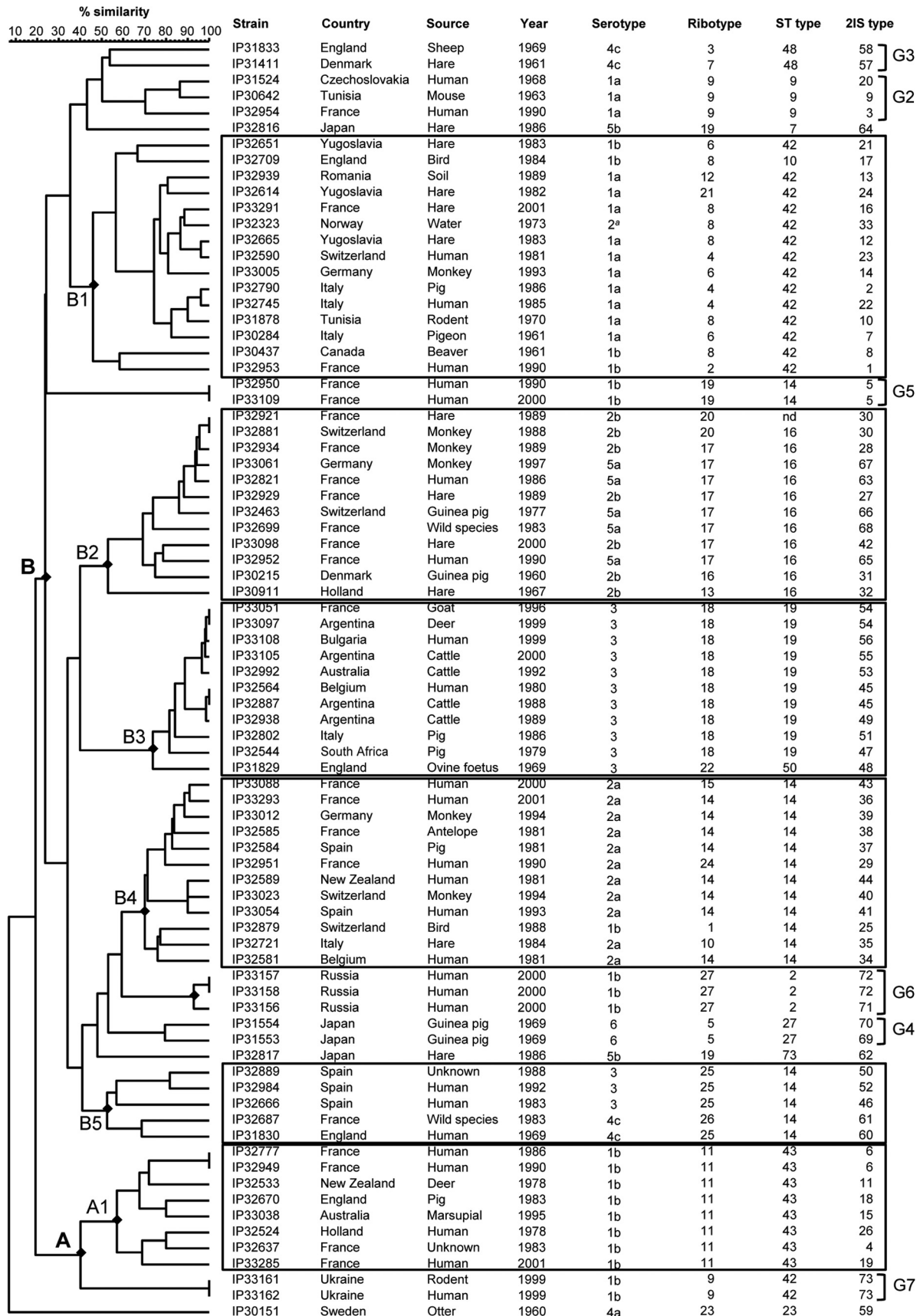
| Strain | Country | Source | Year | Serotype | Ribotype | ST type | 2IS type | |
|--------|---------|--------|------|----------|----------|---------|----------|---|
| IP31833 | England | Sheep | 1969 | 4c | 3 | 48 | 58 | G3 |
| IP31411 | Denmark | Hare | 1961 | 4c | 7 | 48 | 57 | |
| IP31524 | Czechoslovakia | Human | 1968 | 1a | 9 | 9 | 20 | |
| IP30642 | Tunisia | Mouse | 1963 | 1a | 9 | 9 | 9 | G2 |
| IP32954 | France | Human | 1990 | 1a | 9 | 9 | 3 | |
| IP32816 | Japan | Hare | 1986 | 5b | 19 | 7 | 64 | |
| IP32651 | Yugoslavia | Hare | 1983 | 1b | 6 | 42 | 21 | |
| IP32709 | England | Bird | 1984 | 1b | 8 | 10 | 17 | |
| IP32939 | Romania | Soil | 1989 | 1a | 12 | 42 | 13 | |
| IP32614 | Yugoslavia | Hare | 1982 | 1a | 21 | 42 | 24 | |
| IP33291 | France | Hare | 2001 | 1a | 8 | 42 | 16 | |
| IP32323 | Norway | Water | 1973 | 2[a] | 8 | 42 | 33 | |
| IP32665 | Yugoslavia | Hare | 1983 | 1a | 8 | 42 | 12 | |
| IP32590 | Switzerland | Human | 1981 | 1a | 4 | 42 | 23 | |
| IP33005 | Germany | Monkey | 1993 | 1a | 6 | 42 | 14 | |
| IP32790 | Italy | Pig | 1986 | 1a | 4 | 42 | 2 | |
| IP32745 | Italy | Human | 1985 | 1a | 4 | 42 | 22 | |
| IP31878 | Tunisia | Rodent | 1970 | 1a | 8 | 42 | 10 | |
| IP30284 | Italy | Pigeon | 1961 | 1a | 6 | 42 | 7 | |
| IP30437 | Canada | Beaver | 1961 | 1b | 8 | 42 | 8 | |
| IP32953 | France | Human | 1990 | 1b | 2 | 42 | 1 | |
| IP32950 | France | Human | 1990 | 1b | 19 | 14 | 5 | G5 |
| IP33109 | France | Human | 2000 | 1b | 19 | 14 | 5 | |
| IP32921 | France | Hare | 1989 | 2b | 20 | nd | 30 | |
| IP32881 | Switzerland | Monkey | 1988 | 2b | 20 | 16 | 30 | |
| IP32934 | France | Monkey | 1989 | 2b | 17 | 16 | 28 | |
| IP33061 | Germany | Monkey | 1997 | 5a | 17 | 16 | 67 | |
| IP32821 | France | Human | 1986 | 5a | 17 | 16 | 63 | |
| IP32929 | France | Hare | 1989 | 2b | 17 | 16 | 27 | |
| IP32463 | Switzerland | Guinea pig | 1977 | 5a | 17 | 16 | 66 | |
| IP32699 | France | Wild species | 1983 | 5a | 17 | 16 | 68 | |
| IP33098 | France | Hare | 2000 | 2b | 17 | 16 | 42 | |
| IP32952 | France | Human | 1990 | 5a | 17 | 16 | 65 | |
| IP30215 | Denmark | Guinea pig | 1960 | 2b | 16 | 16 | 31 | |
| IP30911 | Holland | Hare | 1967 | 2b | 13 | 16 | 32 | |
| IP33051 | France | Goat | 1996 | 3 | 18 | 19 | 54 | |
| IP33097 | Argentina | Deer | 1999 | 3 | 18 | 19 | 54 | |
| IP33108 | Bulgaria | Human | 1999 | 3 | 18 | 19 | 56 | |
| IP33105 | Argentina | Cattle | 2000 | 3 | 18 | 19 | 55 | |
| IP32992 | Australia | Cattle | 1992 | 3 | 18 | 19 | 53 | |
| IP32564 | Belgium | Human | 1980 | 3 | 18 | 19 | 45 | |
| IP32887 | Argentina | Cattle | 1988 | 3 | 18 | 19 | 45 | |
| IP32938 | Argentina | Cattle | 1989 | 3 | 18 | 19 | 49 | |
| IP32802 | Italy | Pig | 1986 | 3 | 18 | 19 | 51 | |
| IP32544 | South Africa | Pig | 1979 | 3 | 18 | 19 | 47 | |
| IP31829 | England | Ovine foetus | 1969 | 3 | 22 | 50 | 48 | |
| IP33088 | France | Human | 2000 | 2a | 15 | 14 | 43 | |
| IP33293 | France | Human | 2001 | 2a | 14 | 14 | 36 | |
| IP33012 | Germany | Monkey | 1994 | 2a | 14 | 14 | 39 | |
| IP32585 | France | Antelope | 1981 | 2a | 14 | 14 | 38 | |
| IP32584 | Spain | Pig | 1981 | 2a | 14 | 14 | 37 | |
| IP32951 | France | Human | 1990 | 2a | 24 | 14 | 29 | |
| IP32589 | New Zealand | Human | 1981 | 2a | 14 | 14 | 44 | |
| IP33023 | Switzerland | Monkey | 1994 | 2a | 14 | 14 | 40 | |
| IP33054 | Spain | Human | 1993 | 2a | 14 | 14 | 41 | |
| IP32879 | Switzerland | Bird | 1988 | 1b | 1 | 14 | 25 | |
| IP32721 | Italy | Hare | 1984 | 2a | 10 | 14 | 35 | |
| IP32581 | Belgium | Human | 1981 | 2a | 14 | 14 | 34 | |
| IP33157 | Russia | Human | 2000 | 1b | 27 | 2 | 72 | G6 |
| IP33158 | Russia | Human | 2000 | 1b | 27 | 2 | 72 | |
| IP33156 | Russia | Human | 2000 | 1b | 27 | 2 | 71 | |
| IP31554 | Japan | Guinea pig | 1969 | 6 | 5 | 27 | 70 | G4 |
| IP31553 | Japan | Guinea pig | 1969 | 6 | 5 | 27 | 69 | |
| IP32817 | Japan | Hare | 1986 | 5b | 19 | 73 | 62 | |
| IP32889 | Spain | Unknown | 1988 | 3 | 25 | 14 | 50 | |
| IP32984 | Spain | Human | 1992 | 3 | 25 | 14 | 52 | |
| IP32666 | Spain | Human | 1983 | 3 | 25 | 14 | 46 | |
| IP32687 | France | Wild species | 1983 | 4c | 26 | 14 | 61 | |
| IP31830 | England | Human | 1969 | 4c | 25 | 14 | 60 | |
| IP32777 | France | Human | 1986 | 1b | 11 | 43 | 6 | |
| IP32949 | France | Human | 1990 | 1b | 11 | 43 | 6 | |
| IP32533 | New Zealand | Deer | 1978 | 1b | 11 | 43 | 11 | |
| IP32670 | England | Pig | 1983 | 1b | 11 | 43 | 18 | |
| IP33038 | Australia | Marsupial | 1995 | 1b | 11 | 43 | 15 | |
| IP32524 | Holland | Human | 1978 | 1b | 11 | 43 | 26 | |
| IP32637 | France | Unknown | 1983 | 1b | 11 | 43 | 4 | |
| IP33285 | France | Human | 2001 | 1b | 11 | 43 | 19 | |
| IP33161 | Ukraine | Rodent | 1999 | 1b | 9 | 42 | 73 | G7 |
| IP33162 | Ukraine | Human | 1999 | 1b | 9 | 42 | 73 | |
| IP30151 | Sweden | Otter | 1960 | 4a | 23 | 23 | 59 | |

**FIG 4** 2IS-RFLP dendrogram. This dendrogram was generated after combination of the two individual RFLP patterns (IS*285* and IS*1541*). Rectangles indicate clusters of strains of interest, and braces indicate groups of two or three strains with common characteristics.
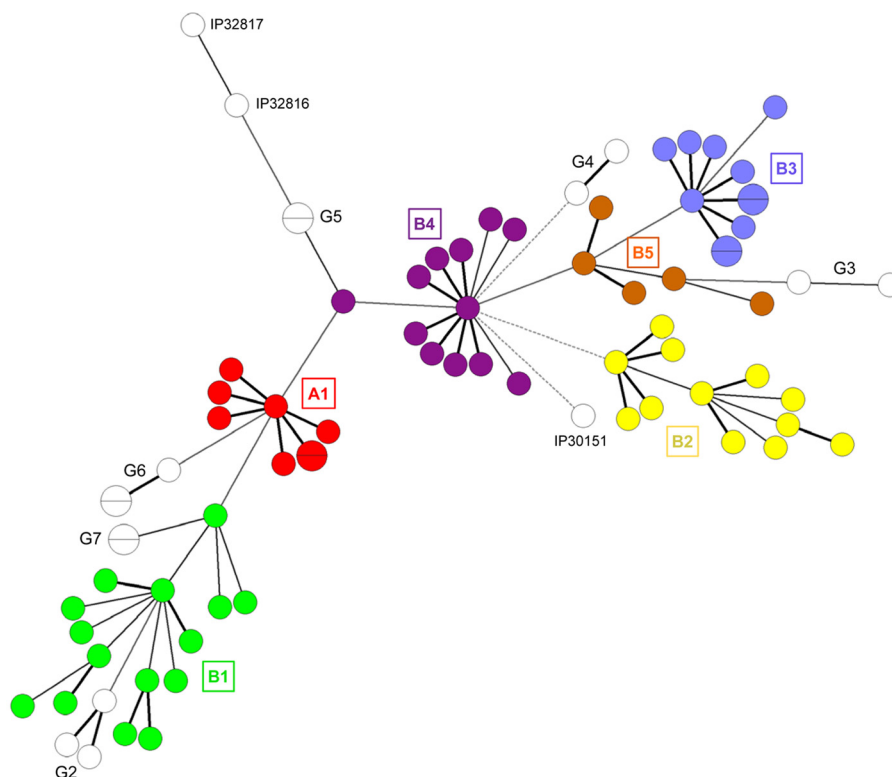
**FIG 5** Minimal spanning tree of the 80 *Y. pseudotuberculosis* strains colored to delineate the different branches (A1, B1 to B5) and groups (G2 to G5) identified in the 2IS-RFLP dendrogram. The tree was constructed using the combination of the serotype, ribotype, ST, and 2IS type for each strain. The thickness of the lines between circles reflects the closeness of the strains, with thicker lines representing a closer relationship.

indicate that IS*1541*-RFLP and 2IS-RFLP have similar efficiencies for grouping genetically related strains.

We observed a remarkable association between ST and 2IS patterns. Except for one strain in cluster B3, all six 2IS clusters and all six 2IS groups contained a single ST. Although less robust, there was also a clear association between serotypes and 2IS clusters, with no more than two different serotypes per cluster of strains. The ribotypes within each 2IS cluster were more diverse but still in limited numbers; this is also indicative of an association between these two genetic traits. The higher discriminatory power of 2IS-RFLP over MLST or ribotyping reflects a higher plasticity of the patterns, most likely due to IS-mediated chromosomal recombination. It is of interest to analyze the 2IS-RFLP patterns of *Y. pseudotuberculosis* strains isolated during one or repeated outbreaks in the same area to evaluate the rate of 2IS-type modification under natural conditions. Although we did not have such isolates in this study, it is still possible that in several instances, strains isolated the same year from the same sources in a specific country clustered together, as exemplified by those forming G4, G6, and G7. In *Y. pestis*, which contains a very high IS copy number (>100), IS-RFLP still robustly clustered strains according to their phylogenetic group and geographical origin (31). Since the IS copy number is 10 times lower in the chromosome of *Y. pseudotuberculosis*, the rate of IS recombination is expected to be lower than in *Y. pestis*, and therefore, the IS patterns are expected to be even more stable.

Our previous MLST study suggested a lack of clear population structure in *Y. pseudotuberculosis* due to mutations and intraspe-

cific recombination events leading to diffuse groups of sequence types (29). In contrast, this IS-RFLP analysis demonstrates the existence of clusters of strains that share identical or nearly identical characteristics. The robustness of these clusters was confirmed when a minimal spanning tree that incorporated the genotype and serotype of each strain was constructed. Discrete populations that matched well with the 2IS clusters were identified in the MST. Since the characteristics analyzed correspond to completely different markers (position and number of IS elements in the genome, sequences of different housekeeping genes, restriction sites flanking the rRNA genes, and composition of the O antigen), the fact that they were tightly associated within each cluster strongly argues for the existence of solid genetic groups in the species *Y. pseudotuberculosis*.

One of these clusters (subbranch B3 in the 2IS dendrogram) was also identified as forming a monophyletic group by MLST and being composed of melibiose nonfermenting strains (29). This cluster contains all four strains from Argentina as well as isolates from Europe, Australia, and South Africa. This and the fact that these strains were isolated mostly from farming animals suggest that this clade spread globally via livestock transportation. Strengthening this hypothesis is the previous observation that these isolates do not circulate among wild animals (41). We also noted that some 2IS clusters (B2, B4, and B5) exclusively contained strains from Europe, suggesting a regional spread of these groups of strains. However, since strains of European origin predominated in our samples (65/80), this may be due to a geographical bias. Similarly, some clusters (B1 to B3) contained mostly

strains of animal origin, suggesting a predominant circulation of these strains in their animal reservoirs, but then again, most of the studied strains were of animal origin, with only 28/80 being human strains. Therefore, a much larger number of strains of diverse geographical origins and sources are needed to establish solid epidemiological links with 2IS clusters.

One strain (IP30151) had a unique serotype, ribotype, and ST and formed an outgroup in both the IS*285* and IS*1541* dendrograms, as it carried the lowest number of IS copies in its genome (one copy of IS*285* and two copies of IS*1541*). Likewise, this strain was previously found to form an outgroup in the ribotype dendrogram (27). The MLST analysis also placed it in a specific cluster designated the Korean group, which was distinct from the *Y. pseudotuberculosis* clusters (29). Although this strain has been isolated in Sweden, it was shown to carry several genes found in Russian isolates that cause a specific and severe disease known as Far East scarlet-like fever (4). We initially thought that this strain was a *Y. pseudotuberculosis* isolate; however, recent phenotypic and genetic characterization of Korean group strains indicate that this group forms a new and potentially pathogenic species of *Yersinia* named *Yersinia wautersii* (42). Our results thus suggest that IS-RFLP might also have the power to differentiate closely related but distinct species within the *Y. pseudotuberculosis* complex. An analysis of additional strains belonging to other species (*Y. wautersii* and *Y. similis*) within this complex is necessary to confirm this observation.

In conclusion, we show here that 2IS-RFLP is a powerful tool for the molecular typing of *Y. pseudotuberculosis*. The superior resolution of the banding patterns, compared to PFGE or ribotyping, should facilitate intra- and interlaboratory comparisons of isolates. 2IS-RFLP is also a useful means to delineate and study the structures of *Y. pseudotuberculosis* populations. To simplify the method, it is also possible to use IS*1541*-RFLP alone, with a minimal loss in sensitivity and grouping efficiency. Since IS-RFLP allows an efficient and reliable analysis of *Y. pseudotuberculosis* isolates of diverse geographical origins, sources, and times of isolation, the incorporation of a large number of new strains in the existing database might help uncover epidemiological links that were not known until now.

## ACKNOWLEDGMENTS

## REFERENCES

1. **Fukushima H, Gomyoda M, Kaneko S.** 1991. Wild animals as the source of infection with *Yersinia pseudotuberculosis* in Shimane Prefecture, Japan. Contrib. Microbiol. Immunol. **12**:1–4.
2. **Mollaret HH.** 1965. Les formes cliniques de l'infection humaine à bacille de Malassez et Vignal. Pathol. Biol. **13**:554–556. (In French.)
3. **Vincent P, Leclercq A, Martin L, *Yersinia* Surveillance Network, Duez JM, Simonet M, Carniel E.** 2008. Sudden onset of pseudotuberculosis in humans, France, 2004–05. Emerg. Infect. Dis. **14**:1119–1122. http://dx.doi.org/10.3201/eid1407.071339.
4. **Eppinger M, Rosovitz MJ, Fricke WF, Rasko DA, Kokorina G, Fayolle C, Lindler LE, Carniel E, Ravel J.** 2007. The complete genome sequence of *Yersinia pseudotuberculosis* IP31758, the causative agent of Far East scarlet-like fever. PLoS Genet. **3**:e142. http://dx.doi.org/10.1371/journal.pgen.0030142.
5. **Mollaret HH, Carniel E, Guilvout I.** 1990. La fièvre scarlatiniforme d'Extrême Orient. Med. Mal. Infec. **20**:519–529. http://dx.doi.org/10.1016/S0399-077X(05)80007-1.
6. **Sato K, Ouchi K, Taki M.** 1983. *Yersinia pseudotuberculosis* infection in children, resembling Izumi fever and Kawasaki syndrome. Pediatr. Infect. Dis. **2**:123–126. http://dx.doi.org/10.1097/00006454-198303000-00011.
7. **Toyokawa Y, Ohtomo Y, Akiyama T, Masuda K, Kasai M, Kaneko S, Maruyama T.** 1993. Large scale outbreak of *Yersinia pseudotuberculosis* serotype 5a infection at Noheji-machi in Aomori Prefecture. Kansenshogaku Zasshi **67**:36–44.
8. **Inoue M, Nakashima H, Ueba O, Ishida T, Date H, Kobashi S, Takagi K, Nishu T, Tsubokura M.** 1984. Community outbreak of *Yersinia pseudotuberculosis*. Microbiol. Immunol. **28**:883–891. http://dx.doi.org/10.1111/j.1348-0421.1984.tb00744.x.
9. **Markov IS, Tkachenko VI, Silin DD.** 1989. Outbreaks of pseudotuberculosis and intestinal yersiniosis among Soviet specialists and members of their families in the Mongolian People's Republic. Zh. Mikrobiol. Epidemiol. Immunobiol. **8**:43–49. (In Russian.)
10. **Tseneva GY, Chesnokova MV, Timofeevich KV, Aleksandrovna VE, Burgasova OA, Sayapina LV, Aleksandrovna TK, Karimova TV.** 2012. Pseudotuberculosis in the Russian federation. Adv. Exp. Med. Biol. **954**:63–68. http://dx.doi.org/10.1007/978-1-4614-3561-7_9.
11. **De Smet P, Van Ussel E.** 1966. Toxi-infection alimentaire collective par *Pasteurella pseudotuberculosis*. Acta Gastroenterol. Belg. **29**:341–350. (In French.)
12. **Jalava K, Hakkinen M, Valkonen M, Nakari UM, Palo T, Hallanvuo S, Ollgren J, Siitonen A, Nuorti JP.** 2006. An outbreak of gastrointestinal illness and erythema nodosum from grated carrots contaminated with *Yersinia pseudotuberculosis*. J. Infect. Dis. **194**:1209–1216. http://dx.doi.org/10.1086/508191.
13. **Jalava K, Hallanvuo S, Nakari UM, Ruutu P, Kela E, Heinäsmäki T, Siitonen A, Nuorti JP.** 2004. Multiple outbreaks of *Yersinia pseudotuberculosis* infections in Finland. J. Clin. Microbiol. **42**:2789–2791. http://dx.doi.org/10.1128/JCM.42.6.2789-2791.2004.
14. **Tertti R, Granfors K, Lehtonen OP, Mertsola J, Mäkelä AL, Välimäki I, Hänninen P, Toivanen A.** 1984. An outbreak of *Yersinia pseudotuberculosis* infection. J. Infect. Dis. **149**:245–250. http://dx.doi.org/10.1093/infdis/149.2.245.
15. **Rimhanen-Finne R, Niskanen T, Hallanvuo S, Makary P, Haukka K, Pajunen S, Siitonen A, Ristolainen R, Pöyry H, Ollgren J, Kuusi M.** 2009. *Yersinia pseudotuberculosis* causing a large outbreak associated with carrots in Finland, 2006. Epidemiol. Infect. **137**:342–347. http://dx.doi.org/10.1017/S0950268807000155.
16. **Tsubokura M, Aleksić S.** 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.
17. **Achtman M, Zurth K, Morelli C, Torrea G, Guiyoule A, Carniel E.** 1999. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. Proc. Natl. Acad. Sci. U. S. A. **96**:14043–14048. http://dx.doi.org/10.1073/pnas.96.24.14043.
18. **Iteman I, Najdenski H, Carniel E.** 1995. High genomic polymorphism in *Yersinia pseudotuberculosis*, p 106–111. *In* Ravagnan G, Chiesa C (ed), Yersiniosis: present and future, vol 13. Karger, Postfach, Basel, Switzerland.
19. **Niskanen T, Fredriksson-Ahomaa M, Korkeala H.** 2002. *Yersinia pseudotuberculosis* with limited genetic diversity is a common finding in tonsils of fattening pigs. J. Food Prot. **65**:540–545.
20. **Nuorti JP, Niskanen T, Hallanvuo S, Mikkola J, Kela E, Hatakka M, Fredriksson-Ahomaa M, Lyytikainen O, Siitonen A, Korkeala H, Ruutu P.** 2004. A widespread outbreak of *Yersinia pseudotuberculosis* O:3 infection from iceberg lettuce. J. Infect. Dis. **189**:766–774. http://dx.doi.org/10.1086/381766.
21. **Odaert M, Berche P, Simonet M.** 1996. Molecular typing of *Yersinia pseudotuberculosis* by using an IS*200*-like element. J. Clin. Microbiol. **34**:2231–2235.
22. **Dolina M, Peduzzi R.** 1993. Population genetics of human, animal, and environmental *Yersinia* strains. Appl. Environ. Microbiol. **59**:442–450.
23. **Goullet P, Picard B.** 1984. Distinctive electrophoretic and isoelectric focusing patterns of esterases from *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. J. Gen. Microbiol. **130**:1471–1480.
24. **Goullet P, Picard B.** 1988. Characterization of *Yersinia enterocolitica*, *Yersinia intermedia*, *Yersinia aldovae*, *Yersinia frederiksenii*, *Yersinia kristensenii* and *Yersinia pseudotuberculosis* by electrophoretic polymorphism of acid phosphatase, esterases, and glutamate and malate dehydrogenases. J. Gen. Microbiol. **134**:317–325.

25. **Fukushima H, Gomyoda M, Kaneko S, Tsubokura M, Takeda N, Hongo T, Shubin FN.** 1994. Restriction endonuclease analysis of virulence plasmids for molecular epidemiology of *Yersinia pseudotuberculosis* infections. J. Clin. Microbiol. **32:**1410–1413.

26. **Halkilahti J, Haukka K, Siitonen A.** 2013. Genotyping of outbreak-associated and sporadic *Yersinia pseudotuberculosis* strains by novel multilocus variable-number tandem repeat analysis (MLVA). J. Microbiol. Methods **95:**245–250. http://dx.doi.org/10.1016/j.mimet.2013.09.007.

27. **Voskressenskaya E, Leclercq A, Tseneva G, Carniel E.** 2005. Evaluation of ribotyping as a tool for molecular typing of *Yersinia pseudotuberculosis* strains of worldwide origin. J. Clin. Microbiol. **43:**6155–6160. http://dx.doi.org/10.1128/JCM.43.12.6155-6160.2005.

28. **Ch'ng SL, Octavia S, Xia Q, Duong A, Tanaka MM, Fukushima H, Lan R.** 2011. Population structure and evolution of pathogenicity of *Yersinia pseudotuberculosis*. Appl. Environ. Microbiol. **77:**768–775. http://dx.doi.org/10.1128/AEM.01993-10.

29. **Laukkanen-Ninios R, Didelot X, Jolley KA, Morelli G, Sangal V, Kristo P, Brehony C, Imori PF, Fukushima H, Siitonen A, Tseneva G, Voskressenskaya E, Falcao JP, Korkeala H, Maiden MC, Mazzoni C, Carniel E, Skurnik M, Achtman M.** 2011. Population structure of the *Yersinia pseudotuberculosis* complex according to multilocus sequence typing. Environ. Microbiol. **13:**3114–3127. http://dx.doi.org/10.1111/j.1462-2920.2011.02588.x.

30. **Bobrov AG, Filippov AA.** 1997. Prevalence of IS285 and IS100 in *Yersinia pestis* and *Yersinia pseudotuberculosis* genomes. Mol. Gen. Mikrobiol. Virusol. (2):36–40. (In Russian.)

31. **Torrea G, Chenal-Francisque V, Leclercq A, Carniel E.** 2006. Efficient tracing of global isolates of *Yersinia pestis* by restriction fragment length polymorphism analysis using three insertion sequences as probes. J. Clin. Microbiol. **44:**2084–2092. http://dx.doi.org/10.1128/JCM.02618-05.

32. **Cabanel N, Leclercq A, Chenal-Francisque V, Annajar B, Rajerison M, Bekkhoucha S, Bertherat E, Carniel E.** 2013. Plague outbreak in Libya, 2009, unrelated to plague in Algeria. Emerg. Infect. Dis. **19:**230–236. http://dx.doi.org/10.3201/eid1902.121031.

33. **Bogdanovich T, Carniel E, Fukushima H, Skurnik M.** 2003. Use of O-antigen gene cluster-specific PCRs for the identification and O-genotyping of *Yersinia pseudotuberculosis* and *Yersinia pestis*. J. Clin. Microbiol. **41:**5103–5112. http://dx.doi.org/10.1128/JCM.41.11.5103-5112.2003.

34. **Carniel E, Mercereau-Puijalon O, Bonnefoy S.** 1989. The gene coding for the 190,000-dalton iron-regulated protein of *Yersinia* species is present only in the highly pathogenic strains. Infect. Immun. **57:**1211–1217.

35. **Guiyoule A, Grimont F, Iteman I, Grimont PAD, Lefèvre M, Carniel E.** 1994. Plague pandemics investigated by ribotyping of *Yersinia pestis* strains. J. Clin. Microbiol. **32:**634–641.

36. **Hunter PR, Gaston MA.** 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J. Clin. Microbiol. **26:**2465–2466.

37. **Podladchikova ON, Dikhanov GG, Rakin AV, Heesemann J.** 1994. Nucleotide sequence and structural organization of *Yersinia pestis* insertion sequence IS*100*. FEMS Microbiol. Lett. **121:**269–274. http://dx.doi.org/10.1111/j.1574-6968.1994.tb07111.x.

38. **Buchrieser C, Brosch R, Bach S, Guiyoule A, Carniel E.** 1998. The high-pathogenicity island of *Yersinia pseudotuberculosis* can be inserted into any of the three chromosomal *asn* tRNA genes. Mol. Microbiol. **30:**965–978. http://dx.doi.org/10.1046/j.1365-2958.1998.01124.x.

39. **McDonough KA, Hare JM.** 1997. Homology with a repeated *Yersinia pestis* DNA sequence IS100 correlates with pesticin sensitivity in *Yersinia pseudotuberculosis*. J. Bacteriol. **179:**2081–2085.

40. **Chain PS, Carniel E, Larimer FW, Lamerdin J, Stoutland PO, Regala WM, Georgescu AM, Vergez LM, Land ML, Motin VL, Brubaker RR, Fowler J, Hinnebusch J, Marceau M, Medigue C, Simonet M, Chenal-Francisque V, Souza B, Dacheux D, Elliott JM, Derbise A, Hauser LJ, Garcia E.** 2004. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. Proc. Natl. Acad. Sci. U. S. A. **101:**13826–13831. http://dx.doi.org/10.1073/pnas.0404012101.

41. **Fukushima H, Matsuda Y, Seki R, Tsubokura M, Takeda N, Shubin FN, Paik IK, Zheng XB.** 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. http://dx.doi.org/10.1128/JCM.39.10.3541-3547.2001.

42. **Savin C, Martin L, Bouchier C, Filali S, Chenau J, Zhou Z, Becher F, Fukushima H, Thomson NR, Scholz HC, Carniel E.** 2014. The *Yersinia pseudotuberculosis* complex: characterization and delineation of a new species, *Yersinia wautersii*. Int. J. Med. Microbiol. **304:**452–463. http://dx.doi.org/10.1016/j.ijmm.2014.02.002.