JOURNAL OF CLINICAL MICROBIOLOGY, Feb. 2005, p. 650–656 Vol. 43, No. 2
0095-1137/05/$08.00+0 doi:10.1128/JCM.43.2.650–656.2005

Identifying Sources of Human Exposure to Plague

Jennifer L. Lowell,[1,2]* David M. Wagner,[3] Bakyt Atshabar,[4] Michael F. Antolin,[2] Amy J. Vogler,[3] Paul Keim,[3] May C. Chu,†[1] and Kenneth L. Gage[1]

Division of Vector-Borne Infectious Diseases, Centers for Disease Control and Prevention,[1] *and Department of Biology, Colorado State University,*[2] *Fort Collins, Colorado; M. Aikimbayev's Kazakh Scientific Center for Quarantine and Zoonotic Diseases, Almaty, Kazakhstan*[4]*; and Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona*[3]

Received 28 July 2004/Returned for modification 30 September 2004/Accepted 17 October 2004

***Yersinia pestis*, the etiologic agent of plague, has shaped the course of human history, killing millions of people in three major pandemics. This bacterium is still endemic in parts of Asia, Africa, and the Americas, where it poses a natural disease threat to human populations. *Y. pestis* has also recently received attention as a possible bioterrorism agent. Thus, rapid methods to distinguish between bioterrorism and naturally occurring plague infections are of major importance. Our study is the first to demonstrate that variable-number tandem repeats (VNTRs) in the *Y. pestis* genome can link human case isolates to those obtained from suspected environmental sources of infection. We demonstrate the valuable utility of VNTR markers in epidemiological investigations of naturally occurring plague and the forensic analysis of possible bioterrorism events.**

Plague, which is caused by the bacterium *Yersinia pestis*, has wreaked devastation around the globe, killing millions of people in three major disease pandemics. Natural transmission of plague to humans remains a possibility in many regions of the world, where foci exist in sylvatic rodent populations (3, 8). Approximately 3,000 human cases occur worldwide annually, with 12 to 15 cases reported each year in the western United States (25). *Y. pestis* has also been identified as a potential bioterrorism agent (12), and the threat of bioterrorism or biocrimes, combined with the continuing occurrence of natural outbreaks, emphasizes the need for methods for differentiating victims of deliberate exposures from those who become infected from natural sources (14). Two of the primary objectives of routine epidemiological plague investigations are to identify the source of human exposure and to assess the exposure site for potential continuing risk. These objectives are sometimes difficult to meet when more than one epizootic source exists or when a patient's history is ambiguous. Despite the epidemic potential of *Y. pestis*, outbreak investigations and prevention efforts are often hampered both by our limited knowledge of how *Y. pestis* spreads through host populations and by a lack of methods for unambiguously identifying individual exposure sites, local sources of infection, and local populations of bacteria. The use of molecular epidemiological techniques in these investigations has been particularly difficult for *Y. pestis* because of its apparent lack of genetic variation (1). *Y. pestis* is currently grouped into three biovars (5), and while previous genotyping techniques are efficient for biovar identification, detection of genetic variability within biovars has not been consistent (9, 11, 13, 19, 22). Furthermore, a lack of high-resolution bacterial strain-typing methods has made molecular epidemiology and surveillance of *Y. pestis* difficult.

The completion of the first *Y. pestis* genome sequence (20) revealed DNA repeats that have the potential to identify variability among plague isolates on small geographic scales, and mutation rates of these DNA repeats have provided additional information on the feasibility of using these markers to identify genetically similar *Y. pestis* isolates on a local scale (7). This information has led to the development of a highly effective typing system for use in molecular epidemiology and forensic analyses (6, 16). We show the applicability of 17 multiple locus variable-number tandem repeat (VNTR) markers (MLVA) (2, 6, 16) to the molecular epidemiology and identification of environmental infection sources for human plague cases. When combined with epidemiological information, the analysis of these highly mutable VNTR markers (16, 20) enabled us to identify exposure sites and likely environmental sources of infection for past human plague cases, including a highly publicized case that occurred in New York City in November 2002 (21).

MATERIALS AND METHODS

Isolate selection. We examined 13 sets of *Y. pestis* isolates collected during epidemic investigations conducted in New Mexico in the early 1980s and in New Mexico, Arizona, and Colorado in 1992, 1996, 1999, 2001, and 2002. Three sets, used as positive location controls, consisted of paired isolates collected from different fleas or hosts but at the same time and location (Table 1). Positive control pair A was collected from an antelope ground squirrel and a flea removed from this animal, control pair B was collected from fleas found in neighboring burrows in the same prairie dog colony, and control pair C was collected from fleas in the same rodent burrow. A fourth isolate set served as a negative location control and consisted of two isolates collected in the same year but at separate sites located approximately 300 km apart (Table 1). The nine remaining sets of isolates were collected during plague case investigations in which isolates were obtained from both human patients and associated environmental samples from other mammalian hosts and fleas (Table 2). Isolates obtained from other mammals or fleas during each plague case investigation were identified and were compared genetically to the corresponding human isolate. A biovar mediavalis isolate from Kazakhstan was included in the phylogenetic analyses as an outgroup.

* Corresponding author. Mailing address: Division of Vector-Borne Infectious Diseases, Centers for Disease Control and Prevention, 3500 Rampart Rd., Fort Collins, CO 80522. Phone: (970) 225-4248. Fax: (970) 225-4257. E-mail: rzl9@cdc.gov.

† Present address: World Health Organization, Dulles, VA 20189-5120.

TABLE 1. Known environmental isolate pairs used as location controls

Control	Location	Source of paired isolates, accession no.		Genetic match[a]
		First of pair	Second of pair	
A	Sandoval County N. Mex.	*Ammospermophilus leucurus* (ground squirrel), NM00293	*Thrassis bacchi* (ground squirrel flea), NM000293-42	Y
B	La Plata County, Colo.	*Oropsylla hirsuta* (prairie dog flea), CO021867-142	*Oropsylla hirsuta* (prairie dog flea), CO021868-143	Y
C	San Miguel County, Colo.	*Thrassis bacchi* (ground squirrel flea), NM8301675-1885	*Oropsylla hirsuta* (prairie dog flea), NM830674-1879	Y
D	Santa Fe County, N. Mex. (first of pair); Bernalillo County, N. Mex. (second of pair)	Human NM012147	*Oropsylla hirsuta* (prairie dog flea), NM0113239-539	N

[a] Y, yes; N, no.

Selection of VNTR markers. A subset of 17 VNTR markers was selected from the 43 VNTR markers previously described for *Y. pestis* (2, 7, 15). The most polymorphic markers were selected because they are considered more effective for forensic analysis and for identifying genetic similarity on small geographic scales (16). In *Y. pestis*, those markers with the highest number of the repeated-motif copies show the highest degree of polymorphism across isolates tested (16) and some of the highest mutation rates in vitro (7).

DNA extraction and PCR amplification. DNA was prepared from *Y. pestis* isolates by a heat soak method (15). Each 20-μl PCR mixture contained 1× PCR buffer with 1.5 mM $MgCl_2$, a 200 μM concentration of the deoxynucleoside triphosphates, 0.5 U of *Taq* polymerase (Promega, Madison, Wis.), 1.0 μl of the DNA template (approximately 0.5 ng of DNA), and one of the following six multiplex phosporamidite linkage dye-labeled primer sets: mix 1, a 0.1 μM concentration of primer M09 and a 0.2 μM concentration each of primers M21 and M18; mix 2, a 0.1 μM concentration of primer M06 and a 0.2 μM concentration of primer M58; mix 3, a 0.1 μM concentration of primer M34 and a 0.2 μM concentration each of primers M23 and M28; mix 4, a 0.1 μM concentration of primer M31 and a 0.2 μM concentration of primer M12; mix 5, a 0.1 μM concentration of primer M27 and a 0.2 μM concentration each of primers M29 and M33; mix 6, a 0.1 μM concentration of primer M22 and a 0.2 μM concentration each of primers M25 and M59; and mix 7, a 0.2 μM concentration of primer M19. Reaction mixtures were placed on a PTC-100 thermal cycler (MJ Research, Waltham, Mass.) at 94°C for 5 min followed by 40 cycles of 94°C for 20 s, 57°C for 20 s, and 72°C for 45 s, with a final extension step of 72°C for 5 min. Following thermal cycling, samples were diluted 1:5 with sterile, DNase-free water for fragment analysis.

Fragment analysis. PCR fragments were analyzed on a CEQ 8000 DNA capillary sequencer (Beckman Coulter, Fullerton, Calif.) by adding 1.25 μl of the amplified samples to 39.5 μl of sample loading solution (Beckman Coulter) and 0.5 μl of a 600-bp D1 dye-labeled size standard (Beckman Coulter). Method parameters consisted of a 35°C capillary temperature, 120 s of denaturation at 90°C, 30 s of injection at 2.0 kV, and 35.0 min of separation at 7.5 kV. PCR fragment sizes were determined from the raw data by using the CEQ 8000 fragment analysis software version 5.0 (Beckman Coulter). After fragment sizes were determined, the number of tandem repeats per allele was calculated in reference to the previously published CO92 repeat sizes (16, 20). Repeat numbers were scored as characters for each taxon, and these data were entered into a data matrix to infer relationships among isolates.

Statistical analysis. The data matrix containing repeat numbers was entered into PAUP version 4.0b10 (23). A strict consensus tree was generated by maximum parsimony analysis, and jackknife support was determined based on 37% deletion and 500 replications. Isolates that were supported in at least 70% of jackknifed parsimony trees fit our first criterion for inferring a match between isolate pairs. A jackknife support of greater than 70% represents a greater than 95% probability of obtaining the correct clade (10).

Database query. To place the genetic relationships within our set of isolates in a global context, each of the nine human isolates and one isolate from each of the four control pairs were compared against a large *Y. pestis* DNA collection at the Keim Genetics Laboratory at Northern Arizona University (NAU). Each of the 13 isolates was compared to the 30 other isolates listed in Tables 1 and 2 and 632 additional isolates from NAU for a total of 662 pairwise comparisons for each sample. Each isolate was compared against the database in a nonnested hierarchal design on worldwide, continental, and local scales. The worldwide scale consisted of 346 biovar orientalis isolates collected outside North America, whereas the local scale consisted of 169 isolates, including our 31, from New Mexico, Arizona, Colorado, and Utah (Four Corners region). The continental scale consisted of 147 isolates from various states in the United States, excluding the Four Corners region (Table 3). Pairwise genetic distances among each of the 13 isolates of interest (one from each location control pair and each of the nine human isolates) and the other 662 isolates were generated with PAUP 4.0b10 (23) based on VNTR fragment sizes. These pairwise distances were converted to the number of marker differences; the average marker difference between samples was 9.6 (99% confidence interval [CI], 8.64 to 10.56). Isolates that matched each other (i.e., had very few or no marker differences) were identified as extreme outliers compared to the lower tail of the data set, thereby fitting our second criterion for inferring a match between isolate pairs.

Epidemiologic data collection. Health officials from federal (Centers for Disease Control and Prevention [CDC]), state, and local agencies routinely conduct investigations of human plague cases in an effort to identify likely sources of infection and persons who might be at risk. As part of these investigations, efforts are made to collect samples from potential mammalian hosts and their fleas. These samples are analyzed for *Y. pestis*, and bacterial isolates are deposited in the plague strain reference collection at CDC's Division of Vector-Borne Infectious Diseases. Investigators also determine the patients' travel histories and potential exposure sites, the proximity of patients' residences to rodent populations known to be common sources of infection (e.g., prairie dogs), patients' recollections of flea bites, patients' direct contact with wild mammals or domestic pets that may have been exposed to a plague epizootic, and other pertinent epidemiologic information. These data were used in addition to statistical analyses as a third criterion for inferring a match between isolate pairs.

Calculation of the VNTR mutation rate. An overall mutation rate for the 17 VNTR markers used in this study was calculated based on data from an in vitro parallel, serial-passage experiment recently described by Girard et al. (7), where mutations observed across ~21,000 *Y. pestis* generations yielded an overall mutation rate of 1.3×10^{-3} mutations/generation for the 43 VNTR markers. Because the markers used in the present study are a subset of the 43 used by Girard et al. (7), it was possible to calculate an overall mutation rate of 1.0×10^{-3} mutations/generation for our 17 VNTR loci. Based on the Poisson distribution, the probability of observing n mutation(s) is maximized at the number of generations that is equal to the inverse of the rate times n. For example, the probability of observing one mutation is maximized at 1,000 generations (95% CI, 26 to 5,370 generations).

Transmission cycle estimates. The number of transmission cycles that occurred between some of the human and paired environmental isolates was estimated using a recently described transmission model for *Y. pestis* (7). This model predicts that ~52 *Y. pestis* generations (doublings) occur in a single transmission cycle, which involves a single infected flea passing on a *Y. pestis* infection to a single mammalian host (7). When coupled with a mutation rate estimate for VNTR markers, this transmission model provides predictions of the number of transmission cycles that have occurred between two isolates. For

TABLE 2. Human cases and associated environmental isolates collected during case investigations

Case, date (mo/yr)	Human isolate CDC accession no., case location	Circumstance of exposure or distance of environmental isolate from potential exposure site	Environmental isolate source	Environmental isolate CDC accession no.	Genetic match[a]
A, 11/2002	NM024452, Santa Fe County, N. Mex.	Residence yard	*Orchopeas sexdentatus* (wood rat flea)	NM021852-138	Y
			Orchopeas (wood rat flea)	NM021856-140	Y
			Peromyscopsylla hesperomys (deer mouse flea)	NM024476-306	Y
				NM024477-309	Y
				NM024479-310	Y
			Anomiopsyllus nudatus (wood rat flea)	NM024484-315	Y
B, 7/1992	AZ921389, Apache County, Ariz.	200 m from residence (site 1)	*Oropsylla hirsuta* (prairie dog flea) (site 1)	AZ921367-360	Y
		27 km from residence (site 2)	*Spermophilus variegates* (rock squirrel) (site 2)	AZ921377	N
C, 1999	NM990061, Santa Fe County, N. Mex.	Patient skinned rabbit	*Sylvilagus auduboni* (rabbit)	NM990030	Y
D, 1992	AZ962456, Coconino County, Ariz.	Patient visited prairie dog town (site 1)	*Oropsylla hirsuta* (prairie dog flea) (site 1)	AZ962544-528	Y
		Patient visited prairie dog town 22.4 km from site 1 (site 2)	No isolate (site 2)		
E, 10/1992	CO92, Chaffee County, Colo.	1 km from residence	*Tamius quadrivittatus* (Colorado chipmunk)	CO921715	N
F, 8/1983	NM830692, San Miguel County, N. Mex.	Gravel pit (site 1)	*Oropsylla hirsuta* (prairie dog flea) (site 1)	NM830651-885	N
			Thrassis bacchi (ground squirrel flea) (site 1)	NM8306741-879	N
		4 km from site 1 (site 2)	No isolate (site 2)		
		4 to 8 km from site 1 (site 3)	No isolate (site 3)		
		Roaming dogs contacted patient	No isolates from roaming dogs (seropositive)		
G, 6/1983	NM830483, Santa Fe County, N. Mex.	300 m (site 1)	*Oropsylla montana* (ground squirrel flea) (site 1)	NM8304881-284	N
		Roaming cats slept with patient	No isolates from roaming cats		
H, 8/1983	NM830694, McKinley County, N. Mex.	Residence (site 1)	*Oropsylla hirsuta* (prairie dog flea) (site 1)	IJ831816-1920	N
		354 km (site 2)	No isolate (site 2)		
I, 4/1983	NM830202, McKinley County, N. Mex.	400 m from residence	*Aetheca wagnerii* (deer mouse flea)	NM83-IJ823	N

[a] Y, yes; N, no.

example, the probability of observing two mutations in the 17 VNTR markers is maximized at 2,000 generations (95% CI, 254 to 6,840 generations), which corresponds to ~38 transmission cycles (95% CI, 5 to 132 transmission cycles).

RESULTS

All positive location controls met our three criteria for inferring a positive match between isolates. First, the most-parsimonious trees generated from these samples and 17 markers had three branches that included both isolates from each positive location control (A, B, and C). Jackknifing analysis showed that support for unique pairing of isolates from each area was 98, 93, and 84%, respectively (Fig. 1). Second, each isolate pair in the positive location controls differed from its match at just one marker, did not match any of the other isolates in the NAU database query, and fell well outside the lower limit of the 99% CI for the average number of marker differences. Third, paired isolates were collected from different fleas or hosts but at the same time and location, as would be the case if matching human and environmental isolate pairs were obtained during an epidemiologic investigation. The geographically distant negative location control pair D was not supported in parsimony analyses and fell outside the upper end of the 99% CI for the average number of marker differences for all pairwise comparisons (11 marker differences).

Epidemiologic information was combined with MLVA data from human and environmental samples to verify specific plague exposure sites for each patient. The 2002 New York City plague case (case A) was a high-profile case, and because it was diagnosed in an area where plague does not cycle in wild rodent populations, urgent identification of the infective source was needed to rule out bioterrorism. The human isolate matched multiple flea isolates collected near the patient's home in Santa Fe, N.Mex., before and after the date when the patient was first exposed (jackknife support, 88%). Interestingly, matching environmental samples included not only *Y. pestis*-infected fleas collected during the follow-up case investigation but also samples obtained during routine surveillance on the patient's New Mexico property 4 months prior to the onset of illness (Table 2). MLVA results for this group of samples yielded high identity between the human isolate and the flea isolates collected before and after the case occurred (Fig. 1). Five of the environmental samples differed at only two

TABLE 3. Origins of the 632 isolates compared from the NAU *Y. pestis* MLVA type database and the 31 isolates typed in this study

Country, U.S. state, or Four Corners state and county	No. of isolates
Countries	
Belgian Congo	1
Bolivia	2
Brazil	3
Burma	2
China	4
Former USSR	1
Germany	3
India	3
Indonesia	2
Madagascar	308
Namibia	5
Senegal	2
South Africa	1
Turkey	2
Vietnam	7
Total	346
U.S. states	
California	127
Kansas	9
Montana	2
Nevada	2
Oregon	1
Texas	5
Wyoming	1
Total	147
Four Corners states and counties	
Arizona	
Apache	10
Coconino	38
Navajo	4
Yavapai	1
Colorado	
Chafee	3
Denver	1
Larimer	7
La Plata	3
Park	1
New Mexico	
Bernalillo	11
Cibola	1
Harding	1
Los Alamos	1
McKinley	3
Rio Arriba	3
San Juan	1
San Miguel	5
Sandoval	5
Santa Fe	31
Unspecified	39
Utah	
Jaub	1
Total	170

markers, while one of the environmental samples differed at three markers. The 2002 New York City human case isolate was also highly dissimilar to isolates collected in surrounding regions during other case investigations and highly dissimilar to compared isolates from the Four Corners, from the United States, and from the world (Fig. 2). Exceptions were two isolates collected in 1998 approximately 61 km from the 2002 human case and one collected in 1991 in the same county. One of the 1998 isolates also differed from the human isolate by two markers, and the other 1998 isolate and the 1991 isolate differed from the human case A isolate at three markers.

Our MLVA also linked certain human case isolates each with an environmental isolate from a single suspected exposure site even when isolates from more than one exposure site existed. Case B had two known potential exposure sites, one approximately 27 km from the patient's residence, where he was collecting wood, and the other 200 m north of his residence, where a plague epizootic had occurred in prairie dogs and other nearby rodent populations. During the case investigation, several field mice and a rock squirrel (*Spermophilus variegatus*) were collected in the immediate vicinity of the wood collection site, and a *Y. pestis* isolate was obtained from the carcass of a rock squirrel that died in a live trap. Several fleas were also collected from different types of rodent burrows near the patient's residence, and *Y. pestis* was isolated from an *Oropsylla montana* flea pool from one of these burrows. When the human isolate from case B was tested against the environmental isolates from the distant wood collection site and the rodent burrows near the patient's residence, it showed high similarity to the nearby site, with a jackknife value of 75% and only two marker differences. The isolate from the wood collection site had very little similarity to either the human or the above-described *Oropsylla montana* isolate (Fig. 1), differing at 10 markers.

In contrast to the above-described case studies, patient cases C, D, E, F, G, H, and I had multiple potential sites of exposure but with environmental isolates from only one of the sites. Case C involved a hunter who shot and skinned a rabbit from an area where plague is enzootic. A tissue sample from the dead rabbit, which was stored in the patient's freezer, yielded a *Y. pestis* isolate that was highly similar to the patient isolate, with 81% jackknife support and only one marker difference, well outside of the 99% CI of mean marker differences for the NAU database query. This match verified that the likely infection source was the rabbit and that it was unlikely that the patient was exposed to *Y. pestis* in other areas where he might have been hunting.

The patient in case D reportedly visited two potential exposure sites approximately 22.4 km apart during a 3-day period. Prairie dog die-offs, suggestive of plague, were observed at both of these sites, and rodents and fleas were sampled from both areas. The epidemiological investigation yielded only positive *Y. pestis* fleas from one of the sites, perhaps because the other site had been affected much earlier by epizootic activity and the burrows no longer harbored infected fleas. Maximum parsimony analysis generated a highly supported clade (jackknife support, 84%) between the positive flea pool isolate and the patient isolate, with three marker differences. This result provided strong evidence that the infection source was from the area in which the positive fleas were recovered.

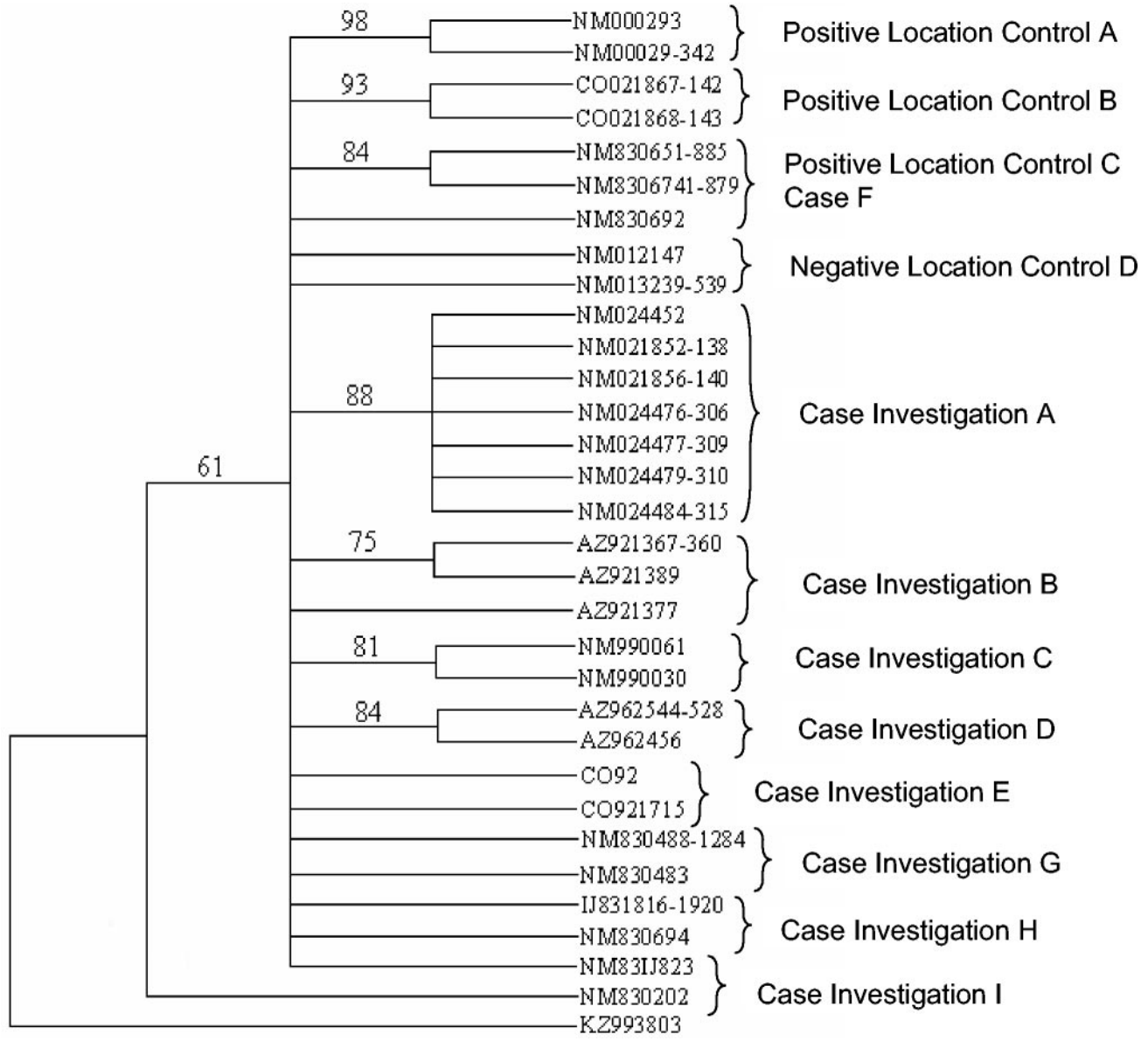

FIG. 1. Construction of a strict consensus tree by using maximum parsimony analysis identified genetically similar isolates. Number labels on the tree are accession numbers and may be referenced in Tables 1 and 2. Jackknife support values are based on 500 simulations. Branches with no numbers had values of less than 50. Isolates NM830651-885 and NM8306741-879 were used both in positive location control C and as environmental isolates for human case F.

Case E represents a situation where epidemiologic evidence clearly indicated a domestic cat as the infective source. The patient presumably became infected while removing this domestic cat from the crawlspace of a friend's home. The sick cat displayed symptoms strongly suggestive of pneumonic plague, and the patient was diagnosed postmortem with primary pneumonic plague (6). Unfortunately, the cat died prior to examination and was incinerated at a local veterinary practice before investigators arrived, precluding *Y. pestis* isolation attempts. One environmental sample was isolated from the carcass of a Colorado chipmunk (*Tamius quadrivittatus*) collected approximately 1 km from the patient's residence. This isolate was paired with the human isolate from case E to see if it may have been related to the infective source. MLVA detected some similarity, with 13 markers in common; however, this isolate pair was not considered a match because it was not collected at the actual exposure site.

Successful environmental sample collection for cases F, G, H, and I ranged from 300 m to 0.4 km from the patients' residences or potential exposure sites; however, in each instance, epidemiological data indicated that patients had traveled as far as 354 km to other potentially plague-affected areas in New Mexico (Table 2). No isolates were obtained from environmental investigations done at these alternative exposure sites. In addition, cases F and G had roaming household pets that potentially covered several kilometers surrounding

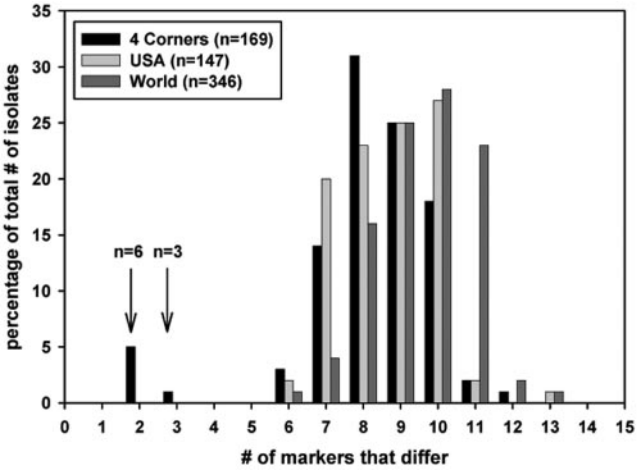

FIG. 2. Case A marker difference distribution when all of the isolates from the NAU database ($n = 632$), as well as the isolates typed in our study ($n = 30$), were compared to the human case A isolate. Isolates showing two marker differences from the human case A sample ($n = 7$) consist of the six environmental isolates associated with the case and one 1998 isolate. Isolates showing three marker differences from the human case A isolate ($n = 2$) consist of one 1998 and one 1991 isolate. Matching isolates fall in the extreme lower tail of the distribution, while the nonmatching isolates fall in the 99% CI or in the extreme upper tail of the distribution.

patients' residences before returning home with dead rodents and live fleas. When tested by MLVA, the environmental and human isolate pairs were not supported in jackknifed parsimony trees, and marker differences ranged from 7 to 15 loci, which is typical of the number of marker differences observed for nonmatching isolates seen in the NAU database. We therefore concluded that the correct exposure sites were not successfully sampled.

DISCUSSION

Three inferences may be used in combination to support our conclusions that particular isolate pairs do indeed represent a match. First, in jackknifed parsimony trees, samples that were considered a match were highly supported, with jackknife values ranging from 75 to 98%, providing >99% confidence of a correct match. Second, isolates that were called matches were extreme outliers from the lower limit of the 99% CI of mean marker differences (663 isolates). Third, patient history and data collected during epidemiological investigations supported the match on a temporal and geographical scale. Based on mutation rate data and transmission modeling, we expected to see some slight genetic variation between matching human and environmental samples in those markers that mutate the fastest, and it is not surprising that those isolate pairs with the strongest statistical support differ at one to three markers (average, two). Based on the transmission model, 2,000 (95% CI, 254 to 6,840) generations or approximately 38 transmission cycles are required to see two mutations (7), and this number of transmission cycles probably would occur during an epizootic period or during one or two seasons of ongoing enzootic transmission in a plague focus such as that identified for case A. In contrast, the genetic similarity seen between the older 1991 and 1998 isolates and the human case A isolate may be the result of these samples arising from a common origin but undergoing limited enzootic transmission and few mutations during those years in which epizootic activity was not very evident. It is also possible that these isolates share alleles that are not identical by descent but are similar because of parallel or convergent evolution. The markers chosen for this comparison are rapidly evolving, and therefore an increased likelihood that the same allelic state could arise through separate mutations (homoplasy) exists; however, we do not feel that this was a common phenomenon in our data set because of the lack of additional randomly matching isolates in the NAU database query. Whether these isolates illustrate an example of homoplasy or they arose from the same epizootic source, the epidemiological data do not support the possibility of a match, as the sample isolations precede the human case by 11 and 4 years, respectively. If the New York City 2002 plague case had been a case of bioterrorism, the human isolate still would have been traced to the correct region and even pinpointed to Santa Fe County, even when it was compared to plague isolates from around the world.

The epidemiological information collected during investigations is helpful in deciding where sampling should occur, but our data indicate that a definitive decision as to where the infective source arose should not be made based on these data alone. For example, in the original 1992 investigation of case B, the identification of an abscess on the patient's abdomen led investigators to believe that he was exposed to an infectious flea bite while carrying wood to his vehicle. Our MLVA results, however, strongly suggest that the patient was exposed near his home, as indicated by the close match between his isolate and the one obtained from the flea pool collected from a prairie dog town near his residence. This example demonstrates the power and importance of using genomic diversity to ascertain likely exposure scenarios when epidemiologic data are inconclusive or contradictory.

Case D visited multiple potential exposure sites, but samples could not be obtained from all of them. It was important in this case to identify the correct exposure site, as the case was fatal and various members of the patient's family resided near two of the rural areas that the patient visited, with an additional site being near the patient's residence and a high school (24). Although plague warnings are posted and appropriate precautions are taken in all suspect areas in cases such as these, a definitive answer as to where the infective source arose can greatly assist public health officials in allocating limited personnel and other resources.

These cases provide examples of how MLVA verified infective plague sources when it was not clear in the original investigation. By combining epidemiological information with matching isolate MLVA data, the likely exposure sites and often the infective sources can be identified. A nonmatching environmental isolate can help investigators appropriately decide whether environmental sampling should be continued at a particular site, whether additional potentially infective sites should be investigated further, or whether a simple warning should be issued in those areas not successfully sampled. The human isolates in cases F, G, H, and I, which all occurred during 1983 in northern New Mexico, did not match the corresponding environmental isolates or any of the isolates in the

NAU database. This result might be expected for a period of widespread, intense epizootic activity that occurred in 1983, when more cases were reported in the United States than had been seen since 1920 (4, 17). Isolates that did not match were collected over a widespread area in the Southwest during the 1980s plague epidemic, suggesting that this outbreak did not arise from a single source but rather from activity in many small plague foci scattered throughout the Southwest. We believe that such results are to be expected when a very widespread outbreak occurs and multiple *Y. pestis* clones arising from many sources spread quickly across a region, eventually overlapping in distribution with each other. The plague outbreak of the early to mid-1980s represented such an event. Because of the high number of cases, only those sites likely to pose threats to other humans were thoroughly investigated by intensive trapping of rodent hosts and collection of flea vectors. Given the genetic dissimilarity among isolates obtained from cases F, G, H, and I, it seems that these cases were infected at alternate sites or by additional widely circulating genotypes that might have spread from neighboring plague-affected areas. While the epidemiological investigations and *Y. pestis* sampling efforts in these cases yielded helpful information, definitive statements about the actual infective source could not be made.

Our study presents an analytic strategy involving both epidemiologic data and MLVA and demonstrates the use of MLVA on multiple case studies, including one where the diagnosis was made a half continent away from the infective source. When combined with epidemiologic information, judicious use of genetic data from nonhuman organisms is highly attractive because of the power of DNA-based analyses to identify exposure sources (14, 16). However, this prospect has proven contentious, as experts disagree upon valid criteria for determining a match among samples (18). Our MLVAs of the above-described human and environmental *Y. pestis* isolates clearly demonstrate the value of this technique for the identification of likely sources of infection and sites of exposure for human plague cases. When coupled with case histories and other epidemiological information, MLVA should also be useful for differentiating naturally occurring cases from those occurring from an intentional *Y. pestis* release.

ACKNOWLEDGMENTS

This work was supported by the ISTC Biotechnology Engagement Program award K-584p.

We thank Mark P. Simmons from the Colorado State University Department of Biology for technical assistance with phylogenetic analyses; Sandra K. Urich from the Division of Vector Borne Infectious Diseases, Centers for Disease Control and Prevention, for laboratory support; and the Arizona Department of Health Services, New Mexico Department of Health, Colorado Department of Health and Environment, New York State Department of Health, and Navajo Area Indian Health Services for assistance with field investigations.

REFERENCES

1. **Achtman, M. A. 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. **Adair, D. M., P. L. Worsham, K. K. Hill, A. M. Klevytska, P. J. Jackson, A. M. Friedlander, and P. Keim.** 2000. Diversity in a variable-number tandem repeat from *Yersinia pestis*. J. Clin. Microbiol. **38:**1516–1519.
3. **Barnes, A. M.** 1982. Surveillance and control of bubonic plague in the United States. Symp. Zool. Soc. Lond. **50:**237–270.
4. **Craven, R. B., G. O. Maupin, M. L. Beard, T. J. Quan, and A. M. Barnes.** 1993. Reported cases of human plague infections in the United States, 1970–1991. J. Med. Entomol. **30:**758–761.
5. **Devignat, R.** 1951. Variétés de l'espèce *Pasteurella pestis*. Nouvelle hypothèse. Bull. W. H. O. **4:**247–263.
6. **Doll, J. M., P. S. Zeitz, P. Ettestad, A. L. Bucholtz, T. Davis, and K. Gage.** 1994. Cat-transmitted fatal pneumonic plague in a person who traveled from Colorado to Arizona. Am. J. Trop. Med. Hyg. **51:**109–114.
7. **Girard, J. M., D. M. Wagner, A. J. Vogler, C. Keys, C. J. Allender, L. C. Drickamer, and P. Keim.** 2004. Differential plague-transmission dynamics determine *Yersinia pestis* population genetic structure on local, regional, and global scales. Proc. Natl. Acad. Sci. USA **101:**8408–8413.
8. **Gratz, N.** 1999. Rodent reservoirs and flea vectors of natural foci of plague, p. 63–95. *In* N. Gratz, J. D. Poland, and E. Tikhomirov (ed.), Plague manual: epidemiology, distribution, surveillance and control. World Health Organization, Geneva, Switzerland.
9. **Guiyoule, A., B. Rasoamanana, C. Buchrieser, P. Michel, S. Chanteau, and E. Carniel.** 1997. Recent emergence of new variants of *Yersinia pestis* in Madagascar. J. Clin. Microbiol. **35:**2826–2833.
10. **Hillis, D. M., and J. J. Bull.** 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. **42:**182–192.
11. **Huang, X.-Z., M. C. Chu, D. M. Engelthaler, and L. E. Lindler.** 2002. Genotyping of a homogenous group of *Yersinia pestis* strains isolated in the United States. J. Clin. Microbiol. **40:**1164–1173.
12. **Inglesby, T. V., D. T. Dennis, D. A. Henderson, J. G. Bartlett, M. S. Ascher, E. Eitzen, A. D. Fine, A. M. Friedlander, J. Hauer, J. F. Koerner, M. Layton, J. McDade, M. T. Osterholm, T. O'Toole, G. Parker, T. M. Perl, P. K. Russell, M. Schoch-Spana, and K. Tonat.** 2000. Plague as a biological weapon medical and public health management. JAMA **283:**2281–2290.
13. **Iteman, I., C. Baril, I. Saint Girons, and E. Carniel.** 1991. Pulse field electrophoresis of the chromosome of the pathogenic yersiniae. Contrib. Microbiol. Immunol. **12:**198–202.
14. **Keim, P.** 2003. Microbial forensics: a scientific assessment. American Academy of Microbiology, Washington, D.C.
15. **Keim, P., L. B. Price, A. M. Klevytska, K. L. Smith, J. M. Schupp, R. Okinaka, P. J. Jackson, and M. E. Hugh-Jones.** 2000. Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J. Bacteriol. **182:**2928–2936.
16. **Klevytska, A. M., L. B. Price, J. M. Schupp, P. L. Worsham, J. Wong, and P. Keim.** 2001. Identification and characterization of variable-number tandem repeats in the *Yersinia pestis* genome. J. Clin. Microbiol. **39:**3179–3185.
17. **Link, V. B.** 1955. A history of plague in the United States of America. Public Health Monograph no. 26. U.S. Department of Health, Education and Welfare, Washington, D.C.
18. **Metzker, M. L., D. P. Mindell, X. M. Liu, R. G. Ptak, R. A. Gibbs, and D. M. Hillis.** 2002. Molecular evidence of HIV-1 transmission in a criminal case. Proc. Natl. Acad. Sci. USA **99:**14292–14297.
19. **Motin, V. L., A. M. Georgescu, J. M. Elliot, H. Ping, P. L. Worsham, L. L. Ott, T. R. Slezak, B. A. Sokhansanj, W. M. Regala, R. R. Brubaker, and E. Garcia.** 2002. Genetic variability of *Yersinia pestis* isolates as predicted by PCR-based IS*100* genotyping and analysis of structural genes encoding glycerol-3-phosphate dehydrogenase (*glpD*). J. Bacteriol. **184:**1019–1027.
20. **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.
21. **Perlman, D. C., R. Primas, B. Raucher, R. Lis, B. Weinberg, A. Davilman, C. Yampierre, J. Protic, D. Weiss, J. Ackelsberg, L. Lee, M. Layton, S. T. Beatrice, P. F. Smith, P. J. Ettestad, P. J. Reynolds, C. M. Sewell, R. E. Enscore, M. Y. Kosoy, K. Kubota, J. L. Lowell, M. Chu, J. Kool, K. L. Gage, C. C. Chow, and C. B. Smelser.** 2003. Imported plague—New York City 2002. Morb. Mortal. Wkly. Rep. **52**:725–728.
22. **Pourcel, C., A. André-Mazeaud, H. Neubauer, F. Ramisse, and G. Vergnaud.** 2004. Tandem repeats analysis for the high resolution phylogenetic analysis of *Yersinia pestis*. BMC Microbiol. **4:**22. (Online.) http://www.biomedcentral.com/1471–2180/4/22.
23. **Swofford, D. L.** 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods), 4.0b10 ed. Sinauer Associates, Sunderland, Mass.
24. **Tenborg, M., B. Davis, D. Smith, C. Levy, B. Koehler, D. Tanda, J. Pape, R. Hoffman, R. Fulgham, B. Joe, and J. Cheek.** 1997. Fatal human plague—Arizona and Colorado, 1996. Morb. Mortal. Wkly. Rep. **46:**617–620.
25. **World Health Organization.** 1999. Weekly epidemiological record. World Health Organization, Geneva, Switzerland.