JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 2004, p. 1694–1702 Vol. 42, No. 4
0095-1137/04/$08.00+0 DOI: 10.1128/JCM.42.4.1694–1702.2004

Multilocus Short Sequence Repeat Sequencing Approach for Differentiating among *Mycobacterium avium* subsp. *paratuberculosis* Strains

Alongkorn Amonsin,[1] Ling Ling Li,[1] Qing Zhang,[1] John P. Bannantine,[2] Alifiya S. Motiwala,[3] Srinand Sreevatsan,[3] and Vivek Kapur[1]*

Department of Microbiology and Biomedical Genomics Center, University of Minnesota, St. Paul, Minnesota 55108[1]; National Animal Disease Center, U.S. Department of Agriculture, Ames, Iowa 50010[2]; and Food Animal Health Research Program, Ohio Agricultural Research and Development Center, and Department of Veterinary Preventive Medicine, The Ohio State University, Wooster, Ohio 44691[3]

Received 14 August 2003/Returned for modification 23 October 2003/Accepted 2 January 2004

We describe a multilocus short sequence repeat (MLSSR) sequencing approach for the genotyping of *Mycobacterium avium* subsp. *paratuberculosis* (*M. paratuberculosis*) strains. Preliminary analysis identified 185 mono-, di-, and trinucleotide repeat sequences dispersed throughout the *M. paratuberculosis* genome, of which 78 were perfect repeats. Comparative nucleotide sequencing of the 78 loci of six *M. paratuberculosis* isolates from different host species and geographic locations identified a subset of 11 polymorphic short sequence repeats (SSRs), with an average of 3.2 alleles per locus. Comparative sequencing of these 11 loci was used to genotype a collection of 33 *M. paratuberculosis* isolates representing different multiplex PCR for IS*900* loci (MPIL) or amplified fragment length polymorphism (AFLP) types. The analysis differentiated the 33 *M. paratuberculosis* isolates into 20 distinct MLSSR types, consistent with geographic and epidemiologic correlates and with an index of discrimination of 0.96. MLSSR analysis was also clearly able to distinguish between sheep and cattle isolates of *M. paratuberculosis* and easily and reproducibly differentiated strains representing the predominant MPIL genotype (genotype A18) and AFLP genotypes (genotypes Z1 and Z2) of *M. paratuberculosis* described previously. Taken together, the results of our studies suggest that MLSSR sequencing enables facile and reproducible high-resolution subtyping of *M. paratuberculosis* isolates for molecular epidemiologic and population genetic analyses.

Mycobacterium avium subsp. *paratuberculosis* (*M. paratuberculosis*) is a slowly growing, acid-fast, mycobactin J-dependent bacterium. Infection with this bacterium leads to a chronic granulomatous enteritis, termed Johne's disease, in cattle and other ruminants and occurs worldwide (11). Clinical signs of the disease include diarrhea, weight loss, fatigue, decreased milk production, and mortality. Infection with this pathogen results in considerable economic losses in the dairy production industry, with estimated annual costs of $40 to $227 per year per cow, totaling industry-wide annual losses of $1.5 billion (20, 21). In addition to the serious health and economic impacts of the pathogen to the dairy industry, several reports suggest a possible link between *M. paratuberculosis* and Crohn's disease in humans (5, 7, 10, 30).

Methods for differentiation or subtyping of bacterial strains provide important information for molecular epidemiologic analysis and assist in providing an understanding of the population genetics of the species. DNA-based molecular subtyping techniques such as multiplex PCR for IS*900* integration loci (MPIL) (4, 20), restriction fragment length polymorphism (RFLP) analysis (6, 8, 23), and amplified fragment length polymorphism (AFLP) analysis (20) have been previously applied to investigate genetic variation in *M. paratuberculosis*. However, the MPIL, AFLP, and RFLP techniques are generally unable to resolve *M. paratuberculosis* isolates into meaningful epidemiologic groups due to the apparently restricted genetic diversity within the subspecies. Furthermore, the data generated by these techniques are biallelic and, hence, are able to provide only limited information regarding the overall genetic diversity and evolutionary mechanisms within the species.

Short sequence repeats (SSRs) or variable-number tandem repeats (VNTRs) in bacterial DNA have been used as markers for the differentiation and subtyping strains of several bacterial species, including *Mycobacterium tuberculosis* (9, 16), *Yersinia pestis* (1), *Salmonella enterica* subsp. *enterica* serovar Typhimurium (17), and *Bacillus anthracis* (15). SSRs consist of simple homopolymeric tracts of a single nucleotide (mononucleotide repeats) or multimeric tracts (homogeneous or heterogeneous repeats), such as di- or trinucleotide repeats, which can be identified as VNTRs in the genome of the organism (35). The variability of the repeats is believed to be caused by slipped-strand mispairing (31); the genetic instability of polynucleotide tracts, especially poly(G-T) (12); and DNA recombination between homologous repeat sequences (32).

The complete genome sequence of *M. paratuberculosis* strain K10 (GenBank accession number AE016958) has recently been characterized in the Department of Microbiology and Biomedical Genomics Center, University of Minnesota (L. L. Li et al., unpublished data). Preliminary bioinformatic analyses led to the identification of numerous SSRs in the *M. paratu-*

* Corresponding author. Mailing address: University of Minnesota, 1500 Gortner Ave., St. Paul, MN 55108. Phone: (612) 625-7712. Fax: (612) 624-6264. E-mail: vkapur@umn.edu.

TABLE 1. *M. paratuberculosis* isolates examined in this study

Strain	Host	Geographic origin	Yr of isolation	MPIL type[a]	AFLP type[b]	MLSSR cluster[c]	MLSSR type[d]
MAP-06[e]	Ovine	Ohio	NA[f]	A1	NA	M1	1
MAP-08[e]	Bovine	Ohio	2001	A18	NA	M2	2
0033	Caprine	Ohio	NA	A18	Z1	M2	3
0016	Bovine	Ohio	NA	A18	Z6	M2	4
0041	Bovine	N.Y.	1983	A18	Z9	M2	4
0029	Caprine	Iowa	NA	A18	Z2	M2	4
0239	Caprine	Ohio	NA	A18	Z2	M2	4
0028	Bovine	Ohio	NA	A18	Z11	M2	5
0240	Bovine	Iowa	1983	A18	Z21	M2	5
MAP-09[e]	Bovine	Ohio	2001	NA	NA	M3	6
0034	Bovine	Minn.	1984	A18	Z1	M3	7
0161-2	Rabbit	Minn.	2001	NA	NA	M3	7
0237	Rabbit	Minn.	2001	NA	NA	M3	7
0560	Soil	Minn.	2001	NA	NA	M3	8
0026	Bovine	Ohio	NA	A18	Z5	M3	9
MAP-K10[e]	Bovine	Wis.	1990	A18	Z2	M3	10
0883	Deer	Minn.	2001	NA	NA	M3	10
0012	Murine	Iowa	1984	A18	Z1	M3	10
0180	Bovine	Ohio	2001	A18	Z3	M3	11
MAP-14[e]	Human	NA	NA	A13	Z15	M3	11
0003	Human	NA	NA	A13	Z15	M3	11
0040	Caprine	Ohio	NA	A18	Z1	M3	12
0558	Bovine	Minn.	2001	NA	NA	M3	13
0161	Bovine	Ohio	2001	A18	Z1	M3	14
0011	Bovine	Ohio	NA	A16	Z12	M3	15
0030	Bovine	Ind.	1984	A18	Z2	M3	16
0015	Caprine	Iowa	1984	A18	Z1	M3	17
0004	Human	NA	NA	A18	Z10	M3	17
0014	Bovine	Ohio	NA	A18	Z4	M3	18
MAP-11[e]	Ovine	S.D.	NA	A8	NA	N	19
0007	Ovine	S.D.	NA	A16	Z18	N	20
0008	Ovine	S.D.	NA	A16	Z7	N	20
0099	Ovine	S.D.	NA	A17	Z8	N	20
D[g]				0.501	0.920		0.967

[a] Isolates were previously identified by MPIL analysis (20).
[b] Isolates were previously identified by AFLP analysis (20).
[c] Isolates were identified by MLSSR analysis based on a genetic distance of 0.43.
[d] Isolates were identified by MLSSR analysis based on 100% similarity.
[e] *M. paratuberculosis* isolates used to examine polymorphisms of 78 SSR loci.
[f] NA, not available.
[g] *D* was calculated by using the equation described previously (13).

berculosis genome. We evaluated the utility of a multilocus SSR (MLSSR)-based typing approach for differentiating among isolates of *M. paratuberculosis.* The results of our studies suggest that MLSSR is a useful approach for strain differentiation and enables the rapid and facile discrimination of epidemiologically and geographically distinct strains of *M. paratuberculosis.*

MATERIALS AND METHODS

Bacterial isolates and DNA isolation. A total of 33 *M. paratuberculosis* isolates from different host species and geographic locations were used in this study, as shown in Table 1. *M. paratuberculosis* isolates were grown on Middlebrook 7H9 broth or 7H11 agar (Difco Laboratories, Detroit, Mich.) with oleic acid-albumin-dextrose-catalase supplement (Becton Dickinson, Sparks, Md.) and mycobactin J (2 mg/100 ml). In some instances the bacterial cultures were incubated at 37°C for 4 to 6 months until colonies were observed. DNA was isolated from the bacterial culture by use of the QIAamp DNA Mini kit (Qiagen Inc., Valencia, Calif.), as described previously (20).

Database search for SSRs and primer design. The whole-genome sequence of *M. paratuberculosis* strain K10 was analyzed for SSRs with Tandem Repeat Finder (version 2.02) software (3). The coordinates of the SSRs were then matched for the regions upstream and downstream to locate the repeats and open reading frame (ORF) flanking the repeat by use of the DNA sequence viewer and annotation software Artemis (28). Primers specific for sequences flanking these repeat sequences were designed with Primer 3 software (27) to yield an average amplification product of ~250 bp for each sequence (Table 2).

MLSSR. A total of 78 loci were amplified by PCR with specific primers, and the amplification products were sequenced to identify sequence polymorphisms in each locus among six strains of *M. paratuberculosis* (reference strain MAP-K-10 and isolates from cattle [isolates MAP-08 and MAP-09], sheep [isolates MAP-06 and MAP-11], and a human [isolate MAP-14]) (Table 1). These six *M. paratuberculosis* isolates were selected because they represent the extent of genetic diversity in the species, as previously identified by MPIL and AFLP analyses (20).

The 25-μl PCR amplification reaction mixture for each SSR comprised 1× PCR buffer II (Perkin-Elmer, Applied Biosystems Division, Foster City, Calif.), 2.0 mM $MgCl_2$ (Perkin-Elmer), 200 μM each deoxynucleoside triphosphate (Roche Diagnostic Co., Indianapolis, Ind.), 0.6 μmol of each primer (Integrated DNA Technologies, Coralville, Iowa), 0.5 U of *AmpliTaq* Gold (Perkin-Elmer), 5% dimethyl sulfoxide (Sigma Chemical Co, St. Louis, Mo.), and 1 μl of DNA. The amplification conditions consisted of an initial denaturation at 94°C for 15 min, followed by 35 cycles of denaturation at 94°C for 45 s, annealing at 60°C for 1 min, and extension at 72°C for 2 min 30 s, with a final extension step at 72°C for 7 min. A 2-μl volume of the PCR products was mixed with 2 μl of loading buffer (0.2% Orange G in 50% glycerol), and the mixture was electrophoresed in 1% agarose with 0.5 μg of ethidium bromide per ml. The gels were photographed

TABLE 2. Primers for 78 sequence repeat loci used for polymorphism analysis of the region

Repeat and locus	Position in genome[a]	SSR	Forward primer	Forward primer size (bp)	Reverse primer	Reverse primer size (bp)
Mononucleotide						
1793091	1793091–1793109	GGGGGGGGGGGGG GGGGGGGG	TCAGACTGTGCGGTATGGAA	20	GTGTTCGGCAAAGTCGTTGT	20
2719085	2719085–2719094	GGGGGGGGGG	GTGACCAGTGTTTCCGTGTG	20	TGCACTTGCACGACTCTAGG	20
Dinucleotide						
3803814	3803814–3803824	GC GC GC GC GC G	ACCGAGCCGATAGTCATCAG	20	CAGCAGCAGCGAGTACGA	18
3840791	3840791–3840801	GC GC GC GC GC G	CTCGATTCGCGATCAGGT	18	CAACTTGACGCCCTGGTACT	20
607784	607784–607794	CG CG CG CG CG C	ATCCAACGCCATGTACTCGT	20	GAGCAGCATCGAGGTGAAAC	20
914968	914968–914978	CG CG CG CG CG C	GACTAGGCCCTTGCGGTATC	20	CGCAACGTGCTGTCGTAG	18
3294684	3294684–3294693	GC GC GC GC GC	GGCAACGCCTCGTACACC	18	CCTACCCGAGCGGTCACT	18
3363782	3363782–3363791	AC AC AC AC AC	AGCGGCTCAATTACACACAA	20	ATCAGGTCCGGAATCACCTT	20
3406364	3406364–3406373	GC GC GC GC GC	GTTCTCGATGGACAGCTTGC	20	GGAGGACGAACCACACTCAT	20
3424469	3424469–3424478	CG CG CG CG CG	ACCGGTTTCAGACAATGGAG	20	GTACGGCGAGCTACGCTATC	20
3456738	3456738–3456747	GC GC GC GC GC	AGCACAAGAAGCACCGGTAT	20	GGATCAACCTCGAGATCCTG	20
3458712	3458712–3458721	CG CG CG CG CG	ACCAACTGCAGATCCTCGAT	20	GTAATCCCAGCCGGTTCAT	19
3573587	3573587–3573596	CG CG CG CG CG	ATCGCCGACTATCTCAATCC	20	ACCGCCAGATAGTGAATTGC	20
3676817	3676817–3676826	CG CG CG CG CG	ACCCGGACACGACGTAGC	18	GTTCGACATCGTGCACCTC	19
3735342	3735342–3735351	GC GC GC GC GC	TGTCGAACTTGCTCTTGGTG	20	CGTCCTGCAACATCTTCTCC	20
3754236	3754236–3754245	GC GC GC GC GC	ACACGTAGTCGGTGGAGACG	20	GTTGCAATCACACGAACCAC	20
3831350	3831350–3831359	GC GC GC GC GC	GCACGCATCTGTTCAACG	18	CACAACAAGATTGGCCTCAG	20
3951841	3951841–3951850	CG CG CG CG CG	ACGCCGCACTACGAATATCT	20	ACATCTCGAAGGACGTTTCG	20
4086404	4086404–4086413	GC GC GC GC GC	AGTCGCTGGGTTTTCAGGA	19	ATACCGCACGGTGTTCTGAT	20
4159074	4159074–4159083	GC GC GC GC GC	TATTCGGGTCCATGCTCAAT	20	ATGTGACGGAGGTCACACTG	20
4176609	4176609–4176618	CG CG CG CG CG	TTCATCGACTACCGGCTCTT	20	TGTTGGGGGACCATGTAACT	20
4465488	4465488–4465497	CG CG CG CG CG	AAGCTGAACTGGTGGGAGAG	20	GTCGAAAACGCTGTCGTAGG	20
4518335	4518335–4518344	CG CG CG CG CG	CTTTGGTTCCCAGCAGCAT	19	GGAGTATTCACCCGACCAAA	20
4717322	4717322–4717331	GC GC GC GC GC	TACCAACATCCCGACCTGAC	20	CGTAGGGGATAACCTGCTGA	20
14474	14474–14483	GC GC GC GC GC	ATACCTGCGTCCGATACCAG	20	GCTTTCATGATCACCGGTTT	20
70461	70461–70470	CG CG CG CG CG	CGATGGCCTACTTCATCGTC	20	CGGCCGAAATAGATGATGTT	20
70930	70930–70939	GC GC GC GC GC	CGATCACCCAGCAGTATGTG	20	ATCAGCTCCTTGGTCACCTG	20
78012	78012–78021	GC GC GC GC GC	CCATGCTACGAGGTCGAGTT	20	TTGGGGAATAAACGACTTGC	20
191084	191084–191093	GC GC GC GC GC	CAGCCGCAACGACTTCTC	18	TTAGGGTTGGCTTCCCATTA	20
219320	219320–210329	GC GC GC GC GC	TTGTCCAGCAAGCAGAAGTG	20	ATAGTGCACCCCCAGCAC	18
274406	274406–274415	TG TG TG TG TG	AGTGGAGGCTGAGATGTTGG	20	GGTGAACACCTTGCCGTTAT	20
429640	429640–429649	GC GC GC GC GC	AAGTACATCCCGCTGCACAC	20	CGATTACCAGGTGCACGAC	19
676774	676774–676783	GC GC GC GC GC	GCGAAATAACCGTTCACCTC	20	GACACCACCTGCGAGTAACC	20
686290	686290–686299	CA CA CA CA CA	AGATCGCATCAAAGAGCACCT	21	GGTGAGTTGTCCGCATCAG	19
762468	762468–762477	CG CG CG CG CG	GTCGACCCGAAGAGTGAGTG	20	GAATTTTTGGGGTCGTGATG	20
899250	899250–899259	CG CG CG CG CG	AGAAAATAGCTGCGGTCGAA	20	GGATCGACACCCTGACCTC	19
909353	909353–909362	CG CG CG CG CG	CGATCTCATACACCGTCGTC	20	AGGTGAACCGTAAGCGACAC	20
970000	970000–970009	CG CG CG CG CG	TCGAGACCTCAAAAGCCTTG	20	GGGGACCTGCTGGTGATAG	19
2892461	2892461–2892470	GT GT GT GT GT	GGGTTGGTCTACGTTTGTCG	20	CCACAGCCCCTCGTAGTG	18
2887399	2887399–2887408	CG CG CG CG CG	AGCTGTACCGCGACTACCAC	20	AGCAACCGCGAATATGGTC	19
2831533	2831533–2831542	GC GC GC GC GC	CTGTTCGCCTACGTGCTGTA	20	GGGGTGGAGACGATGAAATA	20
2808989	2808989–2808998	CG CG CG CG CG	GGTGCGCGATAATGAAACTC	20	AAGACCACGCTGGTGAATCT	20
2103815	2103815–2103194	GC GC GC GC GC	GACCAGTTCCGGGCTCAC	18	AGGAACTGTTCAACCCGATG	20
2716662	2716662–2716671	CA CA CA CA CA	ACCGATCTGGAAGAACATCG	20	GCCCTGGTCATTACACGACT	20
2709505	2709505–2709514	CG CG CG CG CG	GAGGTCGTCTTCCGTTACGA	20	CTCGTCGACCAGCAGTATGG	20
2615343	2615343–2615352	GC GC GC GC GC	AGGAAGGCGTCGACAAGAT	19	GAGGTGTCGTGCTTGGAGAT	20
2583850	2583850–2583859	CG CG CG CG CG	CGGCTTCGTATTGTCGTCTT	20	GGTCATGAGCAGAAACCTTCC	20
2582523	2582523–2582532	CA CA CA CA CA	CAGATCGCATCAAAGAGCAC	20	GGTGAGTTGTCCGCATCAG	19
2540968	2540968–2540977	GC GC GC GC GC	ATTCAACCAACAGCCTCAGC	20	TAGCCGTTACCGGTGTTGAT	20
2519120	2519120–2519129	CG CG CG CG CG	CCCACCAGGTCGAAGAAAT	19	CAACTGGAGGACCAGGTGTT	20
2509961	2509961–2509970	GC GC GC GC GC	GCATCTCCACCCAGAAATTG	20	GAACACGTCGGTCTTGGTTT	20
2289642	2289642–2289651	GC GC GC GC GC	GATCGAAGACCGAAAACGTC	20	AGATCATCGGTGAGCTGGAG	20
2259726	2259726–2259735	GC GC GC GC GC	GAAGGGTCGTTCATGTTGCT	20	CAACTTACGGACCTGGGATG	20
2059817	2059817–2059826	GC GC GC GC GC	GTCGGGTTCTTCGTCAACAC	20	ATCGGTATCCATCAGGTCCA	20
1960967	1960967–1960976	GC GC GC GC GC	GCATTGCGCTACCTGAGTC	19	TCGACGAGAACATCACGAAC	20
1860248	1860248–1860257	CG CG CG CG CG	ACCTGCAGACCGACGATTAC	20	ACTTGCTCACCGAGAACAGC	20
1841444	1841444–1841453	CG CG CG CG CG	GCAGCTTGTCCAGATCGAA	19	AAAAAGCAGCGACACCAGAC	20
1829129	1829129–1829138	CG CG CG CG CG	GAGGACCACGTGAAAATCGT	20	GGATCCTGCACCAGAACCT	19
1787205	1787205–1787214	CG CG CG CG CG	TGATCATCATGGAAGCCAAC	20	GCGGGAATGTTGATAAGGAA	20
1758285	1758285–1758294	CG CG CG CG CG	GATCATCCAATCGGTGTCCT	20	GCACACTCGTAATCGCTCAA	20
1707736	1707736–1707745	GC GC GC GC GC	GACCACCAAAACTGGTTTCC	20	GTCCGGTAGTGGTCGATGTC	20
1684441	1684441–1684450	GC GC GC GC GC	CAGGAACTGTGGGATGTCCT	20	CTGTACGGCTATTCGGTGGT	20
1664073	1664073–1664082	CG CG CG CG CG	TTCTGGCCGAATTGATCTCT	20	ATCGTTTTTGCCTGAATTGG	20
1526624	1526624–1526633	GC GC GC GC GC	GTCACCATCCGGTACATTCC	20	GAGGTGCCCAAGACGTATCT	20
1257495	1257495–1257504	CG CG CG CG CG	GGGATCCTGTGGCAGATAAC	20	CAACTGCTGGACACCTGCTA	20
Trinucleotide						
4286068	4286068–4286084	GCG GCG GCG GCG GC	GAATCGTCTTGCCTCACTGG	20	TCGAGCAACTGATCTCCACA	20
4310932	4310932–4310948	CCG CCG CCG CCG CCG CC	CGGCAATACCTCGAACAGAT	20	GCTGAAGAGGTCGTGCAGAT	20

Continued on following page

TABLE 2—*Continued*

Repeat and locus	Position in genome[a]	SSR	Forward primer	Forward primer size (bp)	Reverse primer	Reverse primer size (bp)
440731	440731–440747	TGG TGG TGG TGG TGG TG	CAGCGTGATCTGCGACCT	18	GATCAGCGAACTGCTCACG	19
1028129	1028129–1028145	GGT GGT GGT GGT GGT GG	AGATGTCGACCATCCTGACC	20	AAGTAGGCGTAACCCCGTTC	20
2955362	2955362–2955378	TGC TGC TGC TGC TGC TG	GACAAGTTCGGGTTGACCAC	20	AGTTCCTCGACCCAGTCGT	19
3558075	3558075–3558090	GCC GCC GCC GCC GCC G	CTGGAACGTGTCCGAATTG	19	GTATTCGGTGCGGATCTCCT	20
1653414	1653414–1653429	CGC CGC CGC CGC CGC C	TGAGCAGGAACCAGATCTCC	20	GTGGGGTGGATGAGTACGAC	20
1651920	1651920–1651935	CGC CGC CGC CGC CGC C	AGCATCTTGAGCCCACATCT	20	CCGAAATCAATTCTGGTCGT	20
3182509	3182509–3182523	GCG GCG GCG GCG GCG	CGGTCAGGTCGCAGATTT	18	GGTCAGCGAGAAACCACTTG	20
3562415	3562415–3562429	CTG CTG CTG CTG CTG	CTGGCATTGGGAATGTTCTT	20	CAGCACCATGTAGCCGATCT	20
212174	212174–212188	CGA CGA CGA CGA CGA	CAGAGCGGACTGCATTGAG	19	TCGGTGTTGTCCGGATTC	18
1737056	1737056–1737070	GCG GCG GCG GCG GCG	GCTCGTTGCAGGTCAGGTAG	20	GGCATGATCACCGAAAGC	18
1536798	1536798–1536812	CCG CCG CCG CCG CCG	CTGGAGTGGAAGAGCAGTCC	20	GCTGCGTTACCTCAACACC	19

[a] The position in the genome is the coordinate of the SSR locus in the *M. paratuberculosis* strain K10 genome (GenBank accession number AE016958).

under UV light with an Eagle Eye II gel documentation system (Stratagene, La Jolla, Calif.). The PCR amplicons were then sequenced with an ABI 3100 automated fluorescent DNA sequencer (Perkin-Elmer) at the University of Minnesota's Advanced Genetic Analysis Center (www.agac.umn.edu).

MLSSR data analysis. The sequences of each SSR locus of 33 isolates were aligned, and the numbers of tandem repeats were identified by use of the MegAlign program (DNASTAR Inc., Madison, Wis.). The nucleotide sequences of 11 polymorphic SSR loci were analyzed for each isolate, and allele numbers were assigned to reflect the number of copies or the number of nucleotide substitutions represented in the SSR sequence for each locus. Statistical analysis for genetic diversity and overall relationships among the isolates was performed with the computer programs ETDIV and ETCLUS, which were modified for use with the SSR data (2). MLSSR types were then assigned on the basis of the unique combination of alleles for each locus. Genetic diversity (*D*) was calculated by using the following equation: $1 - \Sigma(\text{allele frequency})^2$ (22, 29). The unweighted pair group method with arithmetic averages-based cluster analysis and bootstrap analysis with 1,000 replications were performed with the program PAUP (version 4.0; Sinauer Associates, Inc. Sunderland, Mass.), and the index of discrimination (*D*) was determined as described previously (13).

RESULTS

SSRs in *M. paratuberculosis* genome. Analysis of the whole-genome sequence of *M. paratuberculosis* strain K10 (4.83 Mbp) identified 185 SSRs consisting of three or fewer nucleotides per repeat unit. Of these, 78 mono-, di-, and trinucleotide repeats with perfect matches between adjacent copies were identified and were included as candidate polymorphic loci for further analysis (Table 2). These 78 SSR loci were also selected for inclusion in our analysis because they were short (1 to 3 bp), as is common in prokaryotes, and each locus had at least five copies. Dinucleotide repeats were the most frequently identified SSRs in the *M. paratuberculosis* genome and were present at 63 distinct loci, with the copy numbers varying between 5 and 5.5 per repeat. Mono- and trinucleotide repeats were represented at 2 and 13 loci, respectively.

MLSSR analysis revealed that 11 of the 78 loci were polymorphic in the six isolates examined. The ORFs or genes flanking each locus were also identified (Table 3). For example, locus 2 is located within ORF 210_MAP.128, which is unique to *M. paratuberculosis*. Locus 3 was identified in an intergenic region between two ORFs: a 5′ ORF encoding 6-aminohexanoate-cyclic dimer and a 3′ ORF encoding alpha/beta-hydrolase (Table 3). The functional consequences of the presence of the loci and the influence of the locus copy number on the expression of the adjacent genes deserve further investigation.

MLSSR. The 11 polymorphic SSR loci identified in the preliminary screening were characterized in 27 additional *M. paratuberculosis* isolates that were previously characterized by MPIL and AFLP analyses (20). The analysis identified 20 MLSSR types among the 33 *M. paratuberculosis* isolates recovered from different host species and geographic areas (Tables 1 and 4). The *D* value for each SSR locus was calculated on the basis of the allele frequency and the number of alleles and revealed an average number of alleles per locus of 3.20, with an average *D* value of 0.393 and a range of *D* values of 0.100 to 0.700 (21, 28) (Table 3). While the allelic variation observed in this study focused on the number of copies of the SSRs (Fig. 1A), it is noteworthy that some loci also revealed one or two base substitutions in some isolates (Fig. 1B). For instance, the analysis revealed a single polymorphic site each at SSR loci 4 and 10 and four and five polymorphisms at loci 5 and 9, respectively (Fig. 2). It is interesting that the vast majority of the nucleotide substitutions were found in MAP-06, an isolate recovered from a sheep.

Genetic relationships among *M. paratuberculosis* isolates based on MLSSR analysis. The unweighted pair group method with arithmetic averages-based cluster analysis of *M. paratuberculosis* identified 20 distinct MLSSR types among the isolates that were grouped into two major clusters, clusters M and N (Fig. 3). Cluster M contained 87.88% (29 of 33) of the isolates in the sample, including isolates recovered from bovine, caprine, murine, deer, rabbit, and human sources. The isolates in cluster M with the most common MPIL and AFLP fingerprints, A18 and Z1 and Z2, respectively, were further

TABLE 3. SSRs used in MLSSR analysis

Marker locus	Locus name	SSR	Genome ORF[a]	No. of alleles[b]	*D* value[c]
1	1793091	GGGGGGGGGGGGGGGGGGGG	GI: 13881617 phosphatidylethanolamine-binding protein (*Mycobacterium tuberculosis* CDC 1551)	5	0.700
2	2719085	GGGGGGGGGG	No hit[d] (210_MAP.128)	3	0.616
3[e]	607784	CG CG CG CG CG C	IGI: RMMR05031 6-aminohexanoate-cyclic dimer hydrolase (*Mycobacterium marinum*) GI: 13883430 hydrolase, alpha/beta-hydrolase fold family (*M. tuberculosis* CDC1551)	2	0.189
4	3406364	GC GC GC GC GC	GI: 2791627 *fixA* (*M. tuberculosis* H37Rv)	2	0.100
5	3735342	GC GC GC GC GC	GI: 13092881 putative *S*-adenosyl-L-homocysteine hydrolase (*Mycobacterium leprae*)	4	0.363
6	4286068	GCG GCG GCG GCG GCG GC	IGI: RMMR06009 heat shock protein 70 (*M. marinum*)	2	0.395
7	4310932	CCG CCG CCG CCG CCG CC	GI: 10579910 phytoene dehydrogenase (*Halobacterium* spp.)	2	0.100
8	1028129	GGT GGT GGT GGT GGT GG	GI: 3261715 *mfd* (*M. tuberculosis* H37Rv)	4	0.668
9	2955362	TGC TGC TGC TGC TGC TG	GI: 2117199 *narG* (*M. tuberculosis* H37Rv)	4	0.553
10	3558075	GCC GCC GCC GCC GCC G	GI: 1781217 *nuoG* (*M. tuberculosis* H37Rv)	3	0.279
11	1536798	CCG CCG CCG CCG CCG	GI: 5524340 *PstA* (*Mycobacterium avium*)	4	0.363

[a] GI, NCBI gene identification number; IGI, Integrated Genomics gene identification number.
[b] The average number of alleles was 3.20.
[c] *D* was calculated by using the equation $1 - \Sigma(\text{allele frequency})^2$ (22). The average *D* value was 0.393.
[d] No hit, no nucleotide match with any sequences in the GenBank database.
[e] Locus 3 is located in an intergenic region of two ORFs (6-aminohexanoate-cyclic dimer and hydrolase [the alpha/beta-hydrolase fold]).

divided into three groups, clusters M1, M2, and M3. Cluster M1 contained one isolate (isolate MAP-06), which was recovered from a sheep and which had the A1 MPIL fingerprint. Three of the five isolates from caprine sources were assigned to cluster M2. A total of 13 unique genotypes, including a majority (10 of 15) of the bovine isolates included in this study, were represented in cluster M3. In addition, the three isolates from human sources included in the sample used in this study were also found in cluster M3. Interestingly, two isolates recovered from humans (isolates MAP-14 and 0003) were clustered into the same clade as an isolate of bovine origin (isolate 0180). Isolates that were recovered from a mouse (isolate 0012), rabbits (isolates 0237 and 0160 to 0162), a deer (isolate 0883), and soil (isolate 0560) were also grouped along with the M3 genotype.

In contrast to cluster M, which consisted of isolates recovered from a variety of animal species, all four isolates that were included in cluster N were recovered from sheep. Strains of ovine origin (four of five) also showed a distinct allelic profile compared with the profiles of strains from cattle, goats, or humans.

TABLE 4. Profiles of alleles at 11 SSR loci for 20 clones of *M. paratuberculosis*

MLSSR type	Reference isolate	No. of isolates	No. of copies of SSRs or no. of nucleotide substitutions at the following locus										
			1[a]	2[a]	3[b]	4[b]	5[b]	6[c]	7[c]	8[c]	9[c]	10[c]	11[c]
1	MAP-06	1	7	11	5	1	1	5	5	3	1	0	0
2	MAP-08	1	7	10	5	5	5	5	4	4	4	5	5
3	0033	1	7	11	5	5	5	4	5	4	4	5	5
4	0016	4	7	11	5	5	5	4	5	4	3	5	5
5	0028	2	7	10	5	5	5	4	5	4	3	5	5
6	MAP-09	1	11	11	5	5	5	5	5	5	5	5	3
7	0237	3	>14	9	5	5	5	5	5	5	5	5	5
8	0560	1	>14	11	5	5	5	5	5	5	5	5	5
9	0026	1	8	10	5	5	5	5	5	5	5	5	5
10	MAP-K10	3	>14	10	5	5	5	5	5	5	5	5	5
11	MAP-14	3	7	10	5	5	5	5	5	5	5	5	5
12	0040	1	7	9	5	5	3	5	5	5	5	5	5
13	0558	1	7	9	5	5	5	5	5	5	5	5	5
14	0161	1	8	11	5	5	5	5	5	4	5	5	5
15	0011	1	9	11	5	5	5	5	5	4	5	5	5
16	0030	1	9	11	5	5	5	5	5	5	5	5	5
17	0004	2	7	11	5	5	5	5	5	5	5	5	5
18	0014	1	7	11	5	5	5	5	5	6	5	5	5
19	MAP-11	1	>14	11	4	5	4	4	5	3	4	6	4
20	0007	3	>14	10	4	5	4	4	5	3	4	6	4

[a] The numbers of G mononucleotides present (>14, 14 or more).
[b] The numbers of copies of the dinucleotide repeat.
[c] The numbers of copies of the trinucleotide repeat.

Discriminatory power of subtyping methods. The discriminatory power (*D*) of MLSSR in comparison with those of other subtyping methods was determined as described previously (13). MPIL analysis differentiated the 27 *M. paratuberculosis* isolates for which MPIL typing information was available into 6 subtypes with a *D* value of 0.50, indicating only limited discriminatory power, while MLSSR differentiated 27 *M. paratuberculosis* into 17 subtypes with a *D* value of 0.96. In contrast, AFLP analysis differentiated the 24 *M. paratuberculosis* isolates for which AFLP typing information was available into 15 subtypes, with a *D* value of 0.92 (20). In comparison, MLSSR differentiated the 24 *M. paratuberculosis* isolates into 14 subtypes, with higher *D* value of 0.95. Overall, MLSSR distinguished 20 subtypes among the 33 isolates in the sample with a *D* value of 0.96, indicating that it has a relatively high index of discrimination (Tables 3 and 4).

DISCUSSION

SSRs have been used to type many bacterial pathogens associated with human and animal infections (32). Within the genus *Mycobacterium*, VNTR or mycobacterial interspersed repetitive units have been used for the subtype-specific differentiation of several *Mycobacterium* species (19, 26, 35). In the

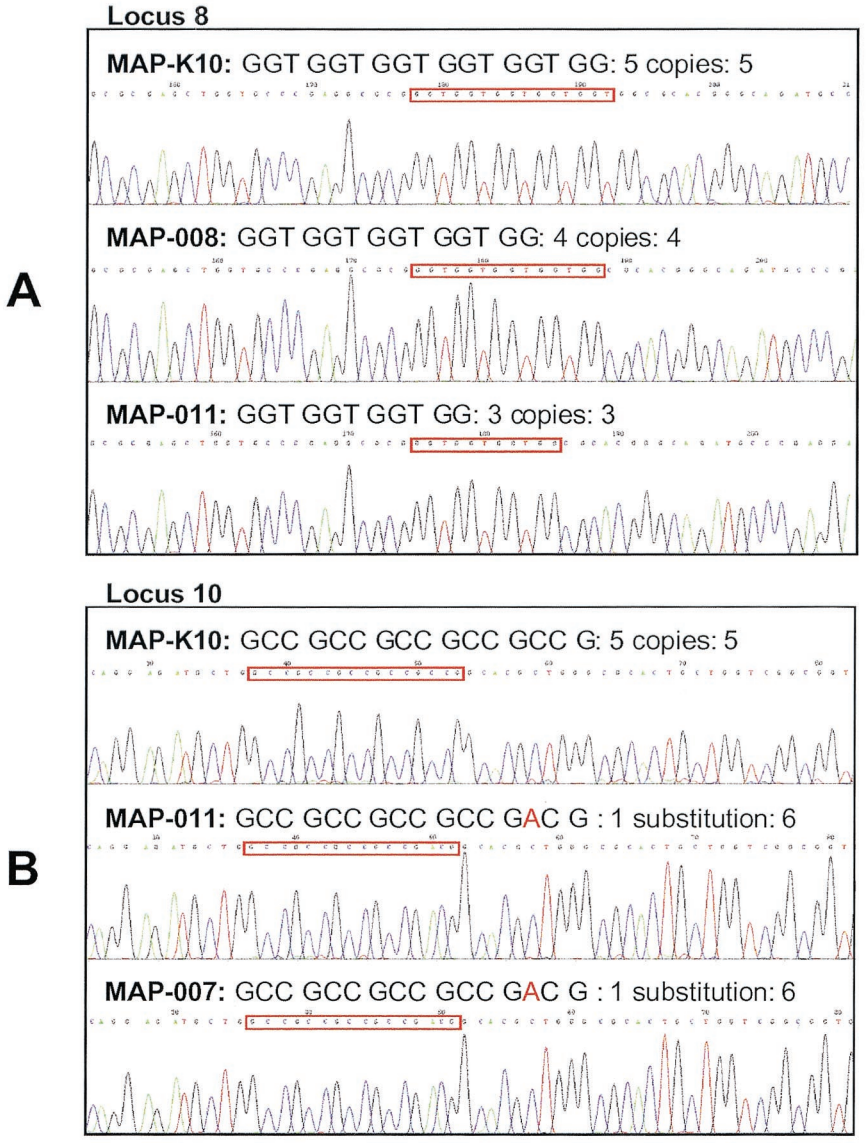


FIG. 1. Sequence analysis of two representative SSR loci. (A) Locus 8 with $(GGT)_5$ repeats. Strain MAP-K10 contains five copies of GGT, while MAP-08 and MAP-11 contain four and three copies of GGT, respectively. (B) Locus 10 with $(GCC)_5$ repeats. Strain MAP-K10 contains five copies of GCC, while MAP-11 and MAP-07 have a A-to-C substitution in a single copy of GCC.

present study we have identified polymorphic SSRs by genomic analysis of *M. paratuberculosis* and used this information to develop a highly discriminatory method for the typing of *M. paratuberculosis* isolates.

The SSRs discovered during our preliminary screening of the *M. paratuberculosis* genome were similar to the repeats that have previously been described in other bacteria, including *Haemophilus*, *Mycoplasma*, and *Mycobacterium* spp. (24, 32, 33). It has previously been recognized that regions of mono-, di-, and trinucleotide tandem repeats are often the most diverse in a bacterial genome, while complex longer repeats generally have lower levels of diversity (14). This is thought to result from slipped-strand mispairing (or replication slippage events) of the DNA polymerase that occurs with greater frequency on the SSRs, a hypothesis that remains to be tested for the SSRs that we have identified in *M. paratuberculosis* (32).

Several important attributes of a strain differentiation assay determine its utility in a clinical or epidemiologic setting. Especially for organisms such as *M. paratuberculosis* that have restricted levels of genetic diversity, the discriminatory power of an assay is a particularly important attribute. Assays such as MPIL and RFLP analysis have been shown to have only moderate abilities to differentiate among epidemiologically distinct isolates of *M. paratuberculosis* and therefore have limited applicabilities in molecular epidemiologic studies (4, 5, 33). The recently described AFLP technique has been shown to have a greater resolving power than the other two approaches but suffers from the limitation that allelic variation is indexed at

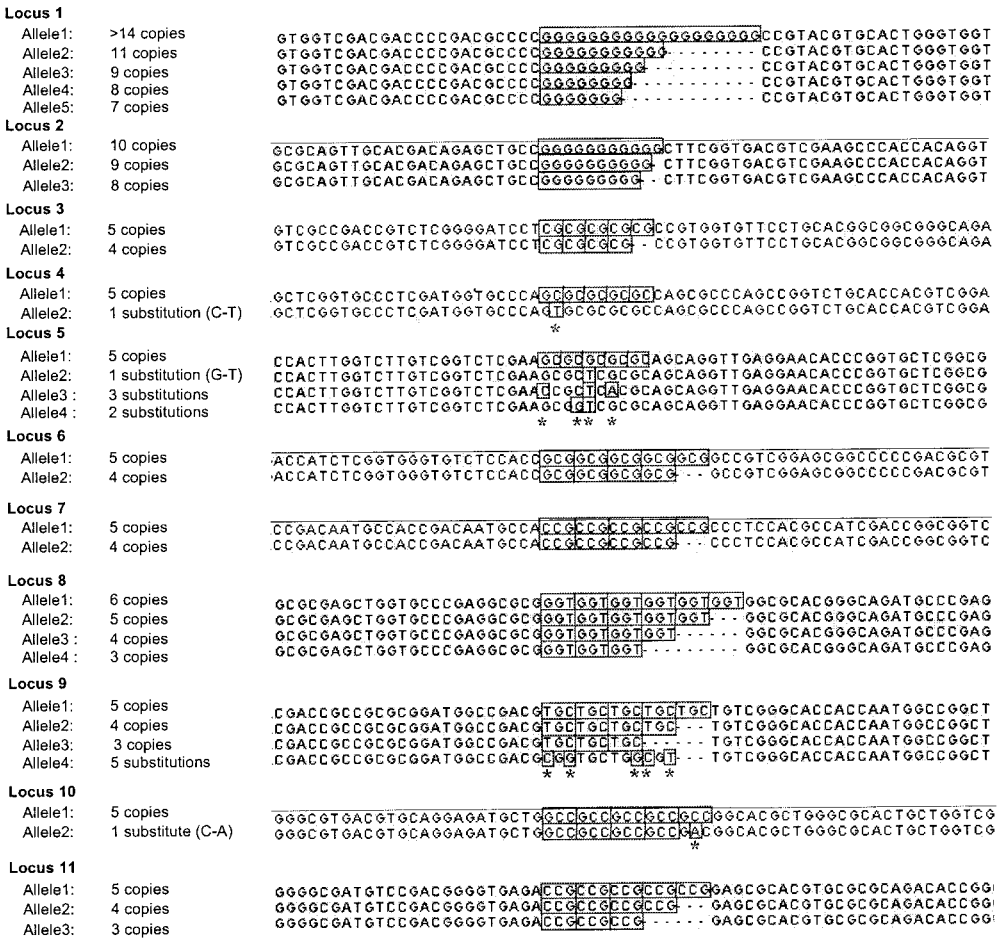
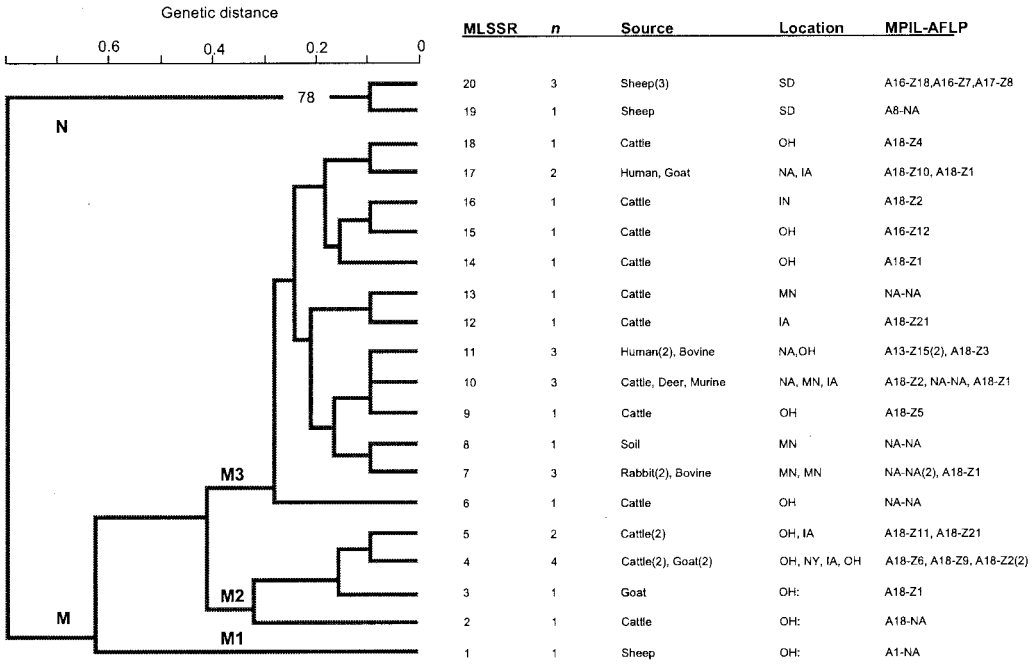


FIG. 2. Allelic variation at 11 SSR loci among 33 *M. paratuberculosis* isolates. The aligned nucleotide sequences of each of the alleles at the 11 SSR loci discovered and characterized during this investigation, along with adjacent conserved sequences, are shown. The SSRs and polymorphic sites are boxed. The locations of single-nucleotide polymorphisms are indicated with asterisks below the aligned sequences.

anonymous biallelic markers (20). In contrast, the MLSSR assay described herein is far more discriminatory, being able to differentiate 33 *M. paratuberculosis* from distinct geographic localities and host species into 20 subtypes on the basis of allelic variation at the 11 SSR loci examined, with a notably high *D* value of 0.96. Consistent with its high discriminatory power, MLSSR enabled the differentiation of seemingly monomorphic *M. paratuberculosis* strains that were indistinguishable by MPIL and AFLP analyses (20). An important advantage of the MLSSR approach is that it also indexes variations at known genetic loci and has the ability to identify multiple alleles per locus. Together, these attributes not only allow an increase in the strain-resolving power of the assay but also enable an understanding of the genetic mechanisms driving strain diversification and evolution within the species.

Another key attribute of a strain differentiation assay is its ability to identify epidemiologically and genetically related strains of a bacterial species. In this context, MLSSR analysis clearly showed that some isolates that are of sheep origin (cluster N) are genetically distinct from those of bovine, caprine, and human origin (cluster M), a finding consistent with those of previous studies (4, 6, 20). It is noteworthy, however, that the five isolates of sheep origin examined during this study were represented by three distinct MLSSR types (MLSSR types 1, 19 and 20), and four isolates clustered together in cluster N. Interestingly, all four of these phylogenetically linked *M. paratuberculosis* isolates were recovered from sheep in South Dakota, suggesting that they are both genetically and epidemiologically related and well distinguishable from the other isolates in the collection. The same isolates were also grouped into four distinct MPIL genotypes (A1, A8, A16, and A17) and three AFLP genotypes (Z7, Z8, and Z18), suggesting that they are indeed genetically distinct from the other isolates in the collection. However, by the MPIL and the AFLP approaches, these isolates do not cluster together as closely as they do by MLSSR analysis (20). Hence, these results suggest that MLSSR analysis may enable molecular epidemiologic investigations that will lead to a better understanding of strain transmission and the spread of *M. paratuberculosis* in natural populations and across host species.

In contrast to the relatively close clustering of the sheep *M. paratuberculosis* strains in the samples examined, far greater diversity was observed in isolates of bovine origin. The analysis showed that while a majority of the *M. paratuberculosis* isolates



MLSSR	*n*	Source	Location	MPIL-AFLP
20	3	Sheep(3)	SD	A16-Z18,A16-Z7,A17-Z8
19	1	Sheep	SD	A8-NA
18	1	Cattle	OH	A18-Z4
17	2	Human, Goat	NA, IA	A18-Z10, A18-Z1
16	1	Cattle	IN	A18-Z2
15	1	Cattle	OH	A16-Z12
14	1	Cattle	OH	A18-Z1
13	1	Cattle	MN	NA-NA
12	1	Cattle	IA	A18-Z21
11	3	Human(2), Bovine	NA,OH	A13-Z15(2), A18-Z3
10	3	Cattle, Deer, Murine	NA, MN, IA	A18-Z2, NA-NA, A18-Z1
9	1	Cattle	OH	A18-Z5
8	1	Soil	MN	NA-NA
7	3	Rabbit(2), Bovine	MN, MN	NA-NA(2), A18-Z1
6	1	Cattle	OH	NA-NA
5	2	Cattle(2)	OH, IA	A18-Z11, A18-Z21
4	4	Cattle(2), Goat(2)	OH, NY, IA, OH	A18-Z6, A18-Z9, A18-Z2(2)
3	1	Goat	OH:	A18-Z1
2	1	Cattle	OH:	A18-NA
1	1	Sheep	OH:	A1-NA

FIG. 3. Dendrogram depicting genetic relationships among 33 *M. paratuberculosis* isolates on the basis of the 11 SSR loci determined by MLSSR analysis. The dendrogram was generated by the unweighted pair-group method with arithmetic averages with the PAUP program. The results of the bootstrap analysis are represented as percentages and are indicated adjacent to the major nodes when the branch order was supported by >50% of the 1,000 replicate trees. Genetic distance is indicated at the top of the dendrogram. Isolate identifications, sources, geographic locations, and MILP and AFLP types are shown to the right of the dendrogram.

of bovine origin clustered together in the M3 subgroup, 60% (three of five) of the caprine isolates were represented by the closely related cluster M2, suggesting that caprine isolates bear greater genetic resemblance to cattle strains than to isolates of ovine origin, a finding that is consistent with the findings of previous studies (34). Similarly, deer and cattle strains also appeared to be more closely related to each other by MLSSR analysis, suggesting a sharing of strains of *M. paratuberculosis* in wildlife species that graze or that may otherwise come into close contact with cattle, as hypothesized previously (25).

Our studies demonstrate that MLSSR analysis offers several advantages over other methods for differentiating among *M. paratuberculosis* isolates. First, as described above, the technique has a high discriminatory power for known multiallelic genetic loci, an essential attribute for the effective differentiation of genetically distinct isolates. Second, MLSSR results are based upon DNA sequencing and, hence, are unambiguous and reproducible and can likely be obtained for most loci of all *M. paratuberculosis* isolates, even those recovered from sheep or wildlife species, as demonstrated by our studies described herein. However, we note the formal possibility that mutations or deletions at the primer sites may render some strains untypeable at some loci, such as loci 10 and 11 in MAP-06. Third, MLSSR analysis is based on the amplification of SSR loci by PCR and thus not only is rapid but also may be performed directly with bacterial colonies without DNA extraction. Fourth, due to the considerable advances in automated DNA sequencing technologies and the falling prices of DNA sequencing, the MLSSR method is amenable to adaptation for high-throughput analysis and can be performed relatively inexpensively as well. Finally, a key advantage of the approach is that the data are sequence based and, hence, enable accurate interlaboratory comparisons to be made and the information used in the development of SSR databases for further molecular epidemiologic studies, which are greatly required in this field (18). While it must be recognized that sequence errors due to strand slippage during either PCR or sequencing reactions may result in an erroneous assignment of genotype, the occurrence of such slippage errors is minimized by increasing the amount of sequence coverage at the locus (by confirming both the forward and the reverse sequences or testing duplicate samples), as is routinely practiced in our laboratory.

In conclusion, we have described here the development of MLSSR-based typing for the subtype-specific differentiation of *M. paratuberculosis* isolates. Our preliminary analyses suggest that this approach will be of considerable utility in enabling detailed molecular epidemiologic and population genetic analyses of this important animal pathogen.

ACKNOWLEDGMENTS

Research in the laboratory of V. Kapur is funded by grants from the U.S. Department of Agriculture, the National Institutes of Health, and the Minnesota Agricultural Experimental Station. Research in the laboratory of S. Sreevatsan is funded through a seed grant from the Ohio Agricultural Research and Development Center's research enhancement competitive grants program.

REFERENCES

1. **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.
2. **Amonsin, A., J. F. Wellehan, L. L. Li, P. Vandamme, C. Lindeman, M.**

Edman, R. A. Robinson, and V. Kapur. 1997. Molecular epidemiology of *Ornithobacterium rhinotracheale*. J. Clin. Microbiol. **35:**2894–2898.
3. **Benson, G.** 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **27:**573–580.
4. **Bull, T. J., J. Hermon-Taylor, I. Pavlik, F. El-Zaatari, and M. Tizard.** 2000. Characterization of IS900 loci in *Mycobacterium avium* subsp. *paratuberculosis* and development of multiplex PCR typing. Microbiology **146**(Pt 9)**:** 2185–2197.
5. **Bull, T. J., E. J. McMinn, K. Sidi-Boumedine, A. Skull, D. Durkin, P. Neild, G. Rhodes, R. Pickup, and J. Hermon-Taylor.** 2003. Detection and verification of *Mycobacterium avium* subsp. *paratuberculosis* in fresh ileocolonic mucosal biopsy specimens from individuals with and without Crohn's disease. J. Clin. Microbiol. **41:**2915–2923.
6. **Cousins, D. V., S. N. Williams, A. Hope, and G. J. Eamens.** 2000. DNA fingerprinting of Australian isolates of *Mycobacterium avium* subsp. *paratuberculosis* using IS900 RFLP. Aust. Vet. J. **78:**184–190.
7. **El-Zaatari, F. A., M. S. Osato, and D. Y. Graham.** 2001. Etiology of Crohn's disease: the role of *Mycobacterium avium paratuberculosis*. Trends Mol. Med. **7:**247–252.
8. **Francois, B., R. Krishnamoorthy, and J. Elion.** 1997. Comparative study of *Mycobacterium paratuberculosis* strains isolated from Crohn's disease and Johne's disease using restriction fragment length polymorphism and arbitrarily primed polymerase chain reaction. Epidemiol. Infect. **118:**227–233.
9. **Gascoyne-Binzi, D. M., R. E. Barlow, R. Frothingham, G. Robinson, T. A. Collyns, R. Gelletlie, and P. M. Hawkey.** 2001. Rapid identification of laboratory contamination with *Mycobacterium tuberculosis* using variable number tandem repeat analysis. J. Clin. Microbiol. **39:**69–74.
10. **Grimes, D. S.** 2003. *Mycobacterium avium* subspecies *paratuberculosis* as a cause of Crohn's disease. Gut **52:**155.
11. **Harris, N. B., and R. G. Barletta.** 2001. *Mycobacterium avium* subsp. *paratuberculosis* in veterinary medicine. Clin. Microbiol. Rev. **14:**489–512.
12. **Henderson, S. T., and T. D. Petes.** 1992. Instability of simple sequence DNA in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **12:**2749–2757.
13. **Hunter, P. R., and M. A. Gaston.** 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J. Clin. Microbiol. **26:**2465–2466.
14. **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.
15. **Kim, W., Y. P. Hong, J. H. Yoo, W. B. Lee, C. S. Choi, and S. I. Chung.** 2002. Genetic relationships of *Bacillus anthracis* and closely related species based on variable-number tandem repeat analysis and BOX-PCR genomic fingerprinting. FEMS Microbiol. Lett. **207:**21–27.
16. **Kremer, K., D. van Soolingen, R. Frothingham, W. H. Haas, P. W. Hermans, C. Martin, P. Palittapongarnpim, B. B. Plikaytis, L. W. Riley, M. A. Yakrus, J. M. Musser, and J. D. van Embden.** 1999. Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. J. Clin. Microbiol. **37:**2607–2618.
17. **Lindstedt, B. A., E. Heir, E. Gjernes, and G. Kapperud.** 2003. DNA fingerprinting of *Salmonella enterica* subsp. *enterica* serovar Typhimurium with emphasis on phage type DT104 based on variable number of tandem repeat loci. J. Clin. Microbiol. **41:**1469–1479.
18. **Maiden, M. C., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. USA **95:**3140–3145.
19. **Mazars, E., S. Lesjean, A. L. Banuls, M. Gilbert, V. Vincent, B. Gicquel, M. Tibayrenc, C. Locht, and P. Supply.** 2001. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. Proc. Natl. Acad. Sci. USA **98:**1901–1906.
20. **Motiwala, A. S., M. Strother, A. Amonsin, B. Byrum, S. A. Naser, J. R. Stabel, W. P. Shulaw, J. P. Bannantine, V. Kapur, and S. Sreevatsan.** 2003. Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis*: evidence for limited strain diversity, strain sharing, and identification of unique targets for diagnosis. J. Clin. Microbiol. **41:**2015–2026.
21. **National Animal Health Monitoring System.** 1997. Johne's disease on U. S. dairy operations. Report N245.1087. USDA, APHIS, VS, CEAH, National Animal Health Monitoring System, Fort Collins, Colo.
22. **Nei, M.** 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA **70:**3321–3323.
23. **Pavlik, I., A. Horvathova, L. Dvorska, J. Bartl, P. Svastova, R. du Maine, and I. Rychlik.** 1999. Standardisation of restriction fragment length polymorphism analysis for *Mycobacterium* avium subspecies *paratuberculosis*. J. Microbiol. Methods **38:**155–167.
24. **Peterson, S. N., C. C. Bailey, J. S. Jensen, M. B. Borre, E. S. King, K. F. Bott, and C. A. Hutchison III.** 1995. Characterization of repetitive DNA in the *Mycoplasma genitalium* genome: possible role in the generation of antigenic variation. Proc. Natl. Acad. Sci. USA **92:**11829–11833.
25. **Riemann, H., M. R. Zaman, R. Ruppanner, O. Aalund, J. B. Jorgensen, H. Worsaae, and D. Behymer.** 1979. Paratuberculosis in cattle and free-living exotic deer. J. Am. Vet. Med. Assoc. **174:**841–843.
26. **Roring, S., A. Scott, D. Brittain, I. Walker, G. Hewinson, S. Neill, and R. Skuce.** 2002. Development of variable-number tandem repeat typing of *Mycobacterium bovis*: comparison of results with those obtained by using existing exact tandem repeats and spoligotyping. J. Clin. Microbiol. **40:**2126–2133.
27. **Rozen, S., and H. Skaletsky.** 2000. Primer 3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132:**365–386.
28. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
29. **Selander, R. K., D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour, and T. S. Whittam.** 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl. Environ. Microbiol. **51:**873–884.
30. **Stabel, J. R.** 1998. Johne's disease: a hidden threat. J. Dairy Sci. **81:**283–288.
31. **Strand, M., T. A. Prolla, R. M. Liskay, and T. D. Petes.** 1993. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature **365:**274–276.
32. **van Belkum, A., S. Scherer, L. van Alphen, and H. Verbrugh.** 1998. Short-sequence DNA repeats in prokaryotic genomes. Microbiol. Mol. Biol. Rev. **62:**275–293.
33. **van Belkum, A., S. Scherer, W. van Leeuwen, D. Willemse, L. van Alphen, and H. Verbrugh.** 1997. Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. Infect. Immun. **65:**5017–5027.
34. **Whittington, R. J., A. F. Hope, D. J. Marshall, C. A. Taragel, and I. Marsh.** 2000. Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis*: IS*900* restriction fragment length polymorphism and IS*1311* polymorphism analyses of isolates from animals and a human in Australia. J. Clin. Microbiol. **38:**3240–3248.
35. **Wiid, I. J., C. Werely, N. Beyers, P. Donald, and P. D. van Helden.** 1994. Oligonucleotide $(GTG)_5$ as a marker for *Mycobacterium tuberculosis* strain identification. J. Clin. Microbiol. **32:**1318–1321.