6-2008

Comparison of DNA Pyrosequencing with Alternative Methods for Identification of Mycobacteria

Loree C. Heller
Old Dominion University, lheller@odu.edu

Michael Jones

Ray H. Widen

Original Publication Citation

Heller, L. C., Jones, M., & Widen, R. H. (2008). Comparison of DNA pyrosequencing with alternative methods for identification of mycobacteria. *Journal of Clinical Microbiology, 46*(6), 2092-2094. doi:10.1128/jcm.02001-07

JOURNAL OF CLINICAL MICROBIOLOGY, June 2008, p. 2092–2094 Vol. 46, No. 6
0095-1137/08/$08.00+0 doi:10.1128/JCM.02001-07

Comparison of DNA Pyrosequencing with Alternative Methods for Identification of Mycobacteria▽

Loree C. Heller,[1]* Michael Jones,[2] and Ray H. Widen[1]

Esoteric Testing, Department of Pathology, Tampa General Hospital, P.O. Box 1289, Tampa, Florida 33601,[1] and Microbiology Laboratory, Sacred Heart Hospital, 5151 N. 9th Avenue, Pensacola, Florida 32504[2]

Received 11 October 2007/Returned for modification 27 December 2007/Accepted 9 April 2008

Identification of mycobacterial clinical isolates by pyrosequencing within the hypervariable A region of the 16S rRNA gene was compared to other identification methods. For >90% of isolates, these identifications correlated to the level of complex or species. For identification of many mycobacteria, pyrosequencing offers an inexpensive alternative to traditional sequencing.

The number of clinically relevant species of nontuberculous mycobacteria is increasing steadily (3, 9). In addition to classic culture-based methods, identification techniques include mycolic acid analysis (gas-liquid chromatography, thin-layer chromatography, and high-performance liquid chromatography [HPLC]) and molecular methods such as genetic probes or sequencing of selected regions of the genome. 16S rRNA gene sequencing is widely used for speciation (2, 5). Restriction fragment length polymorphism (RFLP) analysis, particularly of the *hsp65* gene, may also be used (8).

In pyrosequencing, sequencing is performed by DNA synthesis. An enzyme cascade that produces visible light detects the pyrophosphate released during the reaction (7). Resulting sequences of approximately 30 bases are then compared to public or private databases. For this study, isolates were identified by the pyrosequencing of a region within the hypervariable A region of the 16S rRNA gene (10). DNA was extracted from isolates by use of the UltraClean microbial DNA kit (Mo Bio Laboratories, Carlsbad, CA) per the manufacturer's instructions. PCR amplification was performed as previously described (10) with the substitution of Sensimix PCR reagents (Bioline, Randolph, MA). Pyrosequencing was performed using Pyro Gold reagents on a Pyromark vacuum prep workstation and a Pyromark ID instrument per the manufacturer's instructions (Biotage, Uppsala, Sweden). Acceptable sequences were compared to NCBI GenBank sequences by use of BLAST analysis (1) and to Ribosomal Database Project (RDP II) sequences by use of the sequence match tool (4). Only 100% matches were considered. Species names not included in the DMSZ bacterial nomenclature up-to-date database (http://www.dsmz.de) were excluded. Species not known to be clinically relevant were also excluded. The most likely identification was considered for comparisons.

Deindentified clinical isolates were acquired as residual pathology samples used in institutional review board-approved protocols from Tampa General and Sacred Heart Hospitals. By use of ATCC and clinical isolates, the identifications obtained by pyrosequencing were compared to State of Florida Department of Health (DOH) or HPLC identification. Nine of 10 ATCC (Manassas, VA) isolates tested, including *Mycobacterium fortuitum* (ATCC 6841), *Mycobacterium intracellulare (*ATCC 13950), *Mycobacterium marinum* (ATCC 927), *Mycobacterium avium* (ATCC 25291), *Mycobacterium chelonae* (ATCC 19235, 14472, and 35752), *M. tuberculosis* (ATCC 27294), and *Mycobacterium abscessus* (ATCC 19977), were identified correctly to the species or complex level by 16S rRNA gene pyrosequencing. However, using pyrosequencing, multiple identifications were possible with *Mycobacterium kansasii* (ATCC 12478).

The DOH performs PCR-RFLP analysis of the *hsp65* gene, DNA probes for the identification of *M. tuberculosis* complex, *M. avium* complex, *M. kansasii*, and *Mycobacterium gordonae*, or HPLC to identify isolates (www.doh.state.fl.us). For 144 of 154 (93.5%) clinical isolates, pyrosequencing and DOH identifications correlated to the level of complex or species (Table 1). As discussed in the work of Tuohy et al. (10), *M. gordonae* isolates consistently produced sequences as short as 16 bases with pyrosequencing. However, these shorter sequences were sufficient for species-level identification. Additionally, the species *M. kansasii*, *M. simiae*, *Mycobacterium parascrofulaceum*, and *Mycobacterium scrofulaceum* were indistinguishable by pyrosequencing. Samples of these four species were subjected to pyrosequencing with the forward primer as suggested (10). Although this additional sequence information also resulted in multiple possible identifications, the final identification could sometimes be determined by process of elimination. For example, for the single *M. simiae* isolate tested, BLAST analysis with the sequencing primer resulted in six possible species identifications. With the forward primer, 12 identifications were possible. Only *M. simiae* was common to both groups. However, for the single *M. scrofulaceum* isolate and for two *M. kansasii* isolates, an identification of *M. parascrofulaceum* was also possible. Therefore, an alternative method such as traditional sequencing should be considered as a reflex test for these four species.

Traditional sequencing was used to characterize discordants and strains with multiple possible identifications (Table 2). PCR amplification and traditional sequence analysis were performed per the manufacturer's instructions (Microseq 500 16S

* Corresponding author. Mailing address: Esoteric Testing, Department of Pathology, Tampa General Hospital, P.O. Box 1289, Tampa, FL 33601. Phone: (813) 844-4677. Fax: (813) 844-1312. E-mail: lheller@tgh.org.

▽ Published ahead of print on 16 April 2008.

TABLE 1. Clinical isolates identified by pyrosequencing correlating to DOH identification (n = 144)

DOH identification	Pyrosequencing identification	No. of isolates
M. abscessus or *M. abscessus-M. chelonae* complex	*M. abscessus-M. chelonae* complex	21
M. avium complex	*M. avium-M. intracellulare* complex	52
M. fortuitum	*M. fortuitum* complex	13
M. goodii	*M. goodii*	1
M. gordonae	*M. gordonae*	22
M. malmoense	*M. malmoense*	1
M. marinum	*M. marinum*	1
M. mucogenicum	*M. mucogenicum*	1
M. peregrinum	*M. fortuitum* complex	7
M. szulgai	*M. szulgai*	4
M. tuberculosis complex	*M. tuberculosis* complex	21

rRNA gene bacterial identification kit; Applied Biosystems, Foster City, CA) and electrophoresis was done with an ABI 3130 genetic analyzer. Sequences were compared to GenBank sequences as previously described and to RIDOM Project database sequences (6). An isolate identified as *M. gordonae* by the DOH was identified as *Mycobacterium xenopi* or *Mycobacterium heckeshornense* by pyrosequencing (Table 2). The isolate was categorized as *M. heckeshornense* by traditional 16S rRNA gene sequencing. A second *M. gordonae* isolate was identified as *M. interjectum* by pyrosequencing. Attempts at traditional sequencing were unsuccessful, indicating that the DNA may have been of low quality. Sequencing of isolates identified as *M. avium-M. intracellulare* complex and *M. interjectum* resulted in several possible identifications. Isolates 3 and 13 may have been mixed during processing; these isolates remain discordant.

Pyrosequencing was also compared to HPLC identification performed using the MIDI Sherlock identification system (Newark, DE) using the Agilent 1200 LC system with Chemstation Base software (Agilent Technologies, Santa Clara, CA) per the manufacturer's instructions. Nineteen isolates were repeated from the comparison above. This group did not include any discordant isolates. For 69 of 71 (97%) clinical isolates, the identifications correlated to the level of complex or species (Table 3). One discordant isolate was identified by HPLC as *M. nonchromogenicum* and by pyrosequencing as *M. mucogenicum*. The DOH determined this isolate to be *M. nonchromogenicum*. In this instance, pyrosequencing did not appropriately identify the isolate. A second isolate was identified as *M. fortuitum* by HPLC, *M. kansasii* by pyrosequencing, and *M. scrofulaceum* by the DOH. This isolate most likely belongs to the group discussed above that is not completely identified by pyrosequencing of the 16S hypervariable region.

The techniques compared in this study are described in (Table 4). The techniques vary in advantages and disadvantages. The only nonmolecular identification method tested, HPLC for mycolic acid content (MIDI Sherlock mycobacterium identification system), has been FDA cleared for *M. tuberculosis* identification but can be used to identify other species. The Accuprobe culture identification test has the benefit of FDA approval but only for a limited number of species.

TABLE 2. Isolates discordant for species or complex[a]

Isolate	DOH identification	Pyrosequencing identification by:		Microsequencing identification (BLAST/RIDOM)
		BLAST	RDP	
ATCC 12478	*M. kansasii*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. fortuitum*, *M. conspicuum*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. conspicuum*	*M. kansasii*
3	*M. szulgai*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. fortuitum*, *M. conspicuum*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. conspicuum*	*M. kansasii*
13	*M. kansasii*	*M. szulgai*, *M. heidelbergense*, *M. intracellulare*	*M. szulgai*	*M. szulgai*
26	*M. simiae*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. fortuitum*, *M. conspicuum*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. conspicuum*	*M. simiae*
78	*M. gordonae*	*M. interjectum*	*M. interjectum*	ND[b]
82	*M. scrofulaceum*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. fortuitum*, *M. conspicuum*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. conspicuum*	99.75% *M. parascrofulaceum*[c]
86	*M. avium-M. intracellulare*	*M. celatum*, *M. malmoense*	*Mycobacteria* sp.	98.86% *M. simiae-M. szulgai-M. interjectum*
99	*M. gordonae*	*M. xenopi*, *M. heckeshornense*	*M. xenopi*, *M. heckeshornense*	*M. heckeshornense*
213	*M. kansasii*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. fortuitum*, *M. conspicuum*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. conspicuum*	*M. kansasii*
227	*M. kansasii*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. fortuitum*, *M. conspicuum*	*M. simiae*, *M. parascrofulaceum*, *M. kansasii*, *M. scrofulaceum*, *M. gastri*, *M. conspicuum*	*M. kansasii*
240	*M. interjectum*	*M. celatum*, *M. malmoense*	*Mycobacteria* sp.	*M. celatum* or 99.07% *M. malmoense-M. simiae*

[a] Sequence identity 100% unless otherwise specified.
[b] ND, not determined.
[c] *M. parascrofulaceum* not included in RIDOM database.

TABLE 3. Clinical isolates identified by pyrosequencing correlating to HPLC identification ($n = 69$)

HPLC identification	Pyrosequencing identification	No. of isolates
MAC[a]	*M. avium-M. intracellulare* complex	37
M. abscessus-M. chelonae or *M. chelonae-M. abscessus*	*M. abscessus-M. chelonae* complex	6
M. fortuitum-M. peregrinum	*M. fortuitum* complex	11
M. gordonae I or II	*M. gordonae*	4
M. malmoense	*M. malmoense*	1
M. marinum	*M. marinum*	2
M. nonchromogenicum-M. terrae or *M. terrae-M. nonchromogenicum*	*M. terrae*	2
M. tuberculosis complex	*M. tuberculosis* complex	5
M. xenopi I or II	*M. xenopi*	1

[a] MAC, *M. avium* complex.

TABLE 4. Comparison of mycobacterial identification methods

Method	No. of species/ complexes identified	No. of strains tested	Target and sensitivity	Overall time
Pyrosequencing[a]	Varies[f]	216	10 ng DNA	~4.75 h
Microsequencing[b]	Varies[f]	9	25 ng DNA	~6.5 h
HPLC[c]	25	71	Mycolic acids from very few cells	2.25 h
DNA probes[d]	1	—[g]	RNA from a single colony	50 min
hsp65 RFLP[e]	25	—[g]	Unspecified quantity of DNA	~6.5 h

[a] PCR optimization for pyrosequencing, Biotage, Uppsala, Sweden.
[b] MicroSeq 500 16S rRNA gene bacterial identification kit, Applied Biosystems, Foster City, CA.
[c] MIDI Sherlock identification system (Newark, DE).
[d] Accuprobe culture identification test, Gen-Probe Inc., San Diego, CA.
[e] Telenti et al. (8).
[f] Varies as database content varies.
[g] —, 154 for DNA probes, HPLC, and *hsp65* RFLP in combination.

Isolates must be tested individually. The PCR-RFLP assay requires traditional gel electrophoresis and manual interpretation and is therefore labor-intensive. Pyrosequencing and more traditional Sanger sequencing (microsequencing) are similar assays, although pyrosequencing allows for high-throughput testing, while multiple samples take much longer with traditional sequencing. Interpretation of either assay is performed by comparison to sequence databases, which vary in quality.

This study confirms the accuracy of a previously described method, pyrosequencing, for the identification of mycobacterial species (10) and agrees that this method appropriately identifies >90% of clinical isolates. In the present study, additional species were tested, including *M. goodii*, *M. malmoense*, *M. peregrinum*, *M. szulgai*, and clinical isolates of *M. marinum*. These additional species were correctly identified. Pyrosequencing also compared favorably to identification by HPLC and by classic 16S rRNA sequencing. As noted previously, use of the forward primer as sequencing primer did not misidentify *M. kansasii* or *M. scrofulaceum*. However, multiple identifications were possible.

In conclusion, more than 90% of isolates, including *M. tuberculosis* complex as well as nontuberculous mycobacteria commonly found in the clinical laboratory, were correctly identified to the complex or species level by pyrosequencing with a single primer. Since shorter sequences are generated, this method is not as discriminating as more-traditional sequencing. As with other sequencing identification techniques, the principal limitation is the quality of the available databases. When GenBank is used as a database, some ambiguity on identification may exist due to the free-access nature of the data submitted (5, 11). The RIDOM database in particular censors this information (6). The use of public databases or the Identifire software (Biotage, Uppsala, Sweden) does allow the addition of new species. Although no single test methodology can provide 100% accurate results (9), less than 5% (9 of 206) of isolates tested were truly discordant between pyrosequencing and alternative identification methods. Pyrosequencing has promise as a method for the identification of mycobacteria in the clinical laboratory. It is reasonably inexpensive but it is a technique that requires molecular expertise. Due to the ambiguity in identifications of a few species, an additional identification method will be necessary for specific isolates, perhaps until alternative sequencing targets are described.

We appreciate the technical help of Jeaninne Huffman, Brian Vance, and the Tampa General Microbiology Laboratory. The helpful conversations of Marion Tuohy and Cathy Accurso are also appreciated.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Brenner, D. J., N. R. Krieg, J. T. Staley, and G. M. Garrity.** 2005. Bergey's manual of systematic bacteriology. Springer US, New York, NY.
3. **Brown-Elliott, B. A., D. E. Griffith, and R. J. Wallace, Jr.** 2002. Newly described or emerging human species of nontuberculous mycobacteria. Infect. Dis. Clin. N. Am. **16:**187–220.
4. **Cole, J. R., B. Chai, R. J. Farris, Q. Wang, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, A. M. Bandela, E. Cardenas, G. M. Garrity, and J. M. Tiedje.** 2007. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res. **35:**D169–D172.
5. **Dostal, S., E. Richter, and D. Harmsen.** 2003. Concise guide to mycobacteria and their molecular differentiation. Ridom Press, Wurzburg, Germany.
6. **Harmsen, D., J. Rothganger, C. Singer, J. Albert, and M. Frosch.** 1999. Intuitive hypertext-based molecular identification of micro-organisms. Lancet **353:**291.
7. **Ronaghi, M.** 2001. Pyrosequencing sheds light on DNA sequencing. Genome Res. **11:**3–11.
8. **Telenti, A., F. Marchesi, M. Balz, F. Bally, E. C. Bottger, and T. Bodmer.** 1993. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. J. Clin. Microbiol. **31:**175–178.
9. **Tortoli, E.** 2003. Impact of genotypic studies on mycobacterial taxonomy: the new mycobacteria of the 1990s. Clin. Microbiol. Rev. **16:**319–354.
10. **Tuohy, M. J., G. S. Hall, M. Sholtis, and G. W. Procop.** 2005. Pyrosequencing as a tool for the identification of common isolates of Mycobacterium sp. Diagn. Microbiol. Infect. Dis. **51:**245–250.
11. **Turenne, C. Y., L. Tschetter, J. Wolfe, and A. Kabani.** 2001. Necessity of quality-controlled 16S rRNA gene sequence databases: identifying nontuberculous *Mycobacterium* species. J. Clin. Microbiol. **39:**3637–3648.