Species spectrum of nontuberculous mycobacteria isolated from suspected tuberculosis patients, identification by multi locus sequence analysis

Abdolrazagh Hashemi-Shahraki a, Saeed Zaker Bostanabad b,⇑,1, Parvin Heidarieh c, Leonid Petrovich Titov d, Azar Dokht Khosravi a, Nasrin Sheikhi e, Mostafa Ghalami e, Seyed Ali Nojoumi f

a Infectious and Tropical Disease Research Center, Ahvaz Jundishapur University of Medical Science, Ahvaz, Iran
b Microbiology Department, Islamic Azad University-Parand Branch, Tehran, Iran
c Bacteriology and Virology Department, Alborz University of Medical Science, Karaj, Iran
d The Belarusian Scientific-Research Institute of Epidemiology and Microbiology, Minsk, Belarus
e Microbiology Unit, Masoud Laboratory, Tehran, Iran
f Pasteur Institute of Iran, Tehran, Iran

ARTICLE INFO

Article history:
Received 30 January 2013
Received in revised form 25 August 2013
Accepted 28 August 2013
Available online 24 September 2013

Keywords:
Mycobacterium
Infection
Identification
Multi locus sequence analysis (MLSA)

ABSTRACT

Identification of Mycobacterium species is difficult due to a complex and rapidly changing taxonomy, the failure of 16S rRNA to discriminate many closely related species and the unreliability of phenotypic testing. We investigated a collection of nontuberculous mycobacteria (NTM) strains isolated from suspected tuberculosis patients at Tuberculosis Reference Centre (Ahvaz, Iran) and Masoud Laboratory (Tehran, Iran) during 2008–2012 to evaluate the species spectrum of NTM isolates.

Based on phenotypic tests, the isolates were identified up to species or complex level; however they were heterogonous by hsp65-PCR restriction fragment length polymorphism analysis (PRA) method. Representative isolates from each hsp65-PRA pattern, were subjected to identification using single locus and multi locus sequence analysis (MLSA) based on 16S rRNA, rpoB, hsp65 and 16S–23S internal transcribes spacer (ITS) fragments to determine their taxonomic affiliations.

All 92 NTM isolates from different clinical specimens were considered as etiological agents causing disease according to American Thoracic Society (ATS) guideline. Phenotypic evaluation alone assigned 66 (72%) isolates to a species or complex level and consequently 76 (82%) isolates showed previously reported hsp65-PRA patterns. Although sequence base identification using single locus such as 16S rRNA, rpoB, hsp65 or ITS identified the isolates up to species level, MLSA correctly identified 16 different species of NTM from clinical isolates. In summary, four-locus MLSA is a reliable method for elucidating taxonomic data and reliable species identification of Mycobacterium isolates and therefore, would be more feasible for routine use in Tuberculosis (TB) reference laboratory.

1. Introduction

Members of the genus Mycobacterium are aerobic and Acid Fast Bacilli (AFB) which include more than 150 different species (http://www.bacterio.cict.fr/m/mycobacterium.html). The genus also comprises both the strictly pathogenic species including M. tuberculosis as well as Mycobacterium leprae and saprophytic species or nontuberculous mycobacteria (NTM). While many NTM species are considered to be nonpathogenic, an increasing number is being reported as opportunistic pathogen (Brown-Elliott and Wallace, 2002; Griffith et al., 2007; Katoch, 2004).

Although in developing countries particularly Iran, tuberculosis remains one of the major public health concerns (Bostanabad et al., 2011; Hashemi et al., 2012; Velayati et al., 2009), infections by NTM has rapidly increased over the last years due to advances in isolation and subsequently identification techniques (Shojaei et al., 2011).

According to the American Thoracic Society (ATS) guideline, clinically isolated NTM should be identified to the species level to determine their clinical significance, infection control, epidemiological analysis and patient management (Griffith et al., 2007).

Traditional identification methods including the phenotypic tests have been the most commonly used methods for the species identification in reference laboratories which are cumbersome,
time-consuming, often difficult to standardize and not definitive (Kent and Kubica, 1985; Springer et al., 1996; Tortoli, 2003). PCR restriction fragment length polymorphism analysis (PRA) of digestion products of specific genes, such as hsp65 (Telenti et al., 1993) 16S rRNA (Domenech et al., 1994), rpoB (Lee et al., 2000) and 16S–23S rRNA internal transcribed spacer (ITS) (Roth et al., 2000) has been reported as a rapid and inexpensive identification method. Easy to perform and availability of restriction patterns of most valid species of Mycobacterium in several published algorithms (Devallois et al., 1997; Telenti et al., 1993) and in a free available database (http://app.chuv.ch/prasite) has resulted in the use of hsp65-PRA as a first choice molecular identification method in TB reference laboratory in a low-income country with a high prevalence of TB.

Other molecular methods using sequence analysis of genomic loci including the 16S rRNA gene (Rogall et al., 1990), rpoB (Adekambi et al., 2003), hsp65 (Ringuet et al., 1999), ITS (Roth et al., 1998), recA (Blackwood et al., 2000) and other genetic markers are increasingly used for a rapid and reliable identification. 16S rRNA gene sequencing as gold standard target as well as each of these loci as a single identification target fail to discriminate closely related species of Mycobacterium such as M. abscessus-M. chelonae-M. avium complex, M. farrinogenes-M. senegalense, M. kansasi-M. gastri, M. marinum-M. ulcerans, (Adekambi et al., 2003; Menendez et al., 2002; Rogall et al., 1990).

Recently, multi locus sequence analysis (MLSA) has been suggested as a method for the delineation of species and assignment of strains to define species based on the analysis of concatenated sequences of selected housekeeping genes (Achtman and Wagner, 2008; Gevers et al., 2005).

MLSA approach has been used to identify the species of mycobacteria in recent studies (Castillo-Rodal et al., 2012; Dai et al., 2011; Devulder et al., 2005; Macheras et al., 2011). However, the value of this approaches needs to be more evaluated in clinical setting as a routine identification method.

The primary aim of this study was the identification of clinical isolates of mycobacteria from suspected tuberculosis patients by using the hsp65-PRA method.

For definite and reliable identification, the species also were subjected to sequence base identification, using MLSA of a concatenation of the four genes 16S rRNA, rpoB, hsp65 and ITS. The selected genes used for MLSA of the isolates were based on their suitability for phylogenetic purposes.

2. Materials and methods

2.1. Mycobacterium strains

A total of 92 different isolates of NTM were recovered from the years 2008–2012 by using standard decontamination procedure (NaOH 4%) for non sterile specimens and direct culturing of sterile specimens on Löwenstein-Jensen (LJ) medium (10) from cases of patients, suspected to have Mycobacterium related complications at the TB Reference Centre (Ahvaz, Iran) as well as Masoud Laboratory (Tehran, Iran). The contaminated samples or the isolates not matching with the ATS criteria for definition of NTM disease were excluded from the study. Prior to use, mycobacterial strains were stored at −70 °C and were grown on LJ medium at 37 °C.

2.2. Phenotypic identification

All mycobacterial isolates were grown on LJ medium and examined for growth rate, macroscopic and microscopic morphological features, growth at different temperatures and also a battery of biochemical tests including tween 80 hydrolysis, nitrate reduction, arylsulfatase, urease production, tellurite reduction, salt tolerance, and semiquantitative catalase production according to standard procedures (Kent and Kubica, 1985).

2.3. Molecular identification

2.3.1. DNA extraction

Chromosomal DNA was extracted using lipase (Type VII; final concentration, 2 mg/ml [Sigma]) and a further treatment with proteinase K (100 pg/ml) and 0.5% sodium dodecyl sulfate. The DNA was purified by phenol chloroform-isooamyl alcohol and precipitated with isopropanol. The precipitate was washed in 70% ethanol, dehydrated and dissolved in 100 µl of Milli-Q water and stored in −20 °C freezer until use (Pitcher et al., 1989).

2.3.2. hsp65-PRA-based identification method

An approximately 441 bp fragment of the hsp65 gene was amplified by PCR using two specific primers Tb11 (5′-ACCACACATGTTGTGCCTCA-3′) and Tb12 (5′-CTTGTCGAGCCGA-TACCT-3′). Genomic DNA of M. fortuitum ATCC 49404T and double distilled water were used as positive and negative control, respectively in all PCR experiments in current study. PCR products of hsp65 were digested by the BsrEI and the HaeII (Telenti et al., 1993). The fragments were compared with those of patterns deposited in a free available database (http://app.chuv.ch/prasite) for species identification.

2.3.3. PCR and sequencing of 16S rRNA, rpoB, hsp65 and ITS

2.3.3.1. 16S rRNA. Nearly full lengths of the 16S rRNA genes (1500-bp) from isolates were amplified using primers pA (5′-AGACCTTGTTGATCCTG-3′) and pB (5′-TCACAGACGGCGAAGGA-3′) as described previously (Rogall et al., 1990). The amplified PCR products of 16S rRNA gene for each isolate were purified with the Gene JET™ Gel Extraction Kit (Fermentas, Lithuania) as described in the manufacturer’s instructions. The sequences of the products were determined using an ABI PRISM® 7700 Sequence Detection System (Applied Biosystems, Foster City, Calif.) according to the standard protocol of the supplier.

2.3.3.2. rpoB. A 750-bp fragment of the rpoB gene was amplified and sequenced by two specific primers Mycof (5′-GGAAGCTTGACCACGAAAT-3′) and MycoR (5′-GGAAGCTGGAATCACAT-3′) as previously described Adekambi et al. (2003).

2.3.3.3. hsp65. The amplified PCR products of hsp65 gene for each isolate were purified and the sequences were determined as described above by two specific primers Tb11 and Tb12 (Ringuet et al., 1999).

2.3.3.4. ITS. The universal primers 16S-1511f (5′-AAGTCGTAACAA GGTARCCG-3′) and 23S-23r (5′-TCCTGCAAGCCATCCACC-3′) were used for amplification of the ITS region (~230–350) as previously described Gutell (1994). ITS amplicons were cloned using standard procedures provided by a commercial company (Genexpress, Berlin, Germany) due to interoperon spacer sequence variations (Roth et al., 1998). Thereafter, one clone for each strain was sequenced and analyzed. Distinct ITS sequences found within the mycobacterium species were used to define infrasubspecific taxa, and such subspecies defined by a sequence were called a sequavar (De Smet et al., 1995; Frothingham and Wilson, 1994).

2.3.4. Analysis of sequence data

The obtained sequences for each isolate from different loci were aligned separately and compared with all existing relevant sequences of mycobacteria retrieved from GenBank database using jPhydit program (Jeon et al., 2005). Percentages of similarity
between sequences of each gene were determined by comparing sequences search to an in-house database of 16S rRNA, rpoB, hsp65 genes and ITS region sequences.

2.3.5. Phylogenetic analysis

For phylogenetic analyses, sequences were trimmed using jPhydit in order to start and finish at the same nucleotide position for all the strains under study. Phylogenetic trees were obtained from DNA sequences utilizing the Neighbor-Joining (NJ) method and Kimura's two parameter (K2P) distance correction model with 1000 bootstrap replications supported by the MEGA 4.1 software (Tamura et al., 2007).

To better elucidate the relatedness of the clinical isolates with valid species of Mycobacterium, the respective sequences of 16S rRNA, rpoB, hsp65 and ITS were concatenated in single, 2714 bp-long strands using the jPhydit program, followed by manual editing of the junctions. The sequences were then aligned and analyzed in the MEGA 4.1 beta program with the NJ method bootstrapped with 1000 replicates. The maximum composite likelihood method with the complete deletion option for gaps was used to calculate the evolutionary distances.

2.3.6. Nucleotide sequence accession numbers

The GenBank accession numbers of investigated clinical isolates of NTM determined in this work are as follows: JX266694 to JX266711 for the 16S rRNA, JX294377 to JX294412 for hsp65 and rpoB and JX270654 to JX270671 for ITS.

3. Results

3.1. Mycobacterium strains

During the years 2008–2012, 92 isolates of NTM recovered from 2313 cases of patients suspected of having tuberculosis at the Tuberculosis Reference Centre (Ahvaz, Iran) and Masoud Laboratory (Tehran, Iran). The strains were isolated from bronchoalveolar lavage or BAL (n = 26), sputum (n = 40), abscess (n = 6), blood (n = 6), gastric washing (n = 4), urine (n = 1), soft tissue infection (n = 6) and lymph node biopsy (n = 3). The clinical isolates which recovered from one patient, considered as the same strain according to identical hsp65-PRA profiles and then one isolate randomly selected to further identification. All patients with pulmonary disease were symptomatic. Fever, night perspiration, anorexia and weight loss were the most common symptoms in four out of six patients with disseminated NTM disease; however weight loss and mild fever were recorded for two other patients. CD4+ count were found to be less than 40/µl for all six patients with disseminated NTM disease (Supplementary online Table S1).

Long history (more than 3 weeks) of tender erythematous nodules was the main symptom of patients with soft tissue infection (n = 6) and no response to different course of treatment was their main past medical history. Six different isolates were recovered from painful abscesses with no discharge, which were untreated using routine treatment.

Fever and cervico-facial lymph nodes enlargement were the most important symptoms of patients with cervical lymphadenitis due to NTM disease. Past medical history of patient with urinary tract infection was significant for genitourinary symptoms such as frequency, urgency, pain and dysuria. Additional details about patient’s history with NTM disease can be found at Supplementary online Table S1. Most of the patients (88 cases) had past medical history of different kind of immunosuppressive syndrome such as cancer, chronic bronchitis, chronic obstructive pulmonary disease, diabetes mellitus, hemophilia, HIV and HBV infection, lymphoma, pamphilgus, transplantation, and tuberculosis indicating that patients with immunosuppressive syndrome are more susceptible to NTM disease. Fourteen cases also were healthy which emphasize on the pathogenic potential of NTM species.

All clinical samples except blood samples were positive by direct microscopy examination of Ziehl–Neelsen staining and yielded a pure culture of AFB on LJ medium. The etiologic role of the isolates might be inferred from the fact that AFB were microscopically observed in most of clinical specimens of each patient and the same mycobacterial strains confirmed by hsp65-PRA, were recovered in pure culture from different clinical samples of each patient. This finding confirmed clinical significance of the isolates according to ATS rules (Griffith et al., 2007).

3.2. Species identification by phenotypic tests

On the basis of growth characteristics (Timpe and Runyon, 1954) the isolates studied could be classified within the two groups: fifty clinical isolates were rapidly growing mycobacteria (RGM), whereas, forty-two isolates were slowly growing mycobacteria (SGM).

According to phenotypic tests, the M. fortuitum complex like group was the most frequently encountered (36 isolates), following by M. avium complex (12 isolates) as the second and M. kansasi-like strains (11 isolates) as the third. The remaining strains were mostly rare mycobacteria encompassing one to eight isolates (Table 1). Among the 92 isolates studied, phenotypic evaluation alone assigned 66 (72%) isolates to a species or complex and the remaining isolates were unidentifiable.

3.3. Species identification by molecular tests

3.3.1. hsp65-PRA-based identification

Details on molecular identification by different approaches can be found at Supplementary online Fig. S1. According to hsp65-PRA results, an identical pattern was detected for the isolated microorganisms from each patient. By using hsp65-PRA M. fortuitum was the most frequently encountered (12 isolates), followed by M. kansasi (8 isolates) as the second and M. abscessus subsp. abscessus (8 isolates) and M. avium complex (8 isolates) onward. The remaining strains were mostly rare mycobacteria encompassing one to six isolates (Table 1). Among the 92 isolates, 16 (17.3%) isolates showed unique and unknown patterns in comparison to patterns deposited in free available database (http://app.chuv.ch/prasite). The hsp65 sequences of four unknown patterns which detected by hsp65-PRA (Table 1) were analyzed for point mutation analysis and precise definition of the sizes of the PRA restriction fragments from sequences using RestrictionMapper software (version 3) http://www.restrictionmapper.org. A comparison between restriction fragment size from running agarose gels and the real size of fragments deduced from Restriction Mapper, shown that unknown patterns reported here were related to the confusion in interpretation of actual size of restriction fragments on gels.

All M. abscessus subsp. abscessus, M. abscessus subsp. bolletii, M. avium, M. branderi, M. kansasi, M. peregrinum, M. phieli, M. simiae and M. thermoresistibile (Table 1) displayed similar patterns on BstEII digestion (235/210 bp); however, all of them could be discriminated from the Haelll digestion. Four isolates also had identical hsp65-PRA patterns they were indistinguishable from M. conceptionense and M. senegalense due to the same digestion pattern for BstEII and very similar Haelll pattern.

Four isolates also had identical hsp65-PRA patterns in comparison to those of M. conceptionense or M. senegalense patterns. For better identification and reliable access to species spectrum of the clinical isolates, randomly selected isolates as representative isolates from each cluster of hsp65-PRA, were subjected to sequence based identification using MLSA (Supplementary online Fig. S1).
## Table 1

Results of NTM identification by phenotypic and genotypic tests.

<table>
<thead>
<tr>
<th>Numbers of isolates</th>
<th>Lab designation: AFP-000</th>
<th>Phenotypic characterization of clinical isolates</th>
<th>Phenotypic tests by hsp65-PRA</th>
<th>Identification by PRA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Growth at 37°C</td>
<td>Growth at 42°C</td>
<td>Growth on MacConkey agar</td>
</tr>
<tr>
<td>12</td>
<td>9, 12, 18, 47, 53, 68, 69, 73, 94, 107, 111, SMB†</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>10, 13, 31, 36, 43, 104, NM3</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>16, 22, 23, 52, 54, 63, 89, 215†</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>98, 120, 224†, NM14</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>96, 207†</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>60, 61, 119, 121, 131, 136</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>27, 30, 40, 74†</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>95, 117, 201†, NM5, NM8</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>39, 88, 93, 191†, NM1 NM13</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>77, 83, 137†, NM2</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>19, 20, NM7†</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>25, NM44†</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>NM47†, NM53</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>24, 29, 45, 67, 71, 172, 175</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>32, 50, 59, 218†, NM11, NM26</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>11</td>
<td>3, 21, 37, 38, 42, 55, 56, 58, 118, 209, 220†</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>148, 227†</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

* The isolates randomly selected from each cluster of hsp65-PRA patterns for MLSA.
Table 2
Details of identification of Iranian NTM by sequence analysis.

<table>
<thead>
<tr>
<th>AFPI</th>
<th>Percentage of Similarity (number of nucleotide differences) based on:</th>
<th>hsp65 (–413 bp)</th>
<th>ITS (–214–250 bp)</th>
<th>MLSA (–2714 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM8</td>
<td>100% M. fortuitum, 99.6% (6) M. farcigenes, 99.5% (7) M. senegalense</td>
<td>100% M. fortuitum, 97% (20) M. senegalense</td>
<td>100% M. fortuitum, 99.5% (2) M. houstonense, 99.2% (3) M. senegalense</td>
<td>100% M. fortuitum sequevar Mfo C, 99.5% (1) M. fortuitum sequevar Mfo A</td>
</tr>
<tr>
<td>104</td>
<td>100% M. abscessus subsp. abscessus or M. abscessus subsp. bolletii</td>
<td>100% M. abscessus subsp. abscessus, 96.5% (23) M. abscessus subsp. bolletii</td>
<td>100% M. abscessus subsp. abscessus, 98.5% (6) M. abscessus subsp. bolletii</td>
<td>99.5% (1) M. abscessus sequevars Mab A or Mab B, 98.7% (3) M. abscessus subsp. bolletii</td>
</tr>
<tr>
<td>215</td>
<td>100% M. avium subsp. avium, 99.7% (4) M. colombiense</td>
<td>99.5% (3) M. avium, 95.5% (30) M. indicus</td>
<td>99.5% (2) M. phlei, 95.1% (19) M. splagni</td>
<td>100% M. avium sequevar Mac A, 97% M. intracellularare sequevar Min A</td>
</tr>
<tr>
<td>224</td>
<td>100% M. intracellulare, 99.8% (3) M. avium or M. marseillense</td>
<td>99.5% (3) M. intracellulare, 98.2% (7) M. cheimera, 97.3% (11) M. avium</td>
<td>100% M. phlei sequevar Mph A, 92% (23) M. phlei sequevar Mph B</td>
<td>98.6% (3) M. intracellulare Min A, 97.6% (5) M. cheimera, 96.1% (8) M. avium sequevar Mac A</td>
</tr>
<tr>
<td>207</td>
<td>100% M. phlei, 98.5% (21) M. pulviris</td>
<td>99.5% (1) M. phlei</td>
<td>100% M. phlei sequevar Mph A, 92% (23) M. phlei sequevar Mph B</td>
<td>99.5% (15) M. abscessus subsp. bolletii</td>
</tr>
<tr>
<td>136</td>
<td>100% M. nonchromogenicum, M. abscessus subsp. abscessus</td>
<td>98.4% (11), M. abscessus subsp. bolletii, 96.5 (23) M. abscessus subsp. abscessus</td>
<td>99.3% (3) M. abscessus subsp. bolletii, 98.6% (5) M. abscessus subsp. abscessus</td>
<td>99.6% (1) M. abscessus subsp. bolletii, 98.7% (3) M. abscessus sequevars Mab A or Mab B</td>
</tr>
<tr>
<td>74</td>
<td>100% M. nonchromogenicum</td>
<td>99.5% (2) M. nonchromogenicum</td>
<td>100% M. nonchromogenicum</td>
<td>100% M. nonchromogenicum</td>
</tr>
<tr>
<td>201</td>
<td>99.5% (7) M. thermoresistibile</td>
<td>99.3% (2) M. thermoresistibile</td>
<td>100% M. thermoresistibile</td>
<td>99.53% (13) M. thermoresistibile</td>
</tr>
<tr>
<td>191</td>
<td>100% M. peregrinum, 99.7% (4) M. septicum</td>
<td>99.5% (2) M. peregrinum, 99.3% (3) M. septicum</td>
<td>100% M. peregrinum sequevar MpeA</td>
<td>99.93% (2) M. peregrinum</td>
</tr>
<tr>
<td>137</td>
<td>100% M. conceptionense, 99.7% (4) M. porcinum or M. fortuitum</td>
<td>99.6% (5) M. conceptionense</td>
<td>100% M. conceptionense, 93% (16) M. senegalense</td>
<td>99.65% (1) M. conceptionense</td>
</tr>
<tr>
<td>NM7</td>
<td>100% M. chelonae, 99.7% (4) M. abscessus subsp. abscessus subsp. bolletii</td>
<td>98.5% (3) M. chelonae</td>
<td>100% M. chelonae sequevar Mch A, 99.6% (1) M. chelonae sequevar Mch B, 94.7% (13) M. chelonae sequevar Mch C</td>
<td>99.6% (9) M. chelonae</td>
</tr>
<tr>
<td>NM44</td>
<td>100% M. triplex, 99.5% (7) M. genavense</td>
<td>98.7% (5) M. florentinum, 97.9% (8) M. lentiflavum, 96.4% (14) M. montefiore, 96.4% (14) M. triplex</td>
<td>98% (8) M. lentiflavum, 96.6% (14) M. triplex</td>
<td>99.2% (22) M. triplex</td>
</tr>
<tr>
<td>NM47</td>
<td>100% M. branderi, 98% (30) M. kyroense</td>
<td>100% M. branderi</td>
<td>99.3% (3) M. branderi, 98% (8) M. celatum</td>
<td>100% M. branderi</td>
</tr>
<tr>
<td>175</td>
<td>100% M. simiae, 99.6% (6) M. lentiflavum</td>
<td>98.5% (10) M. simiae</td>
<td>100% M. simiae sequevar Mis A or Mis C, 99.5% (1) M. simiae sequevar Mis D</td>
<td>99.2% (3) M. simiae</td>
</tr>
<tr>
<td>218</td>
<td>100% M. lentiflavum, 99.6% (6) M. simiae</td>
<td>99.3% (3) M. simiae or M. genavense or M. shuoidai</td>
<td>98.6% (3) M. lentiflavum, 97.2% (6) M. genavense</td>
<td>99.6% (1) M. lentiflavum</td>
</tr>
<tr>
<td>220</td>
<td>100% M. kansasii or M. gastri</td>
<td>100% M. kansasii, 98% (8) M. gastri</td>
<td>100% M. kansasii sequevar Mka A, 90.3% (20) M. gastri Mga A, 92.7% (15) M. gastri sequevar Mga A</td>
<td>99.6% (1) M. kansasii</td>
</tr>
<tr>
<td>227</td>
<td>99.7% (4) M. kansasii or M. gastri</td>
<td>96.7% (22) M. kansasii, 96% (26) M. gastri</td>
<td>96.7% (22) M. kansasii, 96% (26) M. gastri</td>
<td>97.8% (58) M. kansasii or M. gastri</td>
</tr>
</tbody>
</table>

* Sequevar variation by ITS, in some species of mycobacteria was adopted based on Roth et al. (1998).
During the current study, sequences of four genes including 16S rRNA, *rpoB* and *hsp65* genes and ITS region were analyzed separately and then a combined database used to create a single alignment dataset for final identification. All mycobacterial species studied showed good separation.

Table 2 summarizes the percentage differences and nucleotide mismatches of almost full 16S rRNA and partial sequences of *rpoB*, *hsp65* and ITS of representative clinical isolates of each group of NTM in detail which was clustered based on *hsp65*-PRA.

### 3.3.2. Identification by 16S rRNA, *rpoB*, *hsp65*, ITS and MLSA

The nucleotide signature sequences of hypervariable A (positions 125–270) and B (positions 408–503) of the 16S rRNA (Patel et al., 2000), region V of the *rpoB* gene from position 2581–3300 (Adekambi et al., 2003), nucleotide differences particularly frequent in two regions (positions 624–664 and positions 683–725) as hypervariable regions of the *hsp65* gene (Ringuet et al., 1999) and sequence variability over the whole spacer sequence of the ITS (Roth et al., 1998) were used to define species. The 16S rRNA sequences of all rapidly test strains contained the short helix at positions 451–482, that is, the sequence which is characteristic of RGM. Similarly, the 16S rRNA of all slowly growing test strains had the extended helix at positions 451–482 which is characteristic of SGM. Selected nucleotide signature sequences of the test strains for hypervariable A of the 16S rRNA together with those of closely related RGM and SGM, are shown in Supplementary online Fig. S2.
Sequence variability were shown over the whole investigated of the segment of the rpoB, hsp65 and ITS, however a motifs between position 3042–3242 from region V of the rpoB gene mostly contribute to species identification. In the case of the hsp65 gene, most of differences are clustered among two motifs between positions 624–664 and positions 683–725. Sequence alignment of the ITS segment showing that the ITS of SGM are approximately 75 base pairs shorter than those of RGM which clearly differentiate two group of mycobacteria. High degree of sequence variability was found over the whole ITS sequence which had important role to define species and differentiate the strains.

Clinical isolates including AFP-00047 (M. branderi), AFP-00074 (M. nonchromogenicum), AFP-000191 (M. peregrinum), AFP-000201 (M. thermoresistibile), AFP-000207 (M. phlei), AFP-000218 (M. lentiflavum), AFP-000NM7 (M. chelonae) and AFP-000SM8 (M. fortuitum) can be confidently identified by each of the 16S rRNA, rpoB, hsp65 and ITS.

**Fig. 2.** rpoB sequence-based phylogenetic tree of clinical isolates of Iranian NTM with those of closely related species which computed by the NJ analyses and K2P model. The support of each branch, as determined from 1000 bootstrap samples, is indicated by percentages at each node. Bar 0.02 substitutions per nucleotide position. T; type strains. *; the isolate AFP-000224 which identified M. intracellulare by MLSA, clustered by M. indicus.
Fig. 3. hsp65 sequence-based phylogenetic tree of clinical isolates of Iranian NTM with those of closely related species which computed by the NJ analyses and K2P model. The support of each branch, as determined from 1000 bootstrap samples, is indicated by percentages at each node. Bar 0.01 substitutions per nucleotide position. T: type strains, *, the isolates AFP-000NM44 that identified M. triplex by MLSA, clustered by M. florentinum. The isolate AFP-000215 which identified and clustered by M. avium subsp. avium, formed unique phylogenetic line by M. colombiense. The isolate AFP-000175 that identified and clustered by M. simiae based on MLSA, clustered by M. simiae, M. genavense and M. shimoidei.
Several clusters were characterized in RGM and SGM groups by phylogenetic constructed trees using the 16S rRNA, rpoB, hsp65 and ITS of 17 NTM clinical isolates (Figs. 1–4). As shown in the Figs. 1–4, AFP-00047, AFP-00074, AFP-000191, AFP-000201, AFP-000207, AFP-000218, AFP-000NM7 and AFP-000SM8 grouped with relevant species by high bootstrap values. In an attempt to improve the taxonomic classification of NTM isolates, concatenated sequences of the 16S rRNA, rpoB, hsp65 and ITS and also concatenated sequences of the rpoB, hsp65 and ITS of representative isolate of each hsp65-PRA pattern were investigated. MLSA based identification and phylogenetic analysis also confirmed the identity of the isolates AFP-000NM44 that identified M. triplex by MLSA, clustered by M. florentinum.

Several clusters were characterized in RGM and SGM groups by phylogenetic constructed trees using the 16S rRNA, rpoB, hsp65 and ITS of 17 NTM clinical isolates (Figs. 1–4). As shown in the Figs. 1–4, AFP-00047, AFP-00074, AFP-000191, AFP-000201, AFP-000207, AFP-000218, AFP-000NM7 and AFP-000SM8 grouped with relevant species by high bootstrap values. In an attempt to improve the taxonomic classification of NTM isolates, concatenated sequences of the 16S rRNA, rpoB, hsp65 and ITS and also concatenated sequences of the rpoB, hsp65 and ITS of representative isolate of each hsp65-PRA pattern were investigated. MLSA based identification and phylogenetic analysis also confirmed the identity of the isolates AFP-00047, AFP-00074, AFP-000191, AFP-000201, AFP-000207, AFP-000218, AFP-000NM7 and AFP-000SM8 (Fig. 5).

16S rRNA and rpoB identified the isolate AFP-00044 as M. triplex species (Figs. 1 and 2), but the isolate shared more similarity by the hsp65 and ITS to those of M. florentinum (Figs. 3 and 4). However MLSA confirmed the identity of isolate as M. triplex (Fig. 5).
16S rRNA, hsp65 and ITS identified the isolate AFP-000137 as *M. conceptionense* species (Fig. 1 and 2 and Fig. 4), but the isolate shared more similarity by the *rpoB* to *M. senegalense*. Identification by MLSA confirmed the identity of isolate as *M. conceptionense* (Fig. 5). 16S rRNA, *rpoB* and ITS identified the isolate AFP-000175 as *M. simiae* species (Figs. 1 and 2 and Fig. 4), but the isolate showed an identical hsp65 sequence to those of *M. simiae*, *M. genavense* and *M. shimoidei*. MLSA confirmed the identity of the isolate as *M. simiae* (Fig. 5).

The isolate AFP-000215 identified as *M. avium* using the 16S rRNA, *rpoB* and ITS (Figs. 1 and 2 and Fig. 4), however the hsp65 sequence identified the strain as *M. colombiense* (Fig. 5). MLSA confirmed the identity of the isolate as *M. avium* (Fig. 5).

The isolate AFP-000224 identified as *M. intracellulare* using the 16S rRNA, hsp65 and ITS (Fig. 1, Figs. 3 and 4), however the *rpoB* sequence identified the strain as *M. indicus* (Fig. 2). MLSA confirmed the identity of the isolate as *M. intracellulare* (Fig. 5).

The identity of the isolates AFP-000104 and AFP-000136 were unclear due to 100% similarity of the 16S rRNA gene of *M. abscessus* subsp. *bolletii* and *M. abscessus* subsp. *abscessus* (Leao et al., 2011; Macheras et al., 2009; Yakrus et al., 2001). The isolate AFP-000104 and AFP-000136 identified as *M. abscessus* subsp. *abscessus* and *M. abscessus* subsp. *bolletii*, respectively by the *rpoB*, hsp65 and ITS (Figs. 2 and 3 and Fig. 4). MLSA confirmed the identity of the isolates (Fig. 5).

As a result of identical 16S rRNA gene of *M. kansasii* and *M. gastri*, real identity of the isolates AFP-000220 and AFP-000227 was unclear. The isolate AFP-000220 identified as *M. kansasii* by the *rpoB*, hsp65 and ITS (Figs. 2–4). MLSA confirmed the identity of the isolate as *M. kansasii* (Fig. 5). The isolate shown similarity to *M. kansasii* by the *rpoB* and hsp65 (Figs. 2 and 3). However the isolate shown 92.7% similarity of the ITS to that of *M. gastri* (Fig. 4). The isolate shown novel species signature by MLSA (Fig. 5).
3.3.3. Sequevar investigation by ITS

At the interspecies level, the clinical isolates AFP-000175, AFP-000191, AFP-000207, AFP-000215, AFP-000220, AFP-000NM7 and AFP-000SM8 showed, respectively an identical ITS sequences with those of references sequence strains of *M. simiae* sequenar Mis A or Mis C, *M. peregrinum* sequenar Mpe A, *M. phlei* sequenar Mph A, *M. avium* sequenar Mac A, *M. kansasi* sequenar Mka A, *M. chelonae* sequenar Mch A and *M. fortuitum* sequenar Mfo C (Table 2).

Intraspecies sequence polymorphism was not detected in *M. nonchromogenicum* (AFP-000074), *M. conceptionense* (AFP-000137), *M. thermoresistibile* (AFP-000201) and *M. branderi* (AFP-000N47). The clinical isolates AFP-000218 and AFP-000224 which were identified as *M. lentiflavum* and *M. intracellulare* showed high number of nucleotide differences with previous known sequevaras (Table 2), and confirmed the new distinct sequevars. The clinical isolate AFP-000104, showed identical ITS sequence to that of *M. abscessus* type I (AV635142). The clinical isolate AFP-000227 was found to have 15 and 17 nucleotide variation from those of *M. gastris* and *M. kansasi*, respectively, and was clearly distinguishable from these species.

4. Discussion

Prokaryotic species are currently identified using a polyphasic approach including combination of genotypic and phenotypic properties. Sequence analysis of 16S RNA allows the rapid classification of prokaryotes using a universally distributed trait (Woese, 1991). Due to low resolution of 16S RNA gene sequence analysis between closely related species of *Mycobacterium*, other genes such as *rpoB* (Adekambi et al., 2003), *hsp65* (Ringuet et al., 1999), ITS (Roth et al., 1998) and recA (Blackwood et al., 2000) that are evolved more rapidly than 16S RNA gene have been introduced for identification of mycobacterial species. However, species identification using single locus is inappropriate, particularly for those loci where rates of recombination are high, since homologous recombination distorts the true relationships between isolates of closely related named species (Cohan, 2002; Ursing et al., 1995). The use of MLSA ensures that recombination at one locus is buffered by the more reliable indications of relatedness provided by the other loci (Coenye et al., 2005; Gevers et al., 2005).

Various investigations have revealed an increase in infections caused by NTM from different parts of the world especially at high tuberculosis setting (Brown-Elliott and Wallace, 2002; Chetchelelskad et al., 2007; Gopinath and Singh, 2010; Katoch, 2004; Lai et al., 2010). As a result of clinical relevance, rapid and reliable identification of NTM at species level should be carried out as a means of effective patient managements and molecular epidemiology (Brown-Elliott and Wallace, 2002; Griffith et al., 2007; Shojaei et al., 2011).

During 2008–2012, a total of 92 isolates of NTM recovered from tuberculosis suspected cases from different regions of Iran which their etiological role was confirmed according to ATS criteria (Griffith et al., 2007). The clinical isolates were accordingly identified and clustered using phenotypic tests and the *hsp65*-PRA method. Out of 92 isolates, 66 (72%) isolates were assigned to species or complex level using phenotypic tests and 81 (83%) isolates showed previously reported *hsp65*-PRA patterns (12 different species). Similar to our study, other reports from different parts of the world confirmed that species identification by *hsp65*-PRA was significantly more accurate than the phenotypic methods (Chimara et al., 2008; Martin et al., 2007). PRA profiles of *hsp65* could differentiate members of MAC as well as *M. abscessus*, *M. chelonae* and *M. peregrinum* from *M. fortuitum* which are biochemically similar but often undistinguishable (Wong et al., 2003). As previous report of Iran (Shojaei et al., 2011), in our study, *M. fortuitum* is the most frequently encounter species of NTM in clinical setting and most of the patients with NTM disease suffered from immunosuppressive syndrome.

Less conclusive identification such as phenotypic tests and the presence of four unknown patterns as a result of gel to gel variation due to small restriction fragment sizes as well as one identical pattern to that of *M. confectionense* or *M. senegalense* based on *hsp65*-PRA method, emphasized the need for more reliable identification system (Supplementary online Fig. S1).

16S rRNA is the most common technique currently used for *Mycobacterium* species identification in reference laboratories (Brown-Elliott and Wallace, 2002; Devulder et al., 2005; Tortoli, 2003). However, the 16S rRNA have limited discriminating power for several closely related Mycobacterium species (Adekambi et al., 2003; Blackwood et al., 2000; Devulder et al., 2005; Ringuet et al., 1999; Roth et al., 1998). For definite and reliable identification, one clinical isolate from each *hsp65*-PRA groups in our research, was subjected to identification by the 16S rRNA, *rpoB*, *hsp65* and ITS.

16S rRNA sequence analysis revealed that twelve isolates had identical sequences to those of species *M. fortuitum*, *M. avium* subsp. *avium*, *M. intracellulare*, *M. phlei*, *M. nonchromogenicum*, *M. peregrinum*, *M. conceptionense*, *M. chelonae*, *M. triplex*, *M. branderi*, *M. simiae* and *M. lentiflavum*. The latter species have few nucleotide differences ranging from 3 to 21 base differences in their 16S rRNA gene compared to other closely related species (Devulder et al., 2005; Menendez et al., 2002) which, in turn, strongly emphasizes the need for complementary genetic marker for better identification.

Based on the *rpoB* gene sequences analyses, most of the clinical isolates were identified more correctly (Table 2). Sequence analysis of the *rpoB* gene identified the isolates AFP-000NM44, AFP-000N47, AFP-000104, AFP-000175, AFP-000191, AFP-000201 and AFP-000SM8 as *M. triplex*, *M. branderi*, *M. abscessus* subsp. *abscessus*, *M. simiae*, *M. peregrinum*, *M. thermoresistibile*, *M. lentiflavum* and *M. fortuitum*, respectively. The isolates belonging to some closely related species like AFP-000104, AFP-000136, AFP-000220 which are poorly discriminated by the 16S rRNA gene sequence were clearly delineated as *M. abscessus* subsp. *abscessus*, *M. abscessus* subsp. *bolletii* and *M. kansasi*, respectively indicating that the 16S RNA gene was less discriminatory than the *rpoB* gene for species identification (Adekambi et al., 2003). However for AFP-000137 identified as *M. conceptionense* based on the 16S rRNA, *hsp65* and ITS, the sequence analysis of the *rpoB* showed most closely relatedness to *M. senegalense*, rather than *M. conceptionense* which emphasizing need to multiple marker for reliable identification (Supplementary online Fig. S1).

Based on the *hsp65* gene sequences analyses, most of the clinical isolates were identified more correctly (Table 2). The isolates belonging to some closely related species like AFP-000104, AFP-000136, AFP-000137 and AFP-000220 which are poorly discriminated by the 16S rRNA gene sequence were clearly delineated as *M. abscessus* subsp. *abscessus*, *M. abscessus* subsp. *bolletii*, *M. conceptionense* and *M. kansasi*, respectively indicating that the 16S RNA gene was more discriminatory than the *rpoB* gene for species identification (Adekambi et al., 2003). However for AFP-000137 identified as *M. conceptionense* based on the 16S rRNA, *hsp65* and ITS, the sequence analysis of the *rpoB* showed most closely relatedness to *M. senegalense*, rather than *M. conceptionense* which emphasizing need to multiple marker for reliable identification (Supplementary online Fig. S1).
Although our result here, indicates that the hsp65 is as suitable tool for identifying most mycobacterium species, but the data indicate that the hsp65 gene sequence alone did not identify the closely related species and resolve the phylogenetic relationships between all currently recognized mycobacterium species (Supplementary online Fig. S1).

Based on ITS sequences, most of the clinical isolates were identified more correctly (Table 2). Sequences analysis of the ITS identified M. nonchromogenicum, M. conceptionense, M. simiae, M. peregrinum, M. thermoresistibile, M. avium, M. kansasii, M. chelonae, M. florentinum, M. branderi, M. fortuitum and also M. abscessus subsp. abscessus, M. abscessus subsp. bolletii and M. kansasii which are poorly discriminated by the 16S rRNA.

Like hsp65, the discrepancy result was obtained for identification of AFP-000NM44 using the ITS. The latter isolate was identified as M. triplex by the 16S rRNA and rpoB, but the isolate showed a highest similarity to that of M. florentinum using the ITS sequence analysis. The ITS region has the potential to be used for clinically significant strain differentiation (De Smet et al., 1995; Frothingham and Wilson, 1994; Hamid et al., 2002; Roth et al., 1998). The ITS based sub-clustering, existing among the clinical isolates of NTM in each part of the world and in our study (Table 2) could be further evaluated if a bigger number of strains were analyzed. The result here, shown that although the ITS provide desirable sensitivity and specificity, it has limited to the identification of a small number of mycobacteria (Supplementary online Fig. S1).

Our result based on sequence analysis of the 16S rRNA, rpoB, hsp65 and ITS, confirmed the need for MLSA rather than single locus sequence analysis.

Although the rpoB, hsp65 or ITS significantly improved the resolution of species identification and complemented the phylogeny of genus Mycobacterium based on 16S rRNA gene sequences, in our study, using the concatenated sequence analysis of 16S rRNA, rpoB, hsp65 and ITS and also concatenated sequences of rpoB, hsp65 and ITS of representative isolate of each hsp65-PRA pattern markedly improved molecular identification and resolution of phylogenetic tree for clinical isolates. Initial inspection of MLSA tree (Fig. 5) reveals that the clinical isolates are clustered with their correspondence reference Mycobacterium species by 100% bootstrap values. In the current study, MLSA as the new standard in the molecular identification (Achtman and Wagner, 2008; Gevers et al., 2005) provided more detailed separations at species level in the clinical isolates with the highest bootstrap values than that obtained by the use of any single locus. The concatenation of genes including the rpoB, hsp65 and ITS also shown to be sensitive for identification and phylogenetic analysis rather than single locus (data not shown). Although the real identity of the isolate AFP-000227 was unclear, according to MLSA data the isolate, it is potentially a new separate species which is closely related to M. kansasii and M. gastro within genus. The evidences obtained based on current multi gene approach, allowed us to unambiguously classify the clinical isolates of NTM at the species level.

MLSA has also shown to be successfully classify the diverse environmental (Castillo-Rodal et al., 2012; Dai et al., 2011) and clinical (Devulder et al., 2005; Macheras et al., 2011) mycobacteria. Our reports revealed that hsp65-PRA as simple, rapid and inexpensive can be useful for the identification of frequently, not rare, encountered species of NTM in clinical setting; but sequence based identification provide more reliable evidence for species assignment. MLSA identified all of the isolates to species level and also identified a novel species within the isolates. On the other hand, our analysis results clearly showed that identification based on single locus is not valuable for reliable separation and combination of multi locus markedly improves the ability of the sequence based identification.

5. Conclusion
In conclusion, these data suggest that sequence based identification using MLSA of 16S rRNA, rpoB, hsp65 and ITS, should provide much better and more reliable discrimination rather than single locus identification in reference laboratories.

6. Conflict of interest
None to declare.

Acknowledgments
The authors are grateful to the Vice Cancellor of Research and Technology Development of Jundishapur University of Medical Science, Ahvaz, Iran (Grant number: 90124-Project Number 83) and the Deputy Director of Islamic Azad University of Parand for financial support of the current study. The current study has been approved by the ethical committee of Voice Cancellor of Research and Technology Development of Jundishapur Medical School University (Ahvaz, Iran) and Masoud Laboratory (Tehran, Iran).

Appendix A. Supplementary data
Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.meegid.2013.08.027.

References


