Molecular identification of clinical *Nocardia* isolates from India

Shivaprakash M. Rudramurthy, Prasanna Honnavar, Harsimran Kaur, Palash Samanta, Pallab Ray, Anup Ghosh and Arunaloke Chakrabarti

Department of Medical Microbiology, Postgraduate Institute of Medical Education and Research, Chandigarh, India

The epidemiology of nocardiosis is evolving with increasing number of *Nocardia* spp. causing human infection. In recent years, molecular techniques have been used to identify *Nocardia* spp. There are limited data available on the spectrum of *Nocardia* spp. isolated from clinical samples in India. Here, a molecular study was carried on 30 clinical isolates maintained in our National Culture Collection to evaluate the techniques used for identifying the agents. The isolates were identified by sequencing two promising genes: the 16S rRNA gene and *hsp65*. Both *hsp65* and the 16S rRNA gene could reliably identify 90% of *Nocardia* isolates, i.e. *N. farcinica*, *N. cyriacigeorgica*, *N. brasiliensis*, *N. otitidiscaviarum*, *N. amibiensis* and *N. pneumoniae*. The mean percentage dissimilarity of sequence identification was higher using the *hsp65* gene (4%, range 0–7.9%) compared with the 16S rRNA gene (2.3%, range 0–8.9%). Two isolates that showed ambiguous results in both the short segment of the 16S rRNA gene and *hsp65* sequences could be resolved by sequencing a larger fragment (~1000 bp) of the 16S rRNA gene. Both of these isolates were identified as *N. beijingensis* with similarities of 99.8 and 100% compared with the standard strain. Genotyping of *N. cyriacigeorgica* strains was performed using *hsp65* gene sequences and compared with previously described genotypes. Our *N. cyriacigeorgica* isolates belonged to genotype 1 (*n*=4) and genotype 2 (*n*=2). The present study highlights a wide spectrum of *Nocardia* spp. in India and emphasizes the need for molecular techniques for identification to the species level.

**INTRODUCTION**

The number of *Nocardia* spp. causing human infections has increased in recent years. Nearly 30 of 102 species of *Nocardia* have been implicated in human infection (Carrasco et al., 2013; see http://www.bacterio.net/nocardia.html and the NCBI taxonomy for *Nocardia* at http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id_1817). Species identification by phenotypic methods is cumbersome and time consuming, and may lead to misidentification in up to 37% of isolates (Baio et al., 2013; Beaman & Beaman, 1994; Carrasco et al., 2013). For rapid and precise speciation of *Nocardia*, various molecular methods have been employed to overcome the problems of conventional methods. Among molecular techniques, DNA sequencing of the conserved region, DNA–DNA hybridization and RFLP analysis have been used (Borriello et al., 2007; Patel et al., 2004). Although partial 16S rRNA gene sequencing is widely used for identifying this group of organisms, this region may not be completely reliable in identification of *Nocardia* spp. (Patel et al., 2004). As the *hsp65* gene has more microheterogeneity regions compared with the 16S rRNA gene and is three to four times more discriminatory, this gene sequence has been targeted for *Nocardia* spp. identification (Rodriguez-Nava et al., 2006; Schlaberg et al., 2008; Yin et al., 2007). However, very few sequences of the *hsp65* gene of *Nocardia* spp. are available in public databases for comparison, which limits its use in identification. The present study was conducted to identify the spectrum of *Nocardia* spp. causing human infection in different...
regions of India and to compare the utility of sequencing of the 16S rRNA and hsp65 genes in identifying Nocardia spp.

METHODS

Clinical isolates. A total of 30 Nocardia spp. collected from April 2006 to March 2014 were included in the study. Nocardia spp. were isolated from respiratory specimens (n = 17) [sputum (n = 8)], fine-needle aspiration from lung lesions (n = 4), bronchoalveolar lavage (n = 3), endotracheal aspirate (n = 2), pyogenic cutaneous lesions (n = 6), blood (n = 4), cerebral abscess (n = 2) and corneal scraping, (n = 1). The isolates were either isolated from patients admitted to the Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh or were received at the National Culture Collection of Pathogenic Fungi (NCCPF). PGIMER, for identification from various provinces in the country including Punjab, Himachal Pradesh, Rajasthan, Maharashtra, Karnataka and Puducherry. All isolates were stored in the facilities of the NCCPF. Nocardia spp. identified previously by standard phenotypic methods (Land, 1992; Winn et al., 2005) were subjected to molecular identification.

DNA extraction and PCR amplification. DNA was extracted from the isolates by a phenol/chloroform method as described by Torres et al. (1996). Amplification was carried out for the 16S rRNA (605 bp) and hsp65 (440 bp) genes (La Scola et al., 1998; Telenti et al., 1993; Weisburg et al., 1991). The primers used for 16S rRNA gene amplification and sequencing were 5′-GCTTAAACGATCCAAATCG-3′ and 5′-AAGTTCCAGTCTGCCCCTG-3′, and for hsp65 were 5′-AACCAGTGTTGTCATCT-3′ and 5′-CCTGTGGAGCCGATTACCC-3′. Sequencing reactions were performed with a Big Dye Terminator Cycle Sequencing kit, version 3.1 (Applied Biosystems) for both strands. The sequencing reaction products were purified and analysed on an ABI 3130 Genetic Analyzer (Applied Biosystems).

Analysis of sequence data (16S rRNA and hsp65 genes). Consensus sequences were prepared for each isolate from sequence data obtained by forward and reverse primers by using Bionumerics software (version 7.1; Applied Maths). These sequences were compared with the sequence data of type strains deposited in GenBank using NCBI BLAST software (http://www.ncbi.nlm.nih.gov/genbank/). Species identity was considered significant when the 16S rRNA gene sequence of the isolate showed a >99 % similarity score with the sequence of the reference or the type strains (Kong et al., 2009). However, no such cut-off of similarity score was established for hsp65 sequence-based identification. The concordance between 16S rRNA and hsp65 gene sequence-based identification was determined using a Kappa test.

For phylogenetic analysis, the sequences of the standard strains were retrieved from GenBank. The sequences were aligned using CLUSTAL_X version 2.0.11 (European Bioinformatics Institute). The evolutionary history and evolutionary distances were computed using the neighbour-joining and the maximum composite likelihood methods, respectively (Felsenstein, 1985; Saitou & Nei, 1987; Tamura et al., 2004). All positions containing missing data and gaps were eliminated. MEGAS was used to conduct evolutionary analyses (Tamura et al., 2011). A bootstrap test was conducted using 1000 replicates (Felsenstein, 1985). The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Streptomyces somaliensis (DSM 40738) was used as an outgroup. Genotyping of Nocardia cyriacigeorgica strains was carried out using hsp65 gene sequences following the above parameters.

RESULTS

The demographic and clinical characteristics of the patients are provided in Table 1. The majority (87 %) of the patients were males with a male : female ratio of 6.5 : 1. The mean age of the patients was 49.6 years, with the majority (63 %) in the age range of 30–60 years. The predisposing factors could not be ascertained in the majority (70 %) of patients. The predisposing features that could be ascertained included chronic obstructive pulmonary disease, human immunodeficiency virus infection, renal transplant, leukaemia and corticosteroid use.

A comparison of the BLAST results of the 16S rRNA and hsp65 gene sequences of all our isolates is presented in Table 1. Figs. 1 and 2 illustrate the phylogenetic trees of the clinical isolates and reference strains constructed by using 16S rRNA gene and hsp65 gene sequences, respectively. The optimal trees (evolutionary distance data of species through time) with the sum of branch lengths = 0.2152527 (16S rRNA gene) and branch lengths = 0.30916202 (hsp65) is shown in Figs. 1 and 2 respectively. These represent 412 and 337 positions in the final dataset of the 16S rRNA gene sequence and hsp65 gene sequences, respectively. The percentage dissimilarities in the sequences of the 16S rRNA and hsp65 gene are provided in Tables S1 and S2, respectively, available with the online Supplementary Material. The mean percentage dissimilarity of sequence identification was higher using hsp65 gene sequences (4 %, range 0–7.9 %) compared with 16S rRNA gene sequences (2.3 %, range 0–8.9 %). Both hsp65 and 16S rRNA gene sequences could reliably identify N. farcinica, N. cyriacigeorgica, N. brasiliensis, N. otitidiscausticum, N. amaniensis and N. pneumoniae. Among these strains, nucleotide variations ranged from 0 to 2.0 % in the 16S rRNA gene region, whereas variations were higher in the hsp65 region (range 0–3.9 %) (Table 2). The concordance between the 16S rRNA gene and hsp65 gene based sequence identification was very good with a k value of 0.867.

Two isolates (NCCPF 260081 and 260040) showed ambiguous results in both the 16S rRNA gene (605 bp) and hsp65 gene (440 bp) sequences. To resolve this ambiguity, a larger fragment of the 16S rRNA gene was sequenced and compared with GenBank sequences. The 16S rRNA gene analysis of isolate NCCPF 260040 showed 99.8 % similarity (1074/1076 bp) with N. beijingensis (DSM 44636), whereas the hsp65 gene sequence of the same isolate exhibited 98.8 % similarity with N. arthritidis (DSM 44731). The hsp65 gene sequence of isolate NCCPF 260081 had 99.1 % similarity with N. asiatica (DSM 44668) and its 16S rRNA gene results showed 100 % similarity (1011/ 1011 bp) with N. beijingensis (DSM 44636).

Genotyping of N. cyriacigeorgica

Two isolates (NCCPF 260082 and N17) identified as N. cyriacigeorgica using the 16S rRNA gene region sequences clustered separately in the hsp65 phylogenetic tree. A detailed investigation of the six N. cyriacigeorgica isolates revealed the existence of intraspecies microheterogeneity
Table 1. Clinical characteristics of patients included in the study and molecular identification of isolates using 16S rRNA and hsp65 gene sequencing

<table>
<thead>
<tr>
<th>Sample no.</th>
<th>NCCPF no./year of isolation</th>
<th>Geographical origin</th>
<th>Age/sex</th>
<th>Chief complaints</th>
<th>Predisposing factor</th>
<th>Sample/site of isolation</th>
<th>16S rRNA gene</th>
<th>Reference strain</th>
<th>Similarity (%)</th>
<th>Hsp65</th>
<th>Reference strain</th>
<th>Similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>260023/2005</td>
<td>Chandigarh</td>
<td>11/M</td>
<td>Fever, cough, haemoptysis</td>
<td>Oral corticosteroids</td>
<td>Sputum</td>
<td>N. otitidiscaviarum</td>
<td>N. otitidiscaviarum ATCC 14629</td>
<td>100</td>
<td>N. farcinica</td>
<td>N. farcinica</td>
<td>98</td>
</tr>
<tr>
<td>2</td>
<td>260024/2006</td>
<td>Chandigarh</td>
<td>28/M</td>
<td>Fever, localized pain over swelling on back</td>
<td>Not known</td>
<td>Aspirate from paraspinal abscess</td>
<td>N. farcinica</td>
<td>N. farcinica ATCC 3318</td>
<td>99</td>
<td>N. farcinica</td>
<td>N. farcinica</td>
<td>99</td>
</tr>
<tr>
<td>3</td>
<td>260112/2006</td>
<td>Chandigarh</td>
<td>21/M</td>
<td>Fever, cough, dyspnoea</td>
<td>Not known</td>
<td>ET aspirate</td>
<td>N. cyriacigeorgica</td>
<td>N. cyriacigeorgica ATCC 14759</td>
<td>99</td>
<td>N. cyriacigeorgica</td>
<td>N. cyriacigeorgica</td>
<td>99</td>
</tr>
<tr>
<td>4</td>
<td>260038/2009</td>
<td>Chandigarh</td>
<td>45/M</td>
<td>Fever, cough</td>
<td>Not known</td>
<td>BAL</td>
<td>N. pneumoniae</td>
<td>N. pneumoniae DSM 44730</td>
<td>99</td>
<td>N. pneumoniae</td>
<td>N. pneumoniae</td>
<td>99</td>
</tr>
<tr>
<td>5</td>
<td>260040/2009</td>
<td>Chandigarh</td>
<td>58/F</td>
<td>Cough, dyspnoea</td>
<td>Not known</td>
<td>Lung FNAC</td>
<td>N. beijingensis*</td>
<td>N. beijingensis DSM 44636</td>
<td>99.8</td>
<td>N. arthritidis</td>
<td>N. arthritidis</td>
<td>98.8</td>
</tr>
<tr>
<td>6</td>
<td>260041/2009</td>
<td>Chandigarh</td>
<td>58/F</td>
<td>Fever, dyspnoea, chest pain</td>
<td>Not known</td>
<td>Lung FNAC</td>
<td>N. otitidiscaviarum</td>
<td>N. otitidiscaviarum ATCC 14629</td>
<td>99</td>
<td>N. otitidiscaviarum</td>
<td>N. otitidiscaviarum</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>260042/2009</td>
<td>Rajasthan</td>
<td>59/M</td>
<td>Loss of consciousness, vomiting, Fever, chills and rigors</td>
<td>Not known</td>
<td>Brain abscess</td>
<td>N. farcinica</td>
<td>N. farcinica DSM 43665</td>
<td>100</td>
<td>N. farcinica</td>
<td>N. farcinica</td>
<td>99</td>
</tr>
<tr>
<td>9</td>
<td>260047/2009</td>
<td>Himachal Pradesh</td>
<td>40/M</td>
<td>Multiple discharging sinuses on foot, swelling</td>
<td>Not known</td>
<td>Discharge from foot</td>
<td>N. brasiliensis</td>
<td>N. brasiliensis ATCC 700358</td>
<td>99</td>
<td>N. brasiliensis</td>
<td>N. brasiliensis</td>
<td>99</td>
</tr>
<tr>
<td>10</td>
<td>N17/2009</td>
<td>Karnataka</td>
<td>75/F</td>
<td>Loss of consciousness, vomiting, Fever, chills and rigors</td>
<td>Not known</td>
<td>Sputum</td>
<td>N. cyriacigeorgica</td>
<td>N. cyriacigeorgica ATCC 14759</td>
<td>99</td>
<td>N. cyriacigeorgica</td>
<td>N. cyriacigeorgica</td>
<td>99</td>
</tr>
<tr>
<td>11</td>
<td>260110/2010</td>
<td>Chandigarh</td>
<td>55/M</td>
<td>Fever</td>
<td>Renal transplant</td>
<td>Blood</td>
<td>N. farcinica</td>
<td>Nocardia farcinica ATCC 3318</td>
<td>100</td>
<td>N. farcinica</td>
<td>N. farcinica</td>
<td>99</td>
</tr>
<tr>
<td>12</td>
<td>260111/2010</td>
<td>Punjab</td>
<td>75/F</td>
<td>Cough, dyspnoea</td>
<td>Not known</td>
<td>Sputum</td>
<td>N. otitidiscaviarum</td>
<td>N. otitidiscaviarum ATCC 14629</td>
<td>100</td>
<td>N. otitidiscaviarum</td>
<td>N. farcinica</td>
<td>99</td>
</tr>
<tr>
<td>13</td>
<td>N3/2010</td>
<td>Chandigarh</td>
<td>63/M</td>
<td>Fever, pain abdomen</td>
<td>Not known</td>
<td>Aspirate from abdominal abscess</td>
<td>N. farcinica</td>
<td>N. farcinica DSM 43665</td>
<td>100</td>
<td>N. farcinica</td>
<td>N. farcinica</td>
<td>99</td>
</tr>
<tr>
<td>14</td>
<td>260055/2012</td>
<td>Karnataka</td>
<td>58/M</td>
<td>Fever, cough, dyspnoea</td>
<td>Not known</td>
<td>Sputum</td>
<td>N. farcinica</td>
<td>N. farcinica DSM 43665</td>
<td>100</td>
<td>N. farcinica</td>
<td>N. farcinica</td>
<td>99</td>
</tr>
<tr>
<td>15</td>
<td>260056/2012</td>
<td>Karnataka</td>
<td>50/M</td>
<td>Fever, cough, dyspnoea</td>
<td>Type II Diabetes mellitus, oral corticosteroids</td>
<td>BAL</td>
<td>N. asiatica</td>
<td>N. asiatica DSM 44668</td>
<td>100</td>
<td>N. asiatica</td>
<td>N. asiatica</td>
<td>99</td>
</tr>
<tr>
<td>Sample no.</td>
<td>NCCPF no./year of isolation</td>
<td>Geographical origin</td>
<td>Age/sex</td>
<td>Chief complaints</td>
<td>Predisposing factor</td>
<td>Sample/site of isolation</td>
<td>16S rRNA gene</td>
<td>Reference strain</td>
<td>Similarity (%)</td>
<td>Hsp65</td>
<td>Reference strain</td>
<td>Similarity (%)</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------</td>
<td>---------------------</td>
<td>---------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>-------------------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>-------</td>
<td>-----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>16</td>
<td>260075/2013</td>
<td>Puducherry</td>
<td>38/M</td>
<td>Multiple discharging sinuses on foot, swelling</td>
<td>Not known</td>
<td>Discharge from foot</td>
<td><em>N. brasiliensis</em></td>
<td>ATCC 700358</td>
<td>99</td>
<td><em>N. brasiliensis</em></td>
<td>ATCC 700358</td>
<td>99</td>
</tr>
<tr>
<td>17</td>
<td>260076/2013</td>
<td>Maharashtra</td>
<td>48/M</td>
<td>Multiple discharging sinuses on foot</td>
<td>Not known</td>
<td>Discharge from foot</td>
<td><em>N. brasiliensis</em></td>
<td>ATCC 700358</td>
<td>99</td>
<td><em>N. brasiliensis</em></td>
<td>ATCC 700358</td>
<td>99</td>
</tr>
<tr>
<td>18</td>
<td>260077/2013</td>
<td>Maharashtra</td>
<td>58/M</td>
<td>Pain, decreased vision, watering</td>
<td>Not known</td>
<td>Corneal scraping BAL</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
<td><em>N. farcinica</em></td>
<td>ATCC 3308</td>
<td>100</td>
</tr>
<tr>
<td>19</td>
<td>260078/2013</td>
<td>Chandigarh</td>
<td>60/M</td>
<td>Fever, cough, dyspnoea</td>
<td>Not known</td>
<td></td>
<td><em>N. cyriacigeorgica</em></td>
<td>ATCC 14759</td>
<td>99</td>
<td><em>N. cyriacigeorgica</em></td>
<td>ATCC 14759</td>
<td>99</td>
</tr>
<tr>
<td>20</td>
<td>260115/2013</td>
<td>Chandigarh</td>
<td>60/M</td>
<td>Fever, haemoptysis, dyspnoea</td>
<td>COPD</td>
<td>Lung FNAC</td>
<td><em>N. otitidisaviarum</em></td>
<td>ATCC 14629</td>
<td>100</td>
<td><em>N. otitidisaviarum</em></td>
<td>ATCC 14629</td>
<td>98</td>
</tr>
<tr>
<td>21</td>
<td>260116/2013</td>
<td>Chandigarh</td>
<td>41/M</td>
<td>Fever, cough, dyspnoea</td>
<td>Oral corticosteroids</td>
<td>ET aspirate</td>
<td><em>N. cyriacigeorgica</em></td>
<td>ATCC 14759</td>
<td>99</td>
<td><em>N. cyriacigeorgica</em></td>
<td>cyriacigeorgica DSM 44484</td>
<td>99</td>
</tr>
<tr>
<td>22</td>
<td>260117/2013</td>
<td>Chandigarh</td>
<td>66/M</td>
<td>Fever, cough, dyspnoea</td>
<td>COPD</td>
<td>Sputum</td>
<td><em>N. otitidisaviarum</em></td>
<td>ATCC 14629</td>
<td>99</td>
<td><em>N. otitidisaviarum</em></td>
<td>ATCC 14629</td>
<td>98</td>
</tr>
<tr>
<td>23</td>
<td>260118/2013</td>
<td>Chandigarh</td>
<td>36/M</td>
<td>Fever, cough, dyspnoea</td>
<td>Not known</td>
<td>Sputum</td>
<td><em>N. cyriacigeorgica</em></td>
<td>ATCC 14759</td>
<td>99</td>
<td><em>N. cyriacigeorgica</em></td>
<td>ATCC 14759</td>
<td>99</td>
</tr>
<tr>
<td>24</td>
<td>260119/2013</td>
<td>Chandigarh</td>
<td>64/F</td>
<td>Cough, dyspnoea</td>
<td>Not known</td>
<td>Sputum</td>
<td><em>N. amamensis</em></td>
<td>DSM 43665</td>
<td>99</td>
<td><em>N. amamensis</em></td>
<td>OAHPP13790</td>
<td>99</td>
</tr>
<tr>
<td>25</td>
<td>260121/2013</td>
<td>Chandigarh</td>
<td>64/M</td>
<td>Fever, dyspnoea</td>
<td>Renal transplant</td>
<td>Sputum</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
</tr>
<tr>
<td>26</td>
<td>260120/2013</td>
<td>Chandigarh</td>
<td>48/M</td>
<td>Fever, vomiting, seizures</td>
<td>Not known</td>
<td>Brain abscess Blood</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
</tr>
<tr>
<td>27</td>
<td>260122/2013</td>
<td>Chandigarh</td>
<td>59/M</td>
<td>Fever, chills</td>
<td>Not known</td>
<td>Blood</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
</tr>
<tr>
<td>28</td>
<td>260123/2013</td>
<td>Chandigarh</td>
<td>2/M</td>
<td>Fever, dyspnoea, chills</td>
<td>ALL</td>
<td>Blood</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
<td><em>N. farcinica</em></td>
<td>DSM 43665</td>
<td>99</td>
</tr>
<tr>
<td>29</td>
<td>260081/2014</td>
<td>Chandigarh</td>
<td>60/M</td>
<td>Fever, pain and swelling over thigh</td>
<td>Not known</td>
<td>Aspirate from abscess in thigh</td>
<td><em>N. beiingensis</em></td>
<td>DSM 44636</td>
<td>100</td>
<td><em>N. asiatica</em></td>
<td>DSM 44668</td>
<td>99.1</td>
</tr>
<tr>
<td>30</td>
<td>260082/2014</td>
<td>Chandigarh</td>
<td>75/M</td>
<td>Cough, dyspnoea</td>
<td>ABPA on oral corticosteroids</td>
<td>Lung FNAC</td>
<td><em>N. cyriacigeorgica</em></td>
<td>ATCC 14759</td>
<td>99</td>
<td><em>N. cyriacigeorgica</em></td>
<td>03-23155</td>
<td>100</td>
</tr>
</tbody>
</table>

*The larger segment of the 16S rRNA gene sequence was used for the analysis.
ET, endotracheal aspirate; BAL, bronchoalveolar lavage; FNAC, fine-needle aspiration cytology; COPD, chronic obstructive pulmonary disorder; ALL, acute lymphocytic leukaemia; ABPA, allergic bronchopulmonary aspergillosis.
in the hsp65 sequences compared with the 16S rRNA gene sequence. Fig. 3 shows the genotyping of N. cyriacigeorgica illustrating three separate clades. Our N. cyriacigeorgica isolates belonged to genotype 1 (n=4) and genotype 2 (n=2).

**DISCUSSION**

The epidemiology of nocardiosis is evolving. Due to better identification methods, the number of Nocardia spp. implicated in human infections has increased (Schlaberg *et al.*, 2005).
The description by some studies favours the increased number of infections by *Nocardia* spp. whilst others consider the cases to be constant in number (Baio et al., 2013). The exact prevalence of nocardiosis in India is not clearly known. The few available case series have shown that the prevalence varies from 1.37 to 1.9% (Gaude et al., 1999; Tendolkar et al., 1994; Tharayil John et al., 2003; Shivaprakash et al., 2007). On average, the NCCPF receives about 10 isolates with presumptive identification as *Nocardia* spp. of which many are finally identified as *Actinomycetes* other than *Nocardia*. The lung is the major organ primarily affected, accounting for...
60–70 % of all human nocardiosis (Minero et al., 2009). Mortality due to nocardiosis ranges from 40 to 80 % depending on the immune status of the patient and intervention therapy (Knollmann et al., 2000; Martinez Tomás et al., 2007; Minero et al., 2009; Peleg et al., 2007; Zeidman et al., 1995). Nocardia spp. can cause disseminated, cutaneous and rarely catheter-related infection and healthcare-associated diseases (Brown-Elliott et al., 2006).

Of all the molecular identification methods, DNA–DNA hybridization was considered the ‘gold standard’ for speciation of prokaryotes (Brown-Elliott et al., 2006). Currently, 16S rRNA gene sequencing is considered a better method for identification of Nocardia to the species level (Borriello et al., 2007), although certain isolates that have been found to have a high sequence similarity (99.8 %) from 16S rRNA sequences are recorded as distinct species by DNA–DNA hybridization (Conville et al., 2004; Rodriguez-Nava et al., 2006). McTaggart et al. (2010) applied multilocus sequence analysis to 190 clinical isolates of Nocardia spp. using sequences of 16S rRNA, hsp65, RNA polymerase (rpoB), gyrase B of the β-subunit of DNA topoisomerase (gyrB) and SecA pre-protein translocase-subunit A (secA1) genes and found this technique to be reliable for accurate speciation of Nocardia (McTaggart et al., 2010). Due to its high discriminatory power and accuracy, multilocus sequence analysis has replaced DNA–DNA hybridization in taxonomy and speciation and for the study of intrageneric relationships (Shneyer, 2007). In spite of the availability of many other molecular methods such as PCR restriction enzyme pattern analysis (Rodriguez-Nava et al., 2006), PCR-RFLP (Conville et al., 2000; Steingrube et al., 1995, 1997) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) (Segawa et al., 2015), gene sequencing remains the best method of identification of Nocardia spp. (McTaggart et al., 2010). Although MALDI-TOF MS is evolving as an important technique for the identification of Nocardia, further refinement is essential to finally accept this technique for routine identification in the clinical laboratory.

It has been proposed that the combination of two or more gene sequences may increase the confidence level of species-specific identification and even the detection of new species (McTaggart et al., 2010). In the present study, we compared the utility of sequencing the most promising genes, hsp65 and the 16S rRNA gene, for identification of 30 Indian clinical isolates of Nocardia spp. We chose these two regions, as the sequence data for these two genes are available (although there are limited data for hsp65) in the public databases for most of the Nocardia spp. and both of these regions are efficient in identification of species. On comparison of these two gene sequences, the mean percentage dissimilarity was found to be higher with the hsp65 gene sequences (4 %, range 0–7.9 %) than the 16S rRNA gene sequences (2.3 %, range 0–8.9 %). The dissimilarity range was higher in previous studies by Rodriguez-Nava et al. (2006) (44 standard strains and 21 clinical isolates; hsp65 sequence, 0–12 %; 16S rRNA gene sequence, 0–9.5 %) and McTaggart et al. (2010) (47 standard strains and 190 clinical isolates; hsp65 sequence, 0–13 %; 16S rRNA gene sequence, 0–9.2 %). We encountered less diversity in our Nocardia isolates compared with these two studies, possibly due to the lower number of isolates included in the present study.

![Fig. 3. Unrooted phylogenetic tree showing the genotypes of N. cyriacigeorgica using hsp65 gene sequence data. The tree was reconstructed by the neighbour-joining method with bootstrap resampling (1000 replicates). N. otitidiscaviarum DSM 44565 was used as the outgroup. The scale bar represented 0.005 substitutions per nucleotide position.](image-url)
Both 16S rRNA and hsp65 gene sequences have several conflicting claims in speciating Nocardi a. Although the complete 16S rRNA gene sequence (1494 bp) exhibits sequence heterogeneity, resulting in identification of distinct species (Roth et al., 2003), analysis of a partial sequence of the 16S rRNA gene (500 bp) is recommended for species identification in clinical settings (Conville & Witebsky, 2005; Patel et al., 2004). One of the major drawbacks in BLAST analysis is misinterpretation of the results due to inaccurate sequences submitted as Nocardia spp. to GenBank (Mellmann et al., 2003; Patel et al., 2000, 2004; Turenne et al., 2001). The RIDOM and MicroSeq 16S rDNA gene databases, which also include hypervariable regions, may be used in such situations for accurate identification of Nocardia spp. (Cloud et al., 2004; Mellmann et al., 2003). However, these databases contain information for a limited number of species (≤30 species). In the present study, both hsp65 and 16S rRNA gene sequences were equally reliable in identifying 90 % of our Nocardia spp., including N. farcinica (n = 10), N. cyriacigeorgica (n = 6), N. oitidiscaviarum (n = 5), N. brasiliensis (n = 3), N. pneumoniae (n = 2) and N. amamensis (n = 1). Although an isolate of N. asiatica (NCCPF 260056) was accurately identified (99–100%) by both these genes, hsp65 alone could not be relied on for identification of this species, as this gene identified N. bei jingensis (NCCPF 260081) as N. asiatica. Two isolates that were inconclusive by analysis of the short segment of the 16S rRNA gene (605 bp) and hsp65 could be resolved by sequencing a larger fragment (~1000 bp) of the 16S rRNA gene. Both of these isolates were identified as N. bei jingensis with a similarity of 99.8 and 100 % with the standard strain (Table 1). As few hsp65 sequences of N. bei jingensis (n = 4), N. arthritidis (n = 2) and N. asiatica (n = 13) are available in GenBank, caution should be taken before concluding on the species based on this gene. Another important finding of our study was identifying the heterogeneity within individual species in hsp65 gene sequence analysis. No heterogeneity was observed in the 16S rRNA gene sequence, which is in contrast to the earlier observation of Patel et al. (2004) for N. oitidiscaviarum. These authors also observed heterogeneity in this region in N. nova and N. transl velensis, which were not evaluated in the present study. In our study, heterogeneity was observed within the hsp65 gene sequence of N. cyriacigeorgica. Intraspecies microheterogeneity in N. cyri acigeorgica has been reported previously by Schlaberg et al. (2008) with more polymorphisms in hsp65 than in the 16S rRNA sequence. These variations have been used to describe three genotypes within this species (Schlaberg et al., 2008). We had six strains of N. cyriacigeorgica, isolated only from respiratory samples. In the study by Schlaberg et al. (2008), the majority of N. cyriacigeorgica was isolated from respiratory samples (5/7). Our isolates belonged to genotype 1 (n = 4) and genotype 2 (n = 2). Due to hypervariable regions in the hsp65 gene, we recommend this gene for genotyping of N. cyriacigeorgica and the 16S rRNA gene for species identification. N. cyriacigeorgica is an emerging species causing a wide variety of infections including brain abscess, pulmonary infection, sepsis and disseminated disease. The data for infections caused by N. cyriacigeorgica are available in the form of case reports and case series. This species appears to be the commonest cause of nocardiosis, especially in Europe, Asia and North America (Baio et al., 2013; Brown-Elliott et al., 2006; Schlaberg et al., 2008). From India, N. cyri acigeorgica has only been reported from ocular infections (Lalitha et al., 2007; Rath et al., 2012). This species accounted for the majority (35 %) of ocular infection reported by Lalitha et al. (2007). Identification of this species in clinical settings is important as it is reported to possess some degree of resistance to cotrimoxazole and minocycline (Minero et al., 2009). Intraspecies variation in susceptibility to imipenem has also been reported in N. cyriacigeorgica (Schlaberg et al., 2014). Recently, genomic sequencing of N. cyriacigeorgica (GUH-2 strain) revealed virulence determinants such as the production of superoxide dismutase and peptide synthesis probably for siderophore/toxin production (Zoropogui et al., 2013).

The incorrect deposition of the 16S rRNA gene sequences and the low availability of hsp65 sequences (total n = 429; N. cyriacigeorgica, n = 79; N. farcinica, n = 56; N. seriole, n = 29; N. nova, n = 27; N. oitidiscaviarum, n = 26; N. brasiliensis, n = 20; N. asteroides, n = 13; N. asiatica, n = 13) for most of the Nocardia spp. in public databases may result in failure of accurate identification. The majority (90 %) of Nocardia spp. (N. farcinica, N. cyriacigeorgica, N. oitidiscaviarum, N. brasiliensis, N. pneumoniae and N. am amensis) isolated from clinical specimens may be identified successfully using either a partial sequence of the 16S rRNA gene or the hsp65 gene. If BLAST analysis results in the identification of N. bei jingensis using the 16S rRNA gene (short fragment) or N. arthritidis/N. asiatica for hsp65 gene analysis, sequencing of a larger fragment of the 16S rRNA gene or other reported genes should be performed. The hsp65 gene sequences may also be exploited for molecular typing of N. cyriacigeorgica spp.

The present study was a case series of 30 different human Nocardia spp. infections occurring in India in different parts of the country over a period of 8 years. This study strengthens the need to use molecular methods for the identification of Nocardia to the species level by identifying all strains by sequencing of the 16S rRNA and hsp65 genes. The results highlight the wide variety of Nocardia spp. received at our centre. Further studies on large collections of Nocardia spp. will be necessary to understand the exact prevalence of Nocardia spp. in India.

ACKNOWLEDGEMENTS

We duly acknowledge the Indian Council of Medical Research (ICMR), New Delhi, India, for the financial support for the research proposal and providing financial aid. There are no conflicts of interest to declare for any of the authors in the study.

REFERENCES


Molecular identification of *Nocardia* spp.


