Infection, Genetics and Evolution xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Genetic characterization of UCS region of *Pneumocystis jirovecii* and construction of allelic profiles of Indian isolates based on sequence typing at three regions

Rashmi Gupta [a,1], Bijay Ranjan Mirdha [a,*], Randeep Guleria [b], Lalit Kumar [c], Kalpana Luthra [d], Sanjay Kumar Agarwal [e], Vishnubhatla Sreenivas [f]

[a] Department of Microbiology, All India Institute of Medical Sciences, New Delhi, India
[b] Department of Medicine, All India Institute of Medical Sciences, New Delhi, India
[c] Department of Medical Oncology, All India Institute of Medical Sciences, New Delhi, India
[d] Department of Biochemistry, All India Institute of Medical Sciences, New Delhi, India
[e] Department of Nephrology, All India Institute of Medical Sciences, New Delhi, India
[f] Department of Biostatistics, All India Institute of Medical Sciences, New Delhi, India

ARTICLE INFO

Article history:
Received 22 January 2012
Received in revised form 29 March 2012
Accepted 25 July 2012
Available online xxxx

Keywords:
Pneumocystis jirovecii
Multilocus typing
UCS
Allelic profiles

ABSTRACT

Pneumocystis jirovecii is an opportunistic pathogen that causes severe pneumonia in immunocompromised patients. To study the genetic diversity of *P. jirovecii* in India the upstream conserved sequence (UCS) region of *Pneumocystis* genome was amplified, sequenced and genotyped from a set of respiratory specimens obtained from 50 patients with a positive result for nested mitochondrial large subunit ribosomal RNA (mtLSU rRNA) PCR during the years 2005–2008. Of these 50 cases, 45 showed a positive PCR for UCS region. Variations in the tandem repeats in UCS region were characterized by sequencing all the positive cases. Of the 45 cases, one case showed five repeats, 11 cases showed four repeats, 29 cases showed three repeats and four cases showed two repeats. By running amplified DNA from all these cases on a high-resolution gel, mixed infection was observed in 12 cases (26.7%, 12/45). Forty three of 45 cases included in this study had previously been typed at mtLSU rRNA and internal transcribed spacer (ITS) region by our group. In the present study, the genotypes at those two regions were combined with UCS repeat patterns to construct allelic profiles of 43 cases. A total of 36 allelic profiles were observed in 43 isolates indicating high genetic variability. A statistically significant association was observed between mtLSU rRNA genotype 1, ITS type Ea and UCS repeat pattern 4.

1. Introduction

Pneumocystis jirovecii is an opportunistic fungal agent that causes severe pneumonia in individuals infected with human immunodeficiency virus (HIV) and in other immunocompromised patients. Due to lack of a suitable *in-vitro* culture system for *P. jirovecii*, epidemiology of this opportunistic pathogen is mostly investigated by using various molecular typing approaches. Such studies describing molecular types of *P. jirovecii* from different parts of the world have greatly increased our present understanding of transmission and evolution of this pathogen (Beard et al., 2000; Miller et al., 2002; Robberts et al., 2004).

A number of loci of *Pneumocystis* genome, including mitochondrial large subunit rRNA (mtLSU rRNA), mitochondrial small subunit rRNA (mtSSU rRNA), internal transcribed spacer (ITS) regions of nuclear ribosomal RNA operon and dihydropteroate synthase (DHPS) have been used to study the genetic diversity of *P. jirovecii*. However, the identification of the best gene targets and best approaches for molecular typing of *P. jirovecii* remains a matter of debate. mtLSU rRNA and ITS are two loci of *Pneumocystis* that are most commonly used to study genetic diversity of *Pneumocystis*. Sequencing of mtLSU rRNA has generated useful data for addressing specific epidemiological questions, like transmission of *Pneumocystis* (Rivero et al., 2008) and geographical variations (Miller et al., 2005; Nahimana et al., 2000; Wakefield et al., 1994) in *P. jirovecii* isolates. In contrast, sequence analysis of ITS region has been useful in detecting a much higher level of intra-species variation due to greater level of sequence variation (Lee et al., 1998). In 2002, a new method for typing of *P. jirovecii* was described based on variations in the tandem repeats in upstream conserved sequence (UCS) of major surface glycoprotein (MSG) expression site (Ma et al., 2002). In comparison to most of the other typing methods, which are based on identification of single nucleotide polymorphisms (SNPs) or combinations of such SNPs, UCS typing method is based on characterization of the size of UCS

* Corresponding author. Address: Department of Microbiology, All India Institute of Medical Sciences, New Delhi 110029, India. Tel.: +91 011 26593288; fax: +91 26588641.
E-mail addresses: rashmi21481@gmail.com (R. Gupta), mirdhabr@hotmail.com (B.R. Mirdha).
[1] Present address: Department of Pathology and Lab Medicine, University of British Columbia, Vancouver, Canada.

1567-1348/$ - see front matter © 2012 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.meegid.2012.07.013

region of the MSG expression site. The expression site or UCS of MSG gene contains a 10 base repeat sequence, with variations in the number of repeats (2–6 repeats) among different isolates as well as in single isolates (Kutty et al., 2001). Such methods, known as variable number of tandem repeats (VNTR), short-sequence repeats or micro-satellite analysis, have been widely used for molecular typing of prokaryotic as well as eukaryotic organisms because of their technical simplicity and high discriminatory power (Nakamura et al., 1987; van Belkum et al., 1998).

Another approach to investigate the epidemiology of *P. jirovecii* is multilocus genotyping (MLG) analysis. MLG offers a greater sensitivity of detecting differences between isolates than genotyping at single locus. In addition, it helps in achieving strain identification and strain typing. Various groups have used MLG to investigate the genetic, temporal and geographical diversity *P. jirovecii* isolates (Esteves et al., 2011, 2010a, 2008; Totet et al., 2003; Tsolaki et al., 1998). Some groups have also used MLG to study the relationship between *P. jirovecii* diversity and clinical data and outcome of patients (Matos et al., 2003; Totet et al., 2003).

In our laboratory, we have previously typed the *P. jirovecii* isolates by sequencing at mtLSU rRNA and ITS regions and observed significant geographical differences (Gupta et al., 2010, 2011). The aim of the present study was to characterize same isolates at the UCS region and to determine their UCS types. The study also aimed to determine the allelic profiles of Indian isolates of *P. jirovecii* based on genotypes at mtLSU rRNA, ITS and UCS regions and to investigate potential correlations between these profiles and disease severity and/or outcome.

2. Materials and methods

2.1. Patients and samples

The present study was performed on 57 respiratory specimens including bronchoalveolar lavage fluid (BALF), sputum and tracheal aspirate (TA) from 50 cases positive for *P. jirovecii*. All these 50 cases were positive by a nested PCR at mtLSU rRNA region of *P. jirovecii* in a previously conducted study on 395 respiratory clinical specimens that were obtained from 323 HIV-infected and HIV-uninfected immunocompromised patients with a high index of clinical suspicion of PCP. These 395 specimens included 174 BALF, 100 sputa, 40 TA, 75 naso-pharyngeal aspirates and six gastric aspirates. All the patients were given an identification, which consisted of abbreviation "PJ" for *P. jirovecii* followed by a number, corresponding to arrival of respiratory samples in the laboratory. Table 1 summarizes underlying diseases, types of respiratory samples and corresponding genotypes at mtLSU rRNA and ITS regions for the fifty patients positive for *P. jirovecii*.

Of the 50 cases, 43 cases were also positive by nested PCR for ITS PCR. Detailed clinical information including symptoms at presentation, chest X-ray and/or HRCT findings, laboratory results including CD4+ T cell counts in HIV-infected individuals, arterial blood gas (ABG) analysis, simultaneous presence of other organisms in the respiratory specimens, treatment against other infections, anti-*Pneumocystis* prophylaxis and treatment, and clinical outcome was obtained from most of the cases. Based on above information, the PCR positive cases were considered to have clinical PCP if any two of the following-mentioned conditions were fulfilled:

1. Clinical findings were consistent with PCP.
2. No other microbial pathogen/s was/were isolated from respiratory samples.
3. Empiric anti-microbial therapy included an anti-*Pneumocystis* agent that led to the resolution of clinical symptoms.

Based on their clinical information and the above-mentioned criteria for clinical PCP, 48 of the 50 cases were found to have clinical PCP.

2.2. DNA extraction and PCR for UCS region

All the specimens except sputa samples were centrifuged at 4000 rpm at 4 °C for 10 min. Sputa samples were first treated with 0.0065 M dithiothreitol (DTT), a mucolytic agent and then centrifuged. The pellet was re-suspended in 1/5th of the supernatant. DNA extraction was performed using Qiagen DNeasy tissue kit (Qiagen Inc., Valencia, CA, USA) following manufacturer's instructions with some modifications. Briefly, 200 µl of pellet was lysed in 200 µl of a sample specific lysis buffer containing 50 mM KCl, 15 mM Tris–HCl (pH 8.3), 0.5% NP-40 and 500 µg of proteinase K. The above mixture was incubated at 56 °C for 45 min. The rest of steps were followed according to the manufacturer's instructions. PCR assay for UCS region consisted of two rounds of amplification. External round was performed using primers ML40 (5′-TTCAGCG-CAGGTTGGTTG-3′), corresponding to the nucleotides 1353–1370 and ML664 (5′-CGAGGCTCCCCCAAATG-3′), corresponding to the nucleotides 1961–1977 and nested round was performed using primers, JQ8 (5′-CTGTGGATTGAGCTATTTCTTGTATCTATGCGCT-3′), corresponding to the nucleotides 1489–1522 and JQ14 (5′-GCATG-CAAGCTGACATTCCGCGCAAAAATAAGC ACT-3′), corresponding to the nucleotides 1577–1612) (Ma et al., 2002). The first round was a touch-down PCR with a 1 °C decrease in each cycle followed by 25 cycles of annealing at 55 °C. For nested PCR, the external PCR products were diluted 1:10 times and were used as the template in the nested round. Nested PCR consisted of 25 cycles of denaturation at 94 °C, annealing at 65 °C and extension at 72 °C. The UCS external PCR assay amplified a product of nearly 600 bases and nested round amplified a product of nearly 123 bases.

2.3. Prevention of false positive cases

To minimize cross-contamination, different procedures such as DNA extraction, PCR amplification, addition of DNA into the master mix, gel run and visualization were performed in different rooms or in different areas of the laboratory. Filter barrier pipette tips were used at every stage of reagent transfer. Multiple negative controls (autoclaved double distilled water or DNA from negative samples and the PCR mixture minus the DNA template) were tested in each run, especially in nested rounds. In case of discrepant results in PCR assays, all the procedures were repeated twice with fresh reagents to re-confirm the results.

2.4. Sequencing of UCS region

The nested PCR products were purified prior to sequencing by using a PCR purification kit by Qiagen (QIAquick PCR purification kit, Qiagen Inc., Valencia, CA, USA). Two cases, which showed some non-specific amplification, were purified by gel extraction kit from Qiagen. Sequencing was performed at both the strands using UCS forward and reverse primers (JQ8/JQ14) on ABI 2720 thermal cycler (ABI Biosystems, California, USA). Each 0.25× reaction was prepared by adding 2.0 µl of ready reaction mix (Big Dye, containing all the four dye-labeled di-deoxy chain terminators, dNTPs and enzyme), 1.0 µl of 5× dilution buffer, 3.2 pmol of primer (JQ8/JQ14) and 20–50 ng of template DNA in a reaction of 10 µl. The contents were mixed well, spun briefly before putting the tube in the thermal cycler and the volume was set to 10 µl. The cycle sequencing reaction consisted of 25 cycles of 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 min. As the excess dye terminators in the sequencing reactions obscure bases in the early part of the

Please cite this article in press as: Gupta, R., et al. Genetic characterization of UCS region of *Pneumocystis jirovecii* and construction of allelic profiles of Indian isolates based on sequence typing at three regions. Infect. Genet. Evol. (2012), http://dx.doi.org/10.1016/j.meegid.2012.07.013