



**Karolinska  
Institutet**

**Institutionen för Mikrobiologi, Tumör och Cell Biologi**

## **Genetic and antigenic diversity in *Pneumocystis jirovecii***

**AKADEMISK AVHANDLING**

som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i Gard Aulan, Smittskyddsinstitutet, Nobels v 18, Solna

**Fredagen den 16 september, 2011, kl 09.00**

av

**Jessica Beser**

*Huvudhandledare:*

Dr Johan Lindh  
Smittskyddsinstitutet  
Avdelningen för Analys och Prevention

*Bihandledare:*

Dr Per Hagblom  
Uppsala

Dr Silvia Botero-Kleiven  
Smittskyddsinstitutet  
Avdelningen för Diagnostik och Vaccin

Docent Qijun Chen  
Karolinska Institutet  
Institutionen för Mikrobiologi, Tumör och Cell  
Biologi

*Fakultetsopponent:*

Professor James Stringer  
University of Cincinnati  
Department of Molecular Genetics,  
Biochemistry and Microbiology  
United States of America

*Betygsnämnd:*

Professor Mats Kalin  
Karolinska Institutet  
Institutionen för Medicin

Professor Jan Albert  
Karolinska Institutet  
Institutionen för Klinisk Mikrobiologi

Docent Christina Welinder Olsson  
Sahlgrenska Universitetssjukhuset  
Institutionen för Klinisk Bakteriologi

**Stockholm 2011**

# ABSTRACT

*Pneumocystis jirovecii* is a human specific, atypical fungus with a worldwide distribution that causes disease in immunocompromised individuals. The fungus proliferates in the lungs where it binds to epithelial alveolar cells, provoking severe pneumonia, denoted pneumocystis pneumonia (PCP). As there is no *in vitro* culture system for the organisms, and no morphological means to differentiate between *P. jirovecii* strains, we have used molecular tools to study the fungus in patient samples. For this purpose we have been targeting different loci in the *P. jirovecii* genome (ITS, DHPS and MSG) to address different aspects of *P. jirovecii* infections. The nucleotide sequence of the internal transcribed spacers (ITS) in *P. jirovecii* has been useful for isolate genotyping. We investigated the genetic diversity in Sweden by analyzing 408 cloned ITS sequences, from 64 clinical specimens. Several globally common haplotypes (combination of ITS1 and ITS2) and one local ITS2 were found. No correlations between certain haplotypes and patient characteristics or geographical associations were uncovered. In this context, a model describing the genealogic relationships of the strains was presented. During this process, we found that the typing system was generating artifactual sequences. We established a set of criteria to determine “*bona fide*” haplotypes, and optimized the typing method to avoid the generation of artifactual recombinants. These improved tools have enabled a more correct assessment of the overestimated genetic diversity of *P. jirovecii* populations. Trimethoprim-sulfamethoxazole (TMP-SMX) is the most widely used drug for prevention and treatment of PCP. Non-synonymous substitutions in the dihydropteroate synthase (DHPS) gene of *P. jirovecii* have been found to be associated with sulphamethoxazole exposure. It has been suggested that this is the result of the fungus developing resistance towards the drug. We conducted a study to investigate the presence of *P. jirovecii* DHPS mutations in the Swedish population and 104 specimens from patients with a suspected PCP were screened. All of the specimens (100%) showed a wild-type DHPS pattern, suggesting that there is no, or a very low prevalence of, DHPS mutations in the country. One surface molecule of *P. jirovecii* with a probable key function in the colonization of the alveoli and in immune evasion is the major surface glycoprotein (MSG). The MSGs are encoded by the *msg*-gene family, and transcription is limited to a single *msg*-gene located in a unique expression site. To investigate the expressed *msg*-genes and the extent of the variability of the MSG antigen, we analyzed *msg*-genes located at the expression site. First, we analyzed a short segment of the 5'-end of the *msg*-genes in 13 patient samples. Second, we extended these studies to two full-length *msg*-sequences from two different patients. We concluded, from these analyses, that there is considerable variation in the potentially expressed MSG-proteins, but that a substantial amount of conservation can be found in the *msg*-gene family, even in samples of unrelated origins. In conclusion, the complexity of *P. jirovecii* populations has been overestimated but typing fidelity can easily be improved. The numbers of ITS haplotypes in Sweden are restricted, and a model depicting the relationships between strains is proposed. Furthermore, *P. jirovecii* DHPS mutations are very rare or possibly even absent in Sweden. Finally, the expressed *msg*-genes display both a remarkable variation and conservation.

ISBN 978-91-7457-434-0