Original Article

Update of phylogenetic and genetic diversity of *Sporothrix schenckii* sensu lato

Lucía Rangel-Gamboa¹, Fernando Martínez-Hernandez¹, Pablo Maravilla¹, Roberto Arenas-Guzmán² and Ana Flisser³.*

¹Departamento de Ecología de Agentes Patógenos, Subdirección de Investigación, Hospital General Dr. Manuel Gea González, Distrito Federal, Mexico, ²Servicio de Micología, Hospital General Dr. Manuel Gea González, Distrito Federal, Mexico and ³Departamento de Microbiología y Parasitología, Facultad de Medicina, Universidad Nacional Autónoma de México, Mexico 04510Distrito Federal, Mexico.

*To whom correspondence should be addressed: Ana Flisser, Departamento de Microbiología y Parasitología, Facultad de Medicina, Universidad Nacional Autónoma de México, Mexico 04510 Distrito Federal, Mexico. Tel: +52 55 56223466; Fax: +52 55 56232382; E-mail: flisser@unam.mx

Received 6 July 2015; Revised 21 September 2015; Accepted 5 October 2015

Abstract

*Sporothrix schenckii sensu lato* causes subcutaneous mycosis. In this article we analysed its phylogeny and genetic diversity using calmodulin DNA sequences deposited in GenBank database. Population genetics indices were calculated, plus phylogenetic and haplotype network trees were built. Five clades with high values of posterior probability, 47 haplotypes and high diversity in the complex were found. Analysis of partial calmodulin sequences alignment revealed conserved and polymorphic regions that could be used as reference for taxonomic identification. The use of population genetics analysis allowed understanding the phylogenetic proximity of *S. schenckii s. str.* and *S. brasiliensis*; scarce genetic flow among them with low migration index and high ancestry coefficient was found. Similarly, *S. globosa, S. mexicana* and *S. pallida* sequences showed highly differentiated species with no genetic exchange. The phylogenetic tree suggests that *S. mexicana* shared a common ancestor with *S. pallida*; while *S. globosa* and *S. brasiliensis* are more related to *S. schenckii s. str.* and showed less haplotype diversity and restrictions in geographic distribution. In the haplotype network tree *S. schenckii s. str.* species displayed worldwide distribution without dispersion centres; while *S. brasiliensis* and *S. globosa*, exhibited Brazil and Euro-Asia as dispersion centres, respectively. Our data suggest that *S. schenckii* complex has been submitted to a divergent evolution process, probably due to the pressure of the environment and of the host. In contrast, *S. brasiliensis* could have been submitted to purifying selection or expansion process.

**Key words:** Calmodulin, genetic population structure, phylogeny, purifying selection, *Sporothrix schenckii*.

Introduction

Sporotrichosis is a subacute or chronic mycosis that has been reported worldwide, especially in temperate, warm, tropical, and subtropical areas with Latin America reporting more cases in the last decades.¹,²,³ This disease is acquired by manipulation of contaminated soil or organic material like decaying vegetation or sphagnum moss; it is also transmitted by contact with mammals, especially...
Lesions are frequently restricted to skin, subcutaneous tissues, and adjacent lymphatic vessels, but there are reports of primary lung, osteoarticular and conjunctival infections as well as disseminated ones. In general, the Sporothrix schenckii sensu lato (s. l.) includes dimorphic fungi that develop mycelia in a saprophytic phase, whereas the invasive form or during special laboratory conditions yeast development is required. Previously, Ophiostoma stenoceras was considered the teleomorph (sexual form) of S. schenckii, but molecular phylogenetic analyses do not support this statement, thus teleomorph is still unknown for these fungi.

During the last decades amplified fragment length polymorphisms (AFLP) and restriction fragment length polymorphisms (RFLP) were used to describe several variants associated with the geographic distribution of S. schenckii isolates. In addition, phylogenetic analyses have been performed with loci coding for internal transcribed spacer (ITS), chitin synthase, β-tubulin and calmodulin genes. Analysis of calmodulin sequences allowed considering S. schenckii as a complex of cryptic species. In human infections, S. brasiliensis (Clade I) and S. schenckii s. str. (Clade II) are the most reported of the Sporothrix complex, which also include S. globosa (Clade III) and S. luriei (Clade VI). Environmental species phylogenetically distant from S. schenckii complex are Sporothrix pallida (synonymy S. albicans; Clade V) and Sporothrix mexicana (Clade IV). Additionally, in agreement with the phylogenetic distribution, Marimon et al. described variations in conidial morphology and preferential assimilation of carbohydrates. Currently, to avoid confusion the distinction between sensu lato (s. l.) and sensu stricto (s. str.) is used; the former referring to S. schenckii in the wide sense and the latter to the species.

S. schenckii s. str. is found mainly in America, Asia, and Africa; S. globosa and S. pallida are cosmopolitan reported in Europe, Asia, and America; S. mexicana has been found in Mexico, Brazil, Italy and Portugal; while S. brasiliensis remains restricted to Brazil, prevailing in the South and Southeast and shows preferential zoonotic transmission by cats. Recently, differences in virulence and histological features related to S. schenckii s. l. were established using mouse experimental models, and concluded that S. brasiliensis is the most virulent species followed by S. schenckii, whereas S. globosa showed little virulence. S. mexicana was initially obtained from environmental sources in Mexico, and was considered saprophytic; afterward it was isolated from a human case in Portugal; from a dog in Italy; and three isolates were found in fungal collections in Brazil where they were classified as S. schenckii since 1955. S. pallida was recovered mainly from environmental samples and cannot produce melanin, it was considered a nonpathogenic mould to humans, closely related to the pathogen S. schenckii; nevertheless in 2013, it was described in a patient as an opportunistic causal agent that caused corneal ulcer; one year later, an isolate from feline sporotrichosis was reported in Brazil. When compared their antifungal drug resistance, S. mexicana was the most resistant and S. brasiliensis was the least resistant. The studies of conidial morphology and biochemical characteristics aid identifying species belonging to the genus Sporothrix; however presently, most authors agree that unequivocal identification depends on calmodulin sequence analysis.

Calmodulin is a highly conserved protein, important as enzymatic cofactor in fungal calcium-calcineurin sensing pathways, that is associated with hyphal growth, antifungal resistance, pathogenesis, and virulence. DNA calmodulin sequences have been used in fungus phylogeny to establish inter-species and intra-species relations. Therefore, in order to provide a conceptual framework for prioritizing, designing and interpreting the results of future studies, the overall goal of the present work was to review the population genetics of S. schenckii s. l. using calmodulin partial sequences recovered from the GenBank database.

**Material and methods**

In this study, 351 partial sequences of calmodulin DNA from S. schenckii s. l. isolated from Africa, Asia, Latin America, and Europe deposited in GenBank/EMBL/DDBJ international databases were analyzed for diversity and genetic structure. The sequences were S. schenckii (n = 139), S. brasiliensis (n = 58), S. globosa (n = 49), S. pallida (n = 21), unidentified Sporothrix sp. (n = 47), S. mexicana (n = 7), and S. luriei (n = 1). To determine the population genetic structure among S. schenckii s. l., the following indices were used: nucleotide diversity (π), as the average proportion of different nucleotides between all possible couples of sequences in a sample; haplotype polymorphism (θ), as the proportion of nucleotide sites expected to be polymorphic in the sample; interpopulation heterozygosity (Hs), as the fraction of individuals between or among populations that are heterozygous for a particular locus; genetic differentiation index (GST), as an estimator based on multilocus versions of the observed and expected heterozygosity or gene diversity; ancestry coefficient (FST), as the reduction in heterozygosity due to subpopulation divergence in allele frequencies, and migration index (Nm), as the expected effective migration rate given as the amount of genetic differentiation among subpopulations in the infinite island model; also Tajima’s D was calculated. All these indices were obtained with DnaSP v5 software.
Results

Nucleotide polymorphism analyses in calmodulin were performed with 322 sequences (Supplementary table 1), because 29 were shorter than 304 base pairs (bp) and were excluded in order to obtain reliable data. Our results showed 23% (148/640) polymorphic sites and 19% (120/640) informative ones; $\pi$ and $\theta$ values for all species were 0.081 and 0.089, respectively, corresponding to 55 different haplotypes ($P < .05$). $S.\text{schenckii s. str.}$ was the most polymorphic with 23, followed by $S.\text{globosa}$ with nine. In contrast, $S.\text{brasiliensis}$ showed three haplotypes, that had lower haplotype diversity, $\theta = 0.18$ and $\pi = 0.0001$ (Table 1). Populations of the $S.\text{schenckii s. l.}$ were initially studied, using the species identified in GenBank, in order to obtain $Hs$, $F_{ST}$, $G_{ST}$ and $Nm$ indices, but the data generated did not allow identifying well-defined populations. Therefore, we decided to compare these indices using the distribution of the haplotype network tree (partial information shown in Figure 1). With this method we found well defined populations with high $F_{ST}$ and $G_{ST}$, low $Nm$, and $Hs$ values were variable among populations (Table 2). Haplotype network tree showed three clades: the first one, seen on the left side of the haplotype network tree, corresponds to $S.\text{globosa}$. $S.\text{schenckii s. str.}$ occupied the central part and showed high diversity. In the right part $S.\text{brasiliensis}$ was found. All isolates reported as $Sporothrix\text{ sp.}$ were situated in one of these clades, most of them in the $S.\text{globosa}$ clade. The isolate reported as $S.\text{luriei}$ (AM747302) shown in light blue in the left upper side, between $S.\text{globosa}$ and $S.\text{schenckii s. str.}$

![Figure 1. Haplotype network analysis of $S.\text{schenckii complex}$ using partial calmodulin gene sequences. Colours represent species: purple = $S.\text{globosa}$, blue = $S.\text{schenckii s. str.}$, light blue = $S.\text{luriei}$, red = $S.\text{brasiliensis}$, and grey = median vectors.](https://academic.oup.com/mmy/article-abstract/54/3/248/2579141)
Table 2. Genetic differentiation among *S. schenckii complex* subpopulations in the infinite island model.

<table>
<thead>
<tr>
<th>Population 1</th>
<th>Population 2</th>
<th>Hs</th>
<th>FST*</th>
<th>GST</th>
<th>Nm</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. schenckii s. str.</em></td>
<td><em>S. brasiliensis</em></td>
<td>0.515</td>
<td>0.913</td>
<td>0.342</td>
<td>0.024</td>
</tr>
<tr>
<td><em>S. schenckii s. str.</em></td>
<td><em>S. mexicana</em></td>
<td>0.727</td>
<td>0.968</td>
<td>0.079</td>
<td>0.008</td>
</tr>
<tr>
<td><em>S. schenckii s. str.</em></td>
<td><em>S. pallida</em></td>
<td>0.720</td>
<td>0.959</td>
<td>0.098</td>
<td>0.011</td>
</tr>
<tr>
<td><em>S. schenckii s. str.</em></td>
<td><em>S. globosa</em></td>
<td>0.671</td>
<td>0.927</td>
<td>0.186</td>
<td>0.020</td>
</tr>
<tr>
<td><em>S. brasiliensis</em></td>
<td><em>S. mexicana</em></td>
<td>0.065</td>
<td>0.998</td>
<td>0.644</td>
<td>0.0005</td>
</tr>
<tr>
<td><em>S. brasiliensis</em></td>
<td><em>S. pallida</em></td>
<td>0.151</td>
<td>0.991</td>
<td>0.598</td>
<td>0.002</td>
</tr>
<tr>
<td><em>S. brasiliensis</em></td>
<td><em>S. globosa</em></td>
<td>0.228</td>
<td>0.970</td>
<td>0.610</td>
<td>0.007</td>
</tr>
<tr>
<td><em>S. mexicana</em></td>
<td><em>S. pallida</em></td>
<td>0.496</td>
<td>0.976</td>
<td>0.315</td>
<td>0.006</td>
</tr>
<tr>
<td><em>S. mexicana</em></td>
<td><em>S. globosa</em></td>
<td>0.458</td>
<td>0.992</td>
<td>0.225</td>
<td>0.002</td>
</tr>
<tr>
<td><em>S. pallida</em></td>
<td><em>S. globosa</em></td>
<td>0.494</td>
<td>0.985</td>
<td>0.280</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Note: Interpopulation heterozygosity (Hs), ancestry coefficient (FST), gene diversity (GST), migration index (Nm).

* The common used values for FST are as follows: 0 to 0.05, small; 0.05 to 0.15, moderate; 0.15 to 0.25, great; values above 0.25 indicate huge genetic differentiation.

Those strongly differentiated have an Nm < 1, whereas an Nm > 4 behaves as a single panmictic unit.

---

**Figure 2.** Haplotype network tree in subpopulations of *S. schenckii s. str.*, colours represent countries of isolation.

str. clades, is clearly distant from all other isolates. Tajima’s D at a range of non-overlapping window sizes, showed that *S. brasiliensis* and *S. globosa* had values of −2.41 (*P* < .01) and −2.13 (*P* < .05), respectively, suggesting a purifying selection process in these groups or a recent expansion for these populations.

Due to the high diversity observed in *S. schenckii s. str.* we performed a population analysis in this group using five clusters identified in haplotype network tree of *S. schenckii s. str.* (Figure 2), four of these populations were well supported subpopulations that showed high FST ≥ 0.36 and no genetic exchange Nm ≤ 0.44, while cluster 4 showed less differentiation with FST = 0.05 and Nm = 4.75 when compared with cluster 3 (Figure 2 and Table 3), indicated that cluster 4 is not a separate population from cluster 3. Cluster 4 is formed by sequences HM143017 to HM143037 that were shorter (≤304 pb), whereas sequences presented in the other clusters have ≥617 pb. In general, the clusters with high FST and low Nm exhibited association with a particular geographic distribution (Figure 3).

The Bayesian phylogenetic tree (Supplementary figure 1) that was in accordance with previous publications, exhibited six different clades with high values of posterior probability in principal branches; clade 1 showed *S. luriei* (AM747302); the second clade is *S. globosa*; the third and fourth clades are *S. brasiliensis* and *S. schenckii s. str.* isolates, the fifth and sixth seen at the end of the tree:
Table 3. Genetic differentiation among *S. schenckii* s. str. sub-populations in the infinite island model.

<table>
<thead>
<tr>
<th>Population 1</th>
<th>Population 2</th>
<th>Hs</th>
<th>G&lt;sub&gt;ST&lt;/sub&gt;</th>
<th>F&lt;sub&gt;ST&lt;/sub&gt;*</th>
<th>Nm&lt;sup&gt;¤&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>Cluster 2</td>
<td>0.172</td>
<td>0.269</td>
<td>0.805</td>
<td>0.060</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>Cluster 3</td>
<td>0.477</td>
<td>0.355</td>
<td>0.917</td>
<td>0.022</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>Cluster 4</td>
<td>0.324</td>
<td>0.377</td>
<td>0.778</td>
<td>0.071</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>Cluster 5</td>
<td>0.166</td>
<td>0.609</td>
<td>0.961</td>
<td>0.010</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>Cluster 3</td>
<td>0.630</td>
<td>0.071</td>
<td>0.556</td>
<td>0.199</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>Cluster 4</td>
<td>0.877</td>
<td>0.051</td>
<td>0.362</td>
<td>0.440</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>Cluster 5</td>
<td>0.419</td>
<td>0.227</td>
<td>0.784</td>
<td>0.068</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>Cluster 4</td>
<td>0.654</td>
<td>0.018</td>
<td>0.050</td>
<td>4.750</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>Cluster 5</td>
<td>0.593</td>
<td>0.157</td>
<td>0.902</td>
<td>0.027</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>Cluster 5</td>
<td>0.637</td>
<td>0.244</td>
<td>0.734</td>
<td>0.091</td>
</tr>
</tbody>
</table>

Note: Interpopulation heterozygosity (Hs), ancestry coefficient (F<sub>ST</sub>), gene diversity (G<sub>ST</sub>), migration index (Nm).

*The common used values for F<sub>ST</sub> are as follows: 0 to 0.05, small; 0.05 to 0.15, moderate; 0.15 to 0.25, great; values above 0.25 indicate huge genetic differentiation.*

<sup>¤</sup> Those strongly differentiated have an Nm < 1, whereas an Nm > 4 behaves as a single panmictic unit.

*S. pallida* and *S. mexicana* isolates. In this tree *S. mexicana* seems to share a common ancestor with *S. pallida*.

All these clades include isolates identified in GenBank as *S. schenckii*, but most of them were published before new species were proposed using calmodulin partial sequences (*S. brasiliensis*, *S. mexicana* and *S. globosa*). In order to confirm this statement, we analyzed carefully the partial calmodulin sequence alignment, and we found conserved and polymorphic regions, some of which are species-specific and could be used as reference to identify the members of *S. schenckii* s. l. (Supplementary figure 2A and 2B). Having this issue in mind, the distribution observed in the phylogenetic and network trees is in agreement with the present classification. Nevertheless, distinction between *S. brasiliensis* and *S. schenckii* s. str. using only the calmodulin sequence could be difficult if they are not sufficiently long.

### Discussion

The genus *Sporothrix* involves a taxon with many species; *S. schenckii* s. str. exhibits high genetic diversity, with four well-defined subpopulations, indicating that they could be cryptic species, while phylogenetic and population genetic analyses point to a process of purifying selection in the recent evolutionary past of *S. brasiliensis* and *S. globosa*, probably followed by clonal expansion. These data are supported by previous findings of electrophoretic karyotype and DNA profiles of *S. brasiliensis* and *S. globosa* isolates showed less variability than those observed in *S. schenckii* s. str. isolates. Ferreira et al. suggested that genetic polymorphism could correlate with diverse degrees of virulence; in addition, the interaction of host-pathogen and environment probably have had a structuring effect on population diversity, frequency, and distribution of *S. schenckii* s. l. Epidemiologic information support this idea and suggests that organic products are associated with *S. globosa* infection in humans; as a consequence exchange in international food markets could be associated with worldwide distribution and rapid clonal expansion. In contrast, domestic cat represents a slow vector of dispersion and limited dispersion of *S. brasiliensis*. Furthermore, population genetics indices might support a purified selection process for *S. globosa* and *S. brasiliensis* due to the low Hs < 0.22, in relation to more polymorphic and ancestral *S. schenckii* s. str. that display high Hs.

![Figure 3. Geographic distribution of five clusters of *S. schenckii* s. str. according of the calmodulin gene. Cluster 1 (green) was found in Asian and Latin America, predominates above tropic of Cancer; Cluster 2 (pink) was found in North American and France; Cluster 3 (blue) was seen mainly between the Equator and Tropic of Capricornio; Cluster 4 (red) only presented in Brazil and Cluster 5 (black) was found below the Equator in Argentina, Ecuador and South Africa.](https://academic.oup.com/mmy/article-abstract/54/3/248/2579141)
values (0.51 to 0.72), as has been argued for other fungi. Phylagenetic and haplotype network trees also suggest recent S. brasiliensis clonal development, since S. schenckii and S. brasiliensis show the lowest ancestry coefficient (FST) between the analysed populations. This information helps understanding the difficulties found in using morphology to distinguish between S. brasiliensis and S. schenckii. On the other hand, population analyses exhibits huge differences between S. brasiliensis and S. mexicana, as seen by the lowest Nm and the highest FST. In conclusion, our data suggest that S. globosa and S. brasiliensis experienced differentiation from a common ancestor with S. schenckii s. str., while S. mexicana shares a common ancestor with S. pallida. Besides, our approach regarding the population genetics analyses support the present classification that includes four pathogenic species, phylogenetically related, namely S. brasiliensis (clade I), S. schenckii (clade II) S. globosa (clade III), and S. luriei (clade VI). The usually saprophytic species S. mexicana (clade IV) and S. pallida (clade V, formerly S. albicans) are phylogenetically distant and therefore are not part of the S. schenckii complex, however in susceptible host they could be pathogenic. These six species presented diverse levels of genetic polymorphism and limited gene flow. Interestingly, our data also suggest the presence of other cryptic clusters; this idea is supported by differences in their virulence profile, protein secretion, and immunogenicity among S. schenckii s. str. In addition, it has been argued that the genetic variation within and between populations of pathogens determines the future evolution, genetic differentiation, and speciation; for example, the epidemiology of the recent S. brasiliensis outbreak is hallmarked by its association with cat infection and transmission, initial clonal spread, and limited geographic distribution in southeast Brazil. In general, the problem of fungal infections in animals is observed as a result of a recent introduction of a pathogen in a susceptible host population; in this case, clonal expansion of S. brasiliensis supported the idea of recent contact of S. brasiliensis with cats. Curiously, clonal expansion of S. schenckii s. str. was reported in Malaysia cats infections recently, which belong to cluster 1 in our analysis. On the other side during the coevolution process of host-pathogens, high diversity of the pathogen and better adaptation to infection of the host is expected. We considered that probably cats have been only recently in strong contact with this fungus due to the particular ecological proximity between wild areas and poor “favelas” in city limits, where semidomestic cats acquire the disease from environment (soil, woods or wild animals, such as rodents, armadillos and hematogenous bats, Desmodus rotundus) and transmit the disease to other cats or humans. S. globosa expansion, in contrast, could be associated with the development of agriculture, implementation of new farming techniques and long-distance travel of infected hosts. In the Australian epidemics, the evidence indicates that hay was the main transmission source; similarly, it could be a plant rather than an animal what caused expansion of S. globosa; Nevertheless, recently it was reported that an S. globosa isolate from feline host in Chiba, Japan. Therefore, it is necessary to study the natural hosts for S. globosa and the patterns of transmission and dispersion, as have been described in Australia outbreak.

Acknowledgments

Lucia Rangel-Gamboa performed her PhD at the Doctorado en Ciencias Biomedicas, Universidad Nacional Autonoma de Mexico (Biomedical Sciences Doctorate Program at the National University of Mexico).

Author contributions

LRG, AF, RA, and FMH conceived and designed the population genetic analyses; LRG and FMH performed the population genetic analyses. LRG, AF, RA, and FMH conceived the population genetic model; LRG and FMH performed the population genetic analyses; LRG, FMH, and PM analysed the data. LRG, AF, FMH, PM, and RA wrote the paper.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and the writing of the paper.

Supplementary material

Supplementary material is available at Medical Mycology online (http://www.mmy.oxfordjournals.org/)

References


