Phylogenetic heterogeneity of the genus *Williopsis* as revealed by 18S rRNA gene sequences

Stephen A. James,¹ Ian N. Roberts¹ and Matthew D. Collins²

Author for correspondence: Stephen A. James. Tel: +44 1603 255274. Fax: +44 1603 458414. e-mail: Steve.James@BBSRC.AC.UK

A phylogenetic investigation of the ascomycetous yeast genus *Williopsis* was performed by using 18S rRNA gene sequence analysis. Comparative sequence analysis revealed the genus to be phylogenetically heterogeneous. The five varieties of *Williopsis saturnus* [var. *mrakii*, var. *sargentensis*, var. *saturnus* (type), var. *suaveolens* and var. *subsufficiens*] were found to have identical 18S rRNA gene sequences and formed a distinct group, quite separate from all other *Williopsis* and non-*Williopsis* species examined. *Williopsis mucosa* was found to be the closest phylogenetic relative to the *Williopsis saturnus* group, however a sequence divergence of approximately 2.3% suggests this species may belong to a separate genus. The recently described species *Williopsis salicorniae* was found to exhibit a relatively close association with *Ogataea minutata* (≡ *Pichia minutata*), the type species of the genus *Ogataea*. The remaining two members of the genus, *Williopsis californica* and *Williopsis pratensis*, were found to form distinct lineages, displaying no specific association with any other *Williopsis* or non-*Williopsis* species. Based on comparative analysis of 18S rRNA genes it is apparent that the genus *Williopsis* as presently constituted is not monophyletic, and that the five currently recognized species form separate sublines each potentially worthy of separate generic status. The genus *Williopsis* should be restricted to the type species *Williopsis saturnus* and its five varieties. Despite the five varieties of *Williopsis saturnus* being genealogically indistinguishable at the 18S rRNA gene level, sequence analysis of the internal transcribed spacer (ITS) region revealed that the five varieties could be differentiated on both their ITS1 and their ITS2 sequences, providing further evidence of the value of ITS sequences for discrimination of yeasts at the subspecies level.

**Keywords:** genus *Williopsis*, 18S rRNA gene sequence, ITS sequences

INTRODUCTION

The genus *Williopsis* was originally introduced to accommodate the saturn-shaped ascospore-forming, nitrate-assimilating species *Williopsis saturnus*. Since the genus was first defined in 1925 by Zender, further species have been accommodated within this genus. In 1979, *Williopsis pratensis* was described as a new *Williopsis* species by Babjeva & Reshetova (1). Barnett et al. (2) also transferred several nitrate assimilating, saturn-shaped ascospore-forming species of the genus *Hansenula*, including *Williopsis californica* to the genus *Williopsis*. In a recent nDNA–nDNA reassociation study, Kurtzman (10) described several new varieties of *W. saturnus* (*W. saturnus* var. *mrakii*, var. *sargentensis*, var. *suaveolens* and var. *subsufficiens*) on the basis of exhibiting intermediate DNA homology values, as well as transferring *Pichia mucosa* (which does not assimilate nitrate) to the genus, as the new combination *Williopsis mucosa*. In 1991, Hinzelin et al. (7) described a fifth species, *Williopsis salicorniae*, isolated from brackish water. This species like *W. mucosa* does not assimilate nitrate. Thus at present the
genus *Williopsis* consists of five species (*W. californica, W. mucosa, W. pratensis, W. salicorniae and W. saturnus* (type species)), one of which (*W. saturnus*) is comprised of five varieties.

Recent partial 18S rRNA (approx. 160 bases) and 26S rRNA (approx. 350 bases) sequencing analyses by Yamada *et al.* (16, 17) revealed the genus to be heterogeneous, with *W. mucosa* and *Pichia anomalata* displaying identical partial 18S rRNA sequences, as did *W. salicorniae* and *Ogataea glucozyma* (= *Pichia glucozyma* (15)). Liu & Kurtzman (11) also commented upon the high rRNA sequence variation observed among *Williopsis* species, compared to that displayed by species within other ascomycetous genera. It is recognized that using different regions of rRNA and/or gene sequences can result in differing tree topologies (8). To improve the reliability of phylogenetic inferences it is essential wherever possible to use complete or near complete gene sequence data and by species within other ascomycetous genera. It is also recognized that using different regions of rRNA and/or gene sequences can result in differing tree topologies (8). To improve the reliability of phylogenetic inferences it is essential wherever possible to use complete or near complete gene sequence data and to standardize the approach to facilitate subsequent comparative analysis. Hence this study aimed to extend the phylogenetic investigations of Yamada *et al.* (16, 17), by using almost complete 18S rRNA gene sequences and a more comprehensive selection of ascomycetous species. The use of internal transcribed spacer (ITS) sequences for the delineation of the five varieties of *W. saturnus* was also investigated.

**METHODS**

**Yeast strains and cultivation.** The yeast strains examined in this study are listed in Table 1. Strains were obtained from the National Collection of Yeast Cultures, Norwich, UK, and the Centraalbureau voor Schimmelcultures, Delft, The Netherlands. All strains were grown on YM agar (0.3% yeast extract, 0.3% malt extract, 0.5% peptone, 1% glucose, 20% agar; pH 5.5) at 24 °C.

**Table 1. Williopsis strains studied and their 18S rRNA gene, ITS1 and ITS2 sequence accession numbers**

<table>
<thead>
<tr>
<th>Species†</th>
<th>Strain†</th>
<th>EMBL accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18S rRNA</td>
<td>ITS1</td>
</tr>
<tr>
<td><em>Williopsis californica</em></td>
<td>NCYC 2590†</td>
<td>Y12108</td>
</tr>
<tr>
<td><em>Williopsis mucosa</em></td>
<td>CBS 6341†</td>
<td>Y12109</td>
</tr>
<tr>
<td><em>Williopsis pratensis</em></td>
<td>CBS 7079†</td>
<td>Y12110</td>
</tr>
<tr>
<td><em>Williopsis salicorniae</em></td>
<td>CBS 8071†</td>
<td>Y12511</td>
</tr>
<tr>
<td><em>Williopsis saturnus var. mrakii</em></td>
<td>NCYC 500†</td>
<td>Y11318</td>
</tr>
<tr>
<td><em>Williopsis saturnus var. sargentensis</em></td>
<td>NCYC 2251</td>
<td>Y11318</td>
</tr>
<tr>
<td><em>Williopsis saturnus var. sargentensis</em></td>
<td>CBS 6342†</td>
<td>Y12112</td>
</tr>
<tr>
<td><em>Williopsis saturnus var. sargentensis</em></td>
<td>CBS 5761†</td>
<td>Y12111</td>
</tr>
<tr>
<td><em>Williopsis saturnus var. sargentensis</em></td>
<td>NCYC 23</td>
<td>Y12111</td>
</tr>
<tr>
<td><em>Williopsis saturnus var. suaveolens</em></td>
<td>NCYC 2586†</td>
<td>Y12103</td>
</tr>
<tr>
<td><em>Williopsis saturnus var. subsufficiens</em></td>
<td>CBS 5763†</td>
<td>Y12102</td>
</tr>
</tbody>
</table>

* The five species are described by Kurtzman (1991) and Liu & Kurtzman (1991).
† Abbreviations: CBS, Centraalbureau voor Schimmelcultures, Delft, The Netherlands; NCYC, National Collection of Yeast Cultures, Norwich, UK.

**PCR amplification of the 18S rDNA and ITS region.** Amplification of the 18S rRNA gene was performed as described by James *et al.* (8). The entire ITS region was amplified as previously described (9), with the exception that primer P3490 ([5'-CCGACGCCGCTACACTGA]; positions 1454–1473, *Saccharomyces cerevisiae* numbering (12)) was used in place of primer pITS1. The amplified products were purified using a QIAGEN QIAquick PCR purification kit according to the manufacturer's instructions.

**Sequence determination and analysis.** Direct sequencing of both the 18S rRNA gene and ITS PCR products was performed using a *Taq* DyDeoxy terminator cycle sequencing kit (Applied Biosystems) and a Hybaid Omnigene thermal cycler according to the manufacturers' recommendations. Complete 18S rDNA and ITS sequences were determined by using the primers described previously (8, 9). Two additional primers; *WIL1* (5'-ATTTCTGCCCCTATC-CAACT; positions 301–318, *S. cerevisiae* numbering) and *WIL2* (5'-AGTTGATAGGGCGAGAAAT; positions 318–301, *S. cerevisiae* numbering) were used for the 18S rRNA gene sequencing of *Williopsis* strains. Purified sequence reaction mixtures were electrophoresed with an Applied Biosystems model 373A automatic DNA sequencer.

**Analysis of sequence data.** The 18S rRNA gene sequences were aligned using the multiple-sequence alignment program *PILEUP* (5) contained within the Genetics Computer Group software package (6), version 8.1. The alignments were adjusted manually. 18S rRNA gene sequence similarity values were calculated using the program *GAP*. Phylogenetic analyses were performed by using the *PHYLIP* phylogeny inference package (4), version 3.572. A distance matrix was obtained by using the *DNADIST* program, and an unrooted phylogenetic tree was constructed by using the neighbour-joining method (14) and the *NEIGHBOR* program. The stability of the individual branches was assessed by using the

International Journal of Systematic Bacteriology 48
RESULTS AND DISCUSSION

18S rRNA gene sequence analysis

The 18S rRNA gene sequences of the type species of all currently recognized Williopsis species (Table 1) were amplified in vitro by PCR, and their nucleotide sequences were determined directly. These newly determined sequences consisted of approximately 1750 nucleotides each and represented more than 95% of the 18S rRNA primary structure. These sequences were aligned with complete or near-complete 18S rRNA sequences retrieved from the GenBank and EMBL databases for species belonging to other ascomycetous genera. Levels of sequence similarity were calculated and the derived distances were used to infer phylogenetic relationships. Fig. 1 shows an unrooted tree constructed by using the neighbour-joining method (14) and depicts the phylogenetic relationship between the species of the genus Williopsis and some other ascomycetous yeasts. The stability of individual branches of the tree was determined by bootstrap analysis (3).

Comparative analysis of the 18S rRNA gene sequence data revealed that the nine Williopsis type strains examined in this study exhibited a substantial degree of sequence variation, with sequence similarity values ranging from 89.7 to 100%. A similar observation was made by Liu & Kurtzman (11), who examined partial 18S and 26S rRNA sequences. In a recent nDNA–nDNA reassociation study, Kurtzman (10) found that the variety of W. saturnus displayed moderate (36–72%) DNA relatedness. However, in this study all five varieties of W. saturnus were found to be genealogically indistinguishable from one another, having identical 18S rRNA gene sequences. It was evident from the results of the phylogenetic analysis (Fig. 1) that the genus Williopsis is not monophyletic. The type species (including its five varieties) formed a distinct phylogenetic lineage which was quite separate from all other Williopsis and non-Williopsis species examined. The closest relative of W. saturnus is W. mucosa. Though both species form saturn-shaped ascospores, W. mucosa can readily be distinguished from W. saturnus by its inability to assimilate potassium nitrate as a sole source of nitrogen (2). A 18S rRNA sequence divergence of 2.3% probably precludes W. mucosa from belonging to the same genus as W. saturnus. It is pertinent to note that Yamada et al. (16) found that W. mucosa and P. anomala had identical partial 18S rRNA sequences based on a comparison of 160 bp (positions 1451–1618; S. cerevisiae numbering). However, in this study employing near-complete 18S rRNA sequences, W. mucosa and P. anomala formed quite separate phylogenetic lineages (Fig. 1), displaying approximately 3% sequence divergence. This distance, coupled with the fact that W. mucosa and P. anomala produce amongst other traits quite different shaped ascospores (W. mucosa, saturn-shaped; P. anomala, hat-shaped), excludes them from belonging to the same genus. This discrepancy between the results of Yamada et al. (16) and the present study further demonstrates the dangers of sampling errors using partial rRNA sequences (in this case approx. 160 bp) and misleading phylogenetic inferences. In 1986, the new genus Waltiozyma was proposed by Müller & Kock (13) to accommodate W. mucosa as the new combination, Waltiozyma mucosa, based on the apparent uniqueness of its fatty acid composition. However, in his recent nDNA–nDNA reassociation study, Kurtzman (10) retained W. mucosa within the genus Williopsis. Nevertheless, the phylogenetic findings of this study support the transfer of W. mucosa to the genus Waltiozyma (13).

Williopsis salicorniae and Ogataea minuta [≡ P. minuta] and now the type species of the genus Ogataea, Yamada et al. (15) were found to form a distinct group supported by a bootstrap value of 100%. The 18S rRNA sequence similarity of this pair of species was 98.6%. In a recent study using partial 18S and 26S rRNA sequences, Yamada et al. (17) found that W. salicorniae and Ogataea glucozyma (≡ P. glucozyma; Yamada et al. (15)) had identical partial 18S rRNA sequences (positions 1451–1618; S. cerevisiae numbering). However, they also noted that W. salicorniae and O. minuta were morphologically and physiologically quite different. Thus they concluded (17) that on phenotypic criteria, W. salicorniae could not be classified in the genus Ogataea. In the present study, W. salicorniae and O. minuta were found to exhibit a significant phylogenetic association with one another. However an 18S rRNA sequence divergence of approximately 1.4% in conjunction with their distinct phenotypes suggests a relationship of two closely related, but nevertheless separate, genera.

The remaining two species of the genus Williopsis, W. californica and W. pratensis, were found to form two distinct phylogenetic lineages which displayed no close association with any other Williopsis or non-Williopsis species. Similar results were obtained by both Liu & Kurtzman (11) and Yamada et al. (16), using partial 18S and 26S rRNA sequences. W. pratensis can be distinguished from other Williopsis species, including W. californica on the basis that it produces 1–2 saturn-shaped ascospores as opposed to 1–4 saturn-shaped ascospores, and its ascii are persistent, in contrast to those from other Williopsis species which are evanescent. Physiologically, W. pratensis can be distinguished from other members of the genus Williopsis by its ability to ferment maltose and its inability to assimilate glycerol. For these reasons, Yamada et al. (16) proposed that this species be transferred to a new genus, Komagataea, as the new combination Komagataea pratensis. In the same study, they also proposed that the genus Zygowilliopsis be reinstated to accommodate the species W. californica, as the com-
Fig. 1. Dendrogram showing the phylogenetic relationship of species of the genus *Williopsis* and some other ascomycetous yeasts based on 18S rRNA gene sequences. The tree was constructed by the neighbour-joining method. Bootstrap values, expressed as percentages of 200 replications, are given at branch points (only values ≥ 50% are shown). Scale bar, 6 estimated base substitutions per 1000 nucleotide positions.
bination Zygowilliopsis californica. The findings of this phylogenetic study support the transfer of both these yeast species to genera quite separate from the genus Williopsis.

**ITS sequence analysis**

The nucleotide sequences of ITS1 and ITS2 regions of the five Williopsis species including all five varieties of *W. saturnus* (Table 1) were determined by directly sequencing PCR-amplified fragments. Length polymorphisms in both ITS1 and ITS2 were observed between all species. For ITS1, *W. californica*, *W. mucosa*, *W. pratensis*, *W. salicorniae* and *W. saturnus* were similar in their ITS1 sequences (Table 1) were determined by directly sequencing PCR-amplified fragments. Length polymorphisms in both ITS1 and ITS2 were observed between all species. For ITS1, *W. californica*, *W. mucosa*, *W. pratensis*, *W. salicorniae* and *W. saturnus* were similar in their ITS1 sequences. However, indels and gross sequence differences precluded reliable spacer sequence alignment essential for phylogenetic inferences. This was particularly apparent for more distantly related taxa (data not shown).

In a recent study (9), we demonstrated that ITS sequences could be used to differentiate between genealogically closely related species. We showed that Zygosaccharomyces bailii and Zygosaccharomyces bisporus, which have almost identical 18S rRNA gene sequences (99.8% sequence similarity), could readily be distinguished on their ITS sequences. Indeed, in the case of Zygosaccharomyces bailii, ITS1 and ITS2 sequences were found to provide resolution at the subspecies level. In the present investigation the five varieties of *W. saturnus* were also found to be indistinguishable at the 18S rRNA gene sequence level, despite the report of considerable divergence in nDNA-nDNA reassociation experiments (10). An examination of the ITS1 and ITS2 sequences for all five *W. saturnus* varieties revealed these to be relatively highly conserved (Figs 2, 3). However, despite the high overall sequence conservation, all five varieties of *W. saturnus* could be distinguished on either their ITS1 or ITS2 sequences. The ITS1 spacer was found to possess a diagnostic region in positions 73–82 (Fig. 2) (numbering based on the ITS2 sequence of *W. saturnus* var. *sartorius*), in which four of the five varieties displayed distinct signature sequences (Figs 2, 3) (numbering based on the ITS2 sequence of *W. saturnus* var. *sartorius*), in which four of the five varieties displayed distinct signature sequences.

**Fig. 2.** Alignment of the ITS1 sequences of the five varieties of *W. saturnus*. Wsat, *W. saturnus* var. *sartorius* (CBS 5761); Wmrak, *W. saturnus* var. *mrakii* (NCYC 500); Wsuav, *W. suaveolens* (NCYC 2586); Wsub, *W. subflavescens* (CBS 6342); Wsarg, *W. sargentensis* (CBS 5763). The diagnostic region is underlined.

**Fig. 3.** Alignment of the ITS2 sequences of the five varieties of *W. saturnus*. Abbreviations as in Fig. 2. The diagnostic region is underlined.
var. mrakii, W. saturnus var. saturnus and W. saturnus var. suaveolens were examined (see Table 1).

The results of our phylogenetic analysis, using near-complete 18S rRNA gene sequences, clearly demonstrate that the genus Williopsis is not monophyletic. In particular, our findings support the conclusions of Yamada et al. (16) that the genus Williopsis should be restricted to the type species W. saturnus and its five varieties. Although the other species (W. californica, W. mucosa, W. pratensis and W. salicorniae) are characterized by their ability to form saturn-shaped ascospores, it is evident from our results that they all form individual sublines, distinct from each other and other yeast taxa examined. Phylogenetically these four species may constitute the nuclei of new genera. In the case of W. saturnus, we have also demonstrated that ITS sequences are not only extremely useful for the delineation of phylogenetically closely related species (9), but also have a use in the differentiation of individual varieties of a species. The presence of indels and the low overall degree of sequence conservation of ITSs, however, precluded their use for assessing phylogenetically distant taxa.

ACKNOWLEDGEMENTS

We would like to thank Dr David Yarrow from the Centraalbureau voor Schimmelcultures at Delft, for kindly providing yeast strains for analysis in this study. The authors are indebted to the European Community (contract CT93-0830) for supporting this work.

REFERENCES