

## Molecular Comparison of Mating Type Loci and Adjacent Chromosomal Regions from Self-fertile and Self-sterile *Cochliobolus* Species

Sung-Hwan Yun and B. Gillian Turgeon\*

Department of Plant Pathology, Cornell University, 334 Plant Science Building, Ithaca, NY 14853, USA

(Received on June 19, 1999)

**In fungi known as ascomycetes, ability to mate is controlled by a single mating type (*MAT*) locus with two dissimilar sequences called idiomorphs carrying genes encoding transcription factors that are unrelated to each other. Fungi requiring strains with different *MAT* genes to complete the sexual process are heterothallic (self-sterile); species in which a single strain is able to undergo sexual reproduction are homothallic (self-fertile). Previous analysis of sequences from several heterothallic and homothallic species of the ascomycete genus *Cochliobolus* showed that homothallics evolve from heterothallics and that each known *Cochliobolus* homothallic species arose independently, from a different heterothallic ancestral species. Here we report detailed comparative analyses of *MAT* sequences and their flanking regions, and show that: (1) The level of *MAT* gene similarity is not correlated with reproductive life style; (2) *MAT* proteins from all *Cochliobolus* species are conserved within the transcription factor signature sequences; they are not conserved in the carboxy terminal half of *MAT-1*, or third of *MAT-2*, except in those from very closely related species; (3) A gene (*ORF1*) of unknown function, consistently found on the *MAT* flank, is more conserved than are the *MAT* genes themselves; (4) The intergenic sequences diverge sharply among species.**

Fungi capable of sexual reproduction can be self-fertile (homothallic) or self-sterile (heterothallic). Previously, to investigate the molecular basis of homothallism, the *MAT* regions from four homothallic species within the ascomycete genus *Cochliobolus* were cloned, sequenced and compared to each other and to *MAT* regions of heterothallic species (Yun et al., 1999). Heterothallic ascomycetes have a single mating type locus called *MAT*; alternate sequences at *MAT* are not "alleles" in the classic sense because they lack significant sequence similarity to each other and each gene carried by an idiomorph encodes a different transcriptional regulator (Turgeon et al., 1993). The dissimilar sequences at *MAT* are termed 'idiomorphs' (Metzenberg and Glass,

1990). This organization is common among all heterothallics analyzed to date, including the *Cochliobolus* spp. (Coppin, 1997).

We found that each homothallic *Cochliobolus* species examined carries both *MAT* genes but, in contrast to heterothallics, the structural organization is unique in each case (Yun et al., 1999). For two species (*C. luttrellii* and *C. homomorphus*) the genes are fused into a single open reading frame; the order of the genes is reversed in *C. homomorphus* with respect to *C. luttrellii*. In the remaining two cases, the genes are not fused. In *C. kusanoi*, the core organization is 5' *MAT-2* 3' - 3' *MAT-1* 5'; the sequence between the genes is similar to a portion of the  $\beta$ -glucosidase gene normally found on the 3' flank of both *MAT* genes in heterothallic *C. heterostrophus* (accession number AF029913 and AF027687). Flanking the core is a perfect inverted repeat of a 561 bp region containing 123 bp of the 5' end of the *MAT-1* ORF and 438 bp of additional sequence including 145 bp of yet a different fragment of the  $\beta$ -glucosidase gene. In the fourth homothallic species, *C. cymbopogonis*, complete copies of the *MAT-1* and *MAT-2* ORFs are present; however, PCR analysis using various combinations of *MAT*-specific primers, along with gel blot analysis, provided no evidence for linkage of the *MAT* genes within 30 kb of each other. Thus the *MAT* genes are closely linked in three homothallics but not in the fourth.

Linked to the *MAT* genes in three of four homothallic species and in all heterothallics is a highly conserved ORF (*ORF1*; Wirsal et al., 1997) that shows similarity to a *Saccharomyces cerevisiae* ORF (accession number U22383) of unknown function. Despite the consistent association with *MAT*, deletion of *ORF1* in *C. heterostrophus* has no effect on mating, or on any other detectable phenotype. In *C. cymbopogonis*, there are two copies of *ORF1* and each is linked to a different *MAT* gene. *C. kusanoi* is an exception; gel blots and PCR amplifications indicate that *ORF1* exists in the genome but suggest it is not closely linked to either *MAT-1* or *MAT-2*.

Here we report a detailed analysis of nucleotide and amino acid sequences of *ORF1*, the *MAT* genes themselves, and the 5' and 3' flanking DNA sequences. Results show that both *MAT* and *ORF1* are highly conserved among

\*Corresponding author.

Phone) +1-607-254-7458, Fax) +1-607-255-4471

E-mail: bgt1@cornell.edu

closely related species, less conserved among distantly related species and that similarity between genes is associated with their phylogenetic relatedness, determined using molecular characters other than *MAT* (Berbee et al., 1999) rather than with their mode of reproduction.

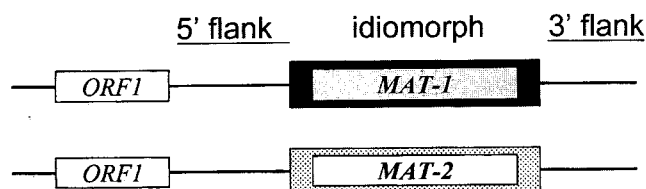
## Materials and Methods

**Strains and sequences.** Comparisons were made among sequences at the *MAT* locus from the following fungi: *C. heterostrophus* heterothallic *MAT-2* strain C4 (Acc. # AF027687) and *MAT-1* strain C5 (Acc. # AF029913), heterothallic *C. ellisii* *MAT-1* strain 81154-2 (Acc. # AF129746) and *MAT-2* strain 81154-7 (Acc. # AF129747), homothallic strains *C. luttrellii* 14643-1 (Acc. # AF129740), *C. cymbopogonis* 88109-1 (Acc. # AF129744 and Acc. # AF129745), *C. kusanoi* Ck2 (Acc. # AF129742) and *C. homomorphus* 13409 (Acc. # AF129741). Isolation and cloning of these sequences has been described (Yun et al. 1999).

**DNA manipulations.** DNA sequences were determined at the Cornell DNA Sequencing Facility and were assembled using the DNASTAR program SeqMan. Similarity searches of the non-redundant database maintained by the National Center for Biotechnology Information (NCBI) were performed using the BLAST suite of programs (Altschul et al., 1990). Comparative analysis of sequences was performed using the DNASTAR program MEGALIGN.

## Results and Discussion

**Size of *MAT* genes and *ORF1*.** Figure 1 shows a generalized diagram of the organization of *Cochliobolus* heterothallic *MAT* loci. For details see Yun et al. (1999). Tables 1 and 2 show the number of nucleotides compared in this study for three genes (*MAT-1*, *MAT-2* and *ORF1*) and adjacent flanking DNA regions. The size of the idiomorphs



**Fig. 1.** Diagram of the *Cochliobolus* heterothallic *MAT* locus and adjacent regions.

**Table 1.** Size of *MAT* ORFs and *ORF1*

Species	<i>MAT-1</i> <sup>a</sup> nt/aa <sup>c</sup>	<i>MAT-2</i> <sup>a</sup> nt/aa <sup>c</sup>	<i>ORF1</i> <sup>b</sup> nt/aa
<i>C. heterostrophus</i> <sup>d</sup>	1,204/384	1,087/344	669/223
<i>C. ellisii</i>	1,190/378	1,087/343	555/184 (at <i>MAT-1</i> ) 531/177 (at <i>MAT-2</i> )
<i>C. luttrellii</i> <sup>e</sup>	856/268	940/296	573/190
<i>C. homomorphus</i> <sup>e</sup>	1,166/371	1,063/336	127/41
<i>C. kusanoi</i>	1,094/335	1,093/346	420/140
<i>C. cymbopogonis</i>	1,118/355	1,035/327	521/172 (at <i>MAT-1</i> ) 599/198 (at <i>MAT-2</i> )

<sup>a</sup> complete gene sequences, nt = nucleotide (bp); aa = amino acid.

<sup>b</sup> partial sequences except for complete sequence for *C. heterostrophus*.

<sup>c</sup> number of amino acids in the ORF encoding the *MAT* protein, within the idiomorph.

<sup>d</sup> *ORF1* sequences (Wirsal et al., 1996) from *C. heterostrophus* *MAT-1* and *MAT-2* flanks are identical.

<sup>e</sup> The *MAT* genes are fused and some amino acids at the fusion junctions are missing.

ranges between 1,094-1,204 nucleotides for *MAT-1* and 1063-1093 nucleotides for *MAT-2* in all cases except for *C. luttrellii* where they are shorter because *MAT-1* is fused to *MAT-2* and 345 nucleotides are missing from the fusion ORF compared to the heterothallic counterparts (Yun et al., 1999). Table 1 also shows sizes of proteins encoded by *MAT-1* (between 335 and 384 amino acids) and *MAT-2* (between 327 and 346 amino acids). In *C. luttrellii*, these proteins are unusually small (*MAT-1* = 268; *MAT-2* = 296) due to the truncated nature of the fusion protein. The *ORF1* sequence is complete only for *C. heterostrophus*.

In Table 2, the region between the 3' end of *ORF1* and the 5' end of the *MAT* gene (termed the 5' flank) is compared among species. The size of this region is remarkably constant (1,105-1,117 bp) for the two heterothallic species (*C. heterostrophus* and *C. ellisii*) and the homothallic species *C. cymbopogonis*, which has two copies of this region and no apparent linkage between the two *MAT* genes. The distance is also highly conserved when closely related pairs (Yun et al., 1999; Berbee et al., 1999) of heterothallic and homothallic species are compared, i.e., heterothallic *C. heterostrophus* and homothallic *C. luttrellii* (for *MAT-1*, 1,110 versus 1,097 bp; for *MAT-2*, 1,117 versus 1,097 bp) or heterothallic *C. ellisii* and homothallic *C. cymbopogonis* (for *MAT-1*, 1,115 versus 1,105 bp, for *MAT-2*, 1,115 versus

**Table 2.** Size (bp) of flanking sequences compared

	<i>C. heterostrophus</i>		<i>C. ellisii</i>		<i>C. luttrellii</i>	<i>C. homomorphus</i>	<i>C. kusanoi</i>		<i>C. cymbopogonis</i>	
	<i>MAT-1</i>	<i>MAT-2</i>	<i>MAT-1</i>	<i>MAT-2</i>	<i>MAT-1/2</i>	<i>MAT-2/1</i>	<i>MAT-1</i>	<i>MAT-2</i>	<i>MAT-1</i>	<i>MAT-2</i>
5' flank <sup>a</sup>	1,110	1,117	1,115	1,115	1,097	883	438	2,046	1,105	1,114
3' flank	1,047	1,262	2,103	1,682	903	482	1,091		837	1,126

<sup>a</sup>Sequence between the 3' end of *ORF1* and the 5' end of the *MAT* ORF except for *C. kusanoi*, where *MAT* and *ORF1* are not linked.

**Table 3.** Percent identities<sup>a</sup> of *MAT-1/MAT-2* genes from *Cochliobolus* spp.

	<i>C. het</i>	<i>C. ell</i>	<i>C. lut</i>	<i>C. hom</i>	<i>C. kus</i>	<i>C. cym</i>
<b>Nucleotide</b>						
<i>C. het</i>	58.9/65.1	90.1/92.1	70.6/77.7	68.3/63.6	56.1/63.6	
<i>C. ell</i>	56.2/64.8	64.2/64.3	35.9/65.4	63.0/68.1	69.1/68.7	
<i>C. lut</i>	92.2/92.2	63.9/62.0	74.7/78.4	70.2/64.4	60.8/56.9	
<i>C. hom</i>	70.9/79.3	54.9/66.0	78.6/79.6	69.7/70.5	46.7/47.7	
<i>C. kus</i>	68.6/73.5	57.1/64.6	73.9/72.5	71.8/74.4	57.2/59.7	
<i>C. cym</i>	49.0/55.9	65.6/66.0	66.1/52.8	48.8/58.4	48.0/58.0	
<b>Amino acid</b>						

Abbreviations: *C. het* = *C. heterostrophus*; *C. ell* = *C. ellisii*; *C. lut* = *C. luttrellii*; *C. hom* = *C. homomorphus*; *C. cym* = *C. cymbopogonis*

<sup>a</sup> % identities were calculated based on pairwise comparison of the *MAT* idiomorphs using Martinez/Needleman-Wunsch Alignment (for nucleotides) and Lipman-Pearson Method (for amino acids) in the DNASTAR program.

1,114 bp). The distance is somewhat less for *C. homomorphus* compared to the others. There is no meaningful comparison to be made for *C. kusanoi*, since *ORF1* appears unlinked to *MAT*. The 3' flank sequence shown in Table 2 represents sequence available at this time; there is no gene (such as *ORF1*) at the 3' end of the idiomorph that can be used as a marker.

#### Relatedness among *MAT-1* and among *MAT-2* genes.

Identities at both the nucleotide and amino acid levels among *MAT* genes are shown in Table 3 and Fig. 2. Previous molecular and phylogenetic analyses showed that species within the genus *Cochliobolus* fall into two large groups (Yun et al., 1999; Berbee et al., 1999; Turgeon, 1998). Heterothallic *C. heterostrophus* is a member of one group, while heterothallic *C. ellisii* represents the other. Homothallic *C. luttrellii*, closely related to *C. heterostrophus*, falls in the first group. Homothallic *C. homomorphus* is also in the first group but does not have a known close heterothallic relative. Homothallic *C. cymbopogonis* and *C. kusanoi* are in the second group, but neither is the closest relative of heterothallic *C. ellisii*, which is also in this group.

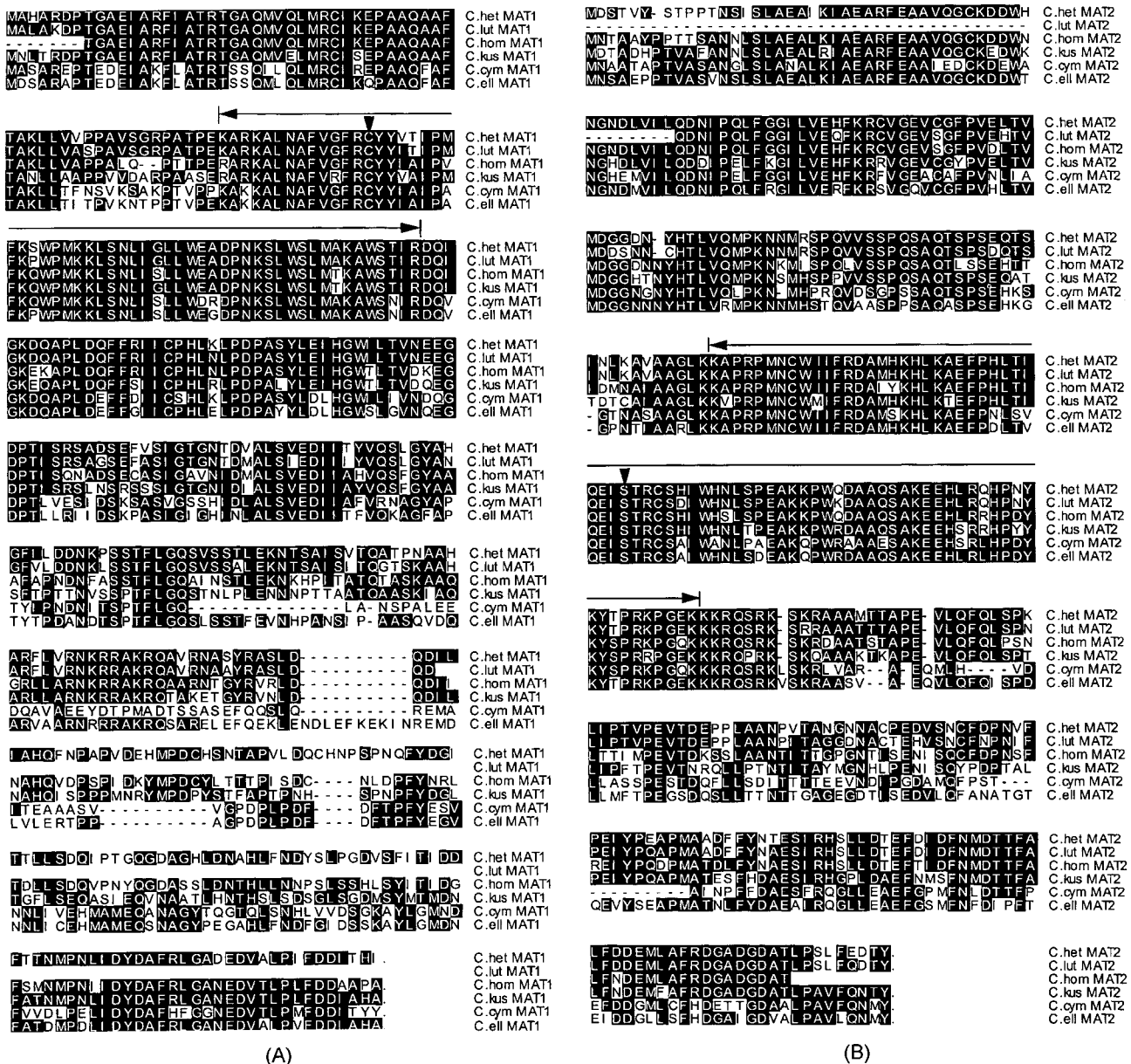
Pairwise percent identity of *MAT* sequences reflects the relative positions of these species in a phylogenetic tree, constructed using ribosomal DNA (ITS) and glyceraldehyde-3-phosphate dehydrogenase (*GPD*) sequences (Berbee et al., 1999; Yun et al., 1999). *MAT* genes from species within a group are more similar to each other than are *MAT* genes from species of similar mating style. For example, percent nucleotide identity between *MAT* genes in the two heterothallic *Cochliobolus* species, each representing one of the large *Cochliobolus* groupings, is only 58.9 for *MAT-1* and 65.1 for *MAT-2*. Average pairwise percent nucleotide identity for all homothallic *MAT-1* genes is 63.1; for *MAT-2*

it is 62.9, illustrating once again that similar reproductive life style does not mean highest nucleotide identity among *MAT* genes. In contrast, percent nucleotide identity between the *MAT* genes of heterothallic *C. heterostrophus* and homothallic *C. luttrellii* is 90.1 for *MAT-1* and 92.1 for *MAT-2*. Similarly, the species with *MAT* genes showing the highest similarity to those of heterothallic *C. ellisii* is homothallic *C. cymbopogonis*. Previous phylogenetic analyses strongly support the hypothesis that each homothallic *Cochliobolus* species arose independently from a different heterothallic ancestor (Yun et al., 1999).

The *MAT-1* and *MAT-2* signature sequences and intron positions within these domains are highly conserved among *MAT* proteins of all species, both homothallic and heterothallic (Fig. 2 A and B, arrowed over-lined regions and arrowheads, respectively). The amino terminal end of the *MAT-1* proteins is highly conserved among members within each large *Cochliobolus* group. The extreme carboxy terminal end is conserved among all *MAT-1* proteins. For *MAT-2*, the amino terminal end is conserved in all. Just 5' of the carboxy terminal end of both *MAT-1* and *MAT-2* is a region of little conservation. For example, *C. ellisii* *MAT-1* has an insert of 10 amino acids not found in the other *MAT-1* proteins, including that of *C. cymbopogonis*. Note that since the 3' end of the *C. luttrellii* *MAT-1* protein is missing 115 amino acids and the 5' end of the *C. luttrellii* *MAT-2* protein is missing 49 amino acids, due to the *MAT* gene fusion, these regions are missing in Figure 2.

**Relatedness among *ORF1* genes.** *ORF1* sequences are conserved among all species, ranging from 62.5% identity for two unrelated homothallic species *C. kusanoi* and *C. cymbopogonis* to 94.8% for the closely related species *C. heterostrophus* and *C. luttrellii* (Table 4). Interestingly, the two copies of *ORF1* in *C. cymbopogonis* are not identical (93.3% identity), suggesting that the duplication of this gene in the genome did not occur recently. The average pairwise percent identity among *ORF1* sequences in all homothallic species is 75.0. This value indicates that *ORF1* is more conserved than either *MAT* gene (*MAT-1* average = 63.1%; *MAT-2* = 62.9%). The high level of conservation among *ORF1* sequences has proven useful as a basis to design primers for PCR reactions aimed at cloning *ORF1*-linked *MAT* genes from fungi whose *MAT* genes themselves are too divergent for cloning by usual procedures (Yun et al., 1999).

**Relatedness among 5' and 3' flanking regions.** Similarity among noncoding DNAs flanking *MAT* varies with the relatedness of the species (Table 5). Within species percent identity in the 5' flank from the 3' end of *ORF1* to the 5' end of the *MAT* ORF is 96.6 for *C. heterostrophus* *MAT-1* and *MAT-2* and 97.0% for *C. ellisii* *MAT-1* and *MAT-2*. Interspecies percent identity in the same region for two closely



**Fig. 2.** Amino acid alignment of MAT proteins from homothallic and heterothallic *Cochliobolus* spp.: (A) MAT-1 and (B) MAT-2. The putative DNA binding domain in each protein (alpha box in MAT-1 and HMG box in MAT-2) is indicated by arrows; the position of a putative intron, which has been spliced out in each domain is shown by arrowheads. Black boxes = identities. Abbreviations: C. het = *C. heterostrophus*; C. lut = *C. luttrellii*; C. hom = *C. homomorphus*; C. kus = *C. kusanoi*; C. cym = *C. cymbopogonis*; C. ell = *C. ellisii*.

related species, heterothallic *C. heterostrophus* and homothallic *C. luttrellii* is 79.4 (MAT-1) - 81.2 (MAT-2). For *C. ellisii* and *C. cymbopogonis* it is 63.2-69.4%. Percent identity in the same region when any other pair of species is compared ranges from 10.5-51.1%. Within this region is a short stretch (about 190 bp, corresponding to bp positions 4,280-4,475 of the *C. heterostrophus* MAT-1 locus (Acc. # AF029913) that is highly conserved (80.0 to 85.7% identity) among all species (Table 5, footnote). Interestingly the 5' flanking regions of the two MAT genes of *C. cymbopogo-*

*nis* are only 77.6% identical. Note that the 100% identity for the 5' flank of *C. kusanoi* is a reflection of the presence of an exact repeat of the MAT-1 5' region, which is fused to MAT-2 (Yun et al., 1999).

DNA 3' of MAT (the 3 flank) is more divergent than that of the 5' flank; only *C. luttrellii* shares significant identity (78.0-82.0%) with the *C. heterostrophus* MAT-1 or MAT-2 flanks; 3 flanks of the other three homothallic species range from 13.1 to 49.0% identity with either of the two heterothallics. Furthermore, there is no conserved region (not even a

**Table 4.** Percent identities<sup>a</sup> among *ORF1* genes adjacent to *MAT* loci of *Cochliobolus* spp.

	<i>C. het</i>		<i>C. ell</i>		<i>C. lut</i>	<i>C. hom</i>	<i>C. kus</i>	<i>C. cym</i>	
	<i>MAT-1</i>	<i>MAT-2</i>	<i>MAT-1</i>	<i>MAT-2</i>				<i>MAT-1</i>	<i>MAT-2</i>
	<b>Nucleotide</b>								
<i>C. het</i>		70.9	69.5		94.8	78.2	66.9	73.5	74.5
<i>C. ell</i>									
<i>MAT-1</i>	85.8		96.4		76.3	80.3	68.8	84.9	86.7
<i>MAT-2</i>	84.1	98.9			69.3	80.9	67.9	86.0	87.5
<i>C. lut</i>	98.4	85.3	83.5			78.2	68.2	74.1	75.5
<i>C. hom</i>	86.5	83.8	81.1		86.5		89.3	81.6	80.9
<i>C. kus</i>	77.4	79.8	78.6		75.0	100.0 <sup>b</sup>		63.3	62.5
<i>C. cym</i>									
<i>MAT-1</i>	81.1	91.4	90.2		80.5	83.8	77.1		93.3
<i>MAT-2</i>	84.3	92.9	92.6		77.3	81.1	84.2	90.8	
	<b>Amino acid</b>								

<sup>a</sup>% identities were calculated based on pairwise comparison of *ORF1* (see footnote in Table 1).

<sup>b</sup>over 24 amino acids. This value reflects a direct repeat of this region.

**Table 5.** Percent nucleotide identity<sup>a</sup> in the 5' flank<sup>b</sup> and 3' flank<sup>c</sup> regions of *MAT* from *Cochliobolus* spp.

	<i>C. het</i>		<i>C. ell</i>		<i>C. lut</i>	<i>C. hom</i>	<i>C. kus</i>		<i>C. cym</i>	
	<i>MAT-1</i>	<i>MAT-2</i>	<i>MAT-1</i>	<i>MAT-2</i>	<i>MAT-1/2</i>	<i>MAT-2/1</i>	<i>MAT-1</i>	<i>MAT-2</i>	<i>MAT-1</i>	<i>MAT-2</i>
	<b>5' flanks</b>									
<i>C. het</i>										
<i>MAT-1</i>		96.6	50.5 <sup>d</sup>	50.7 <sup>d</sup>	81.2	51.1 <sup>d</sup>	44.0	18.4	32.4 <sup>d</sup>	41.0 <sup>d</sup>
<i>MAT-2</i>	95.9		45.7 <sup>d</sup>	45.5 <sup>d</sup>	79.4	51.1 <sup>d</sup>	34.3	13.5	32.6 <sup>d</sup>	39.0 <sup>d</sup>
<i>C. ell</i>										
<i>MAT-1</i>	27.9	17.9		97.0	23.9 <sup>d</sup>	36.4 <sup>d</sup>	47.4	15.1	67.8	69.4
<i>MAT-2</i>	11.8	13.8	96.1		42.9 <sup>d</sup>	44.4 <sup>d</sup>	49.8	11.3	67.9	63.2
<i>C. lut</i>										
<i>MAT-1/2</i>	78.0	82.0	28.1	33.4		49.6 <sup>d</sup>	-	29.4	33.2 <sup>d</sup>	38.1 <sup>d</sup>
<i>C. hom</i>										
<i>MAT-2/1</i>	49.0	47.9	26.4	18.7	26.9		39.7	18.1	35.3 <sup>d</sup>	31.1 <sup>d</sup>
<i>C. kus</i>										
<i>MAT-1</i>	27.9	43.4	25.5	20.8	29.4	42.3		100.0 <sup>e</sup>	10.5	31.8
<i>MAT-2</i>	27.9	43.4	25.5	20.8	29.4	42.3			14.8	33.0
<i>C. cym</i>										
<i>MAT-1</i>	39.0	39.0	29.9 <sup>f</sup>	30.3 <sup>f</sup>	33.1	42.2	13.1			77.6
<i>MAT-2</i>	42.3	42.1	41.0 <sup>f</sup>	45.9 <sup>f</sup>	22.3	24.0	36.7		62.9	
	<b>3' flanks</b>									

<sup>a</sup>% identity was calculated based on pairwise comparison of the sequences; the identity values from pairs of sequences showing significant similarity over the entire region (Table 2) are in bold.

<sup>b</sup>the noncoding DNA region between the 3' end of *ORF1* and the 5' end of *MAT* from each *Cochliobolus* spp. except for *C. kus*: size of this region from each fungus is shown in Table 2.

<sup>c</sup>the noncoding DNA region (except for *C. kus*) flanking the 3' end of the *MAT* idiomorph (Figure 1); see Table 2 for the sizes.

<sup>d</sup>Each pair has short stretches (about 190 bp) showing 80.0 to 85.7% identity, corresponding to the 5' flank (bp 4,280 to 4,475) of the *C. het* *MAT-1* locus (Accession # AF029913).

<sup>e</sup>438 bp of perfect inverted repeat.

<sup>f</sup>There are several stretches (ranging from 40 bp to 100 bp) showing to 70 to 95% identity to each other.

short stretch as on the 5' flanks) among the 3' flanks of all species, except between those of *C. cymbopogonis* and *C. ellisii* where several islands of similarity (70 to 95%) are found (Table 5, footnote). Note, again that the 3' flanking regions of the two *MAT* genes of *C. cymbopogonis* are non-

identical (only 62.9%).

Do differences in the molecular structure of the homothallic *MAT* loci among filamentous ascomycetes (such as presence of only one or both *MAT* genes in the haploid genome, or variations in regions flanking the *MAT*

idomorphs) reflect a variation in molecular mechanisms leading to homothallism? Nucleotide identities in the centromere proximal flank of the *MAT* idiomorphs from *Neurospora* spp. range from 20% (unrelated) to >90% (closely related) when comparisons are made among homothallic and heterothallic spp. (Randall and Metzenberg, 1995). Moreover, the observation that homothallics have a conserved unique region not found in heterothallics suggests a functional role for this region in determining homothallism (Randall and Metzenberg, 1995). We have not found a unique region in the corresponding flank (3' flank) of the homothallic *Cochliobolus* spp. In fact the organization of this region varies: 0.9 kb of *C. luttrellii* is 80.2% identical to the same region of closely related heterothallic *C. heterostrophus* (no variable region is found between these two fungi), 482 bp of *C. homomorphus* and 1.0 kb of *C. kusanoi* do not show any significant similarity to the heterothallics, and 0.8 kb and 1.1 kb of the 3' flanks from the *C. cybopogonis* *MAT-1* and *MAT-2* loci carry several islands of similarity (70 to 95%) to closely related heterothallic *C. ellisii* (Tables 2 and 3). In no case are the 3' flanks of the homothallics similar to each other. If a recombination event (Yun et al. 1999) is responsible for evolution of homothallic from heterothallic species, this variation in structure suggests a different event in every case.

The amount of similarity among *MAT* sequences including the conserved regions (idiomorph and *ORF1*) and highly variable flanking regions reflects phylogenetic distance of these species (compare Tables 3, 4, 5). Closely related species share higher *MAT* similarity (Fig. 2), than do species that share a common reproductive mechanism (heterothallic *versus* homothallic).

## References

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. L. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Berbee, M. L., Pirseyedi, M. and Hubbard, S. 1999. *Cochliobolus* phylogenetics and the origin of pathogens, inferred from ITS and glyceraldehyde-3-phosphate dehydrogenase gene sequences. *In press*.
- Coppin, E., Debuchy, R., Arnais, S. and Picard, M. 1997. Mating types and sexual development in filamentous ascomycetes. *Microbiol. Mol. Biol. Rev.* 61:411-428.
- Metzenberg, R. L. and Glass, N. L. 1990. Mating type and mating strategies in *Neurospora*. *Bioessays* 12:53-59.
- Randall, T. A. and Metzenberg, R. L. 1995. Species-specific and mating type-specific DNA regions adjacent to mating type idiomorphs in the genus *Neurospora*. *Genetics* 141:119-136.
- Turgeon, B. G., Bohlmann, H., Ciuffetti, L. M., Christiansen, S. K., Yang, G., Schafer, W. and Yoder, O. C. 1993. Cloning and analysis of the mating type genes from *Cochliobolus heterostrophus*. *Mol. Gen. Genet.* 238:270-284.
- Turgeon, B. G. 1998. Application of mating type gene technology to problems in fungal biology. *Annu. Rev. Phytopathol.* 36:115-137.
- Wirsel, S., Horwitz, B., Yamaguchi, K., Yoder, O. C. and Turgeon, B. G. 1998. Single mating type-specific genes and their 3' UTRs control mating and fertility in *Cochliobolus heterostrophus*. *Mol. Gen. Genet.* 259:272-281.
- Yun, S-H., Berbee, M. L., Yoder, O. C. and Turgeon, B. G. 1999. Evolution of the fungal self-fertile reproductive life style from self-sterile ancestors. *Proc. Natl. Acad. Sci. USA* 96:5592-5597.