MOLECULAR PLANT PATHOLOGY (2007) **8**(1), 111–120 DOI: 10.1111/J.1364-3703.2006.00376.X

Isolation and characterization of the mating type locus of *Mycosphaerella fijiensis*, the causal agent of black leaf streak disease of banana

LAURA CONDE-FERRÁEZ[1,*], CEES WAALWIJK[2], BLONDY B. CANTO-CANCHÉ[1], GERT H. J. KEMA[2], PEDRO W. CROUS[3], ANDREW C. JAMES[1] AND EDWIN C. A. ABELN[3,†]

[1]*Centro de Investigación Científica de Yucatán (CICY), Calle 43 no. 130, Chuburná de Hidalgo, C.P. 97200, Mérida, Yucatán, México*
[2]*Plant Research International B.V., PO Box 16, 6700 AA, Wageningen, The Netherlands*
[3]*Centraalbureau voor Schimmelcultures, Fungal Diversity Centre, PO Box 85167, 3508 AD, Utrecht, The Netherlands*

SUMMARY

Idiomorphs *mat1-1* and *mat1-2* from *Mycosphaerella fijiensis*, the causal agent of black leaf streak disease of banana, were isolated. Degenerate oligos were used to amplify the HMG box of the *mat1-2* idiomorph from *M. fijiensis*, showing homology with the HMG box of *Mycosphaerella graminicola*. Using a DNA walking strategy, anchored on the DNA lyase gene towards the HMG box, a 9-kb-long region of *mat1-2* was obtained. A 5-kb fragment from the *mat1-1* region was obtained by long-range PCR using primers on the flanking regions, which have close to 100% identity between both idiomorphs. High-identity (77–89%), inverted regions within both idiomorphs were found, which suggest unique inversion events, which have not been found before, and that could have been significant in the evolution of this species. The predicted genes showed the conserved introns in both idiomorphs as well as an additional intron within the alpha box. The implications for the evolution of species in the *Mycosphaerella* complex on banana are discussed.

INTRODUCTION

Bananas (*Musa* spp.) are grown in all tropical regions of the world and play a key role in the economies of many developing countries. World consumption during 1998–2000 in developing countries was 21 kg per capita (mostly domestically produced), while the total value of the international banana trade ranges between US$4.5 and 5 billion per year (Arias *et al.*, 2003).

* *Correspondence*: Tel.: +52 99 99813914; Fax: +52 99 99813900; E-mail: lauracf@cicy.mx
† *Present address*: Promega Benelux BV, Schipholweg 1, PO Box 391, 2300 AJ Leiden, The Netherlands.

The crop is affected by several diseases and pests such as the foliar fungal pathogens *Mycosphaerella fijiensis*, *M. musicola* and *M. eumusae*, which all share similar morphologies and symptom development. *M. fijiensis* (anamorph *Pseudocercospora fijiensis*; Mycosphaerellaceae) is the causal agent of Black Sigatoka or black leaf streak disease (BLSD), which rapidly became the most devastating disease of banana production world-wide. It decreases photosynthesis, reduces fruit size and induces premature maturation. The cost of controlling the disease in large plantations is about US$1000 per hectare (Arias *et al.*, 2003), but it is higher in smaller plantations where fungicides cannot be applied by air and in which crop losses can be up to 50% (Mobambo *et al.*, 1993). Repercussions of the frequent and high input of fungicides include the development of both reduced sensitivity and resistance to these active compounds (Romero and Sutton, 1997; Sierotzki *et al.*, 2000). This was recently exemplified by the rapid development and spread of resistance to strobilurin fungicides in Central America (Marín *et al.*, 2003).

Owing to the fact that *Mycosphaerella* is one of the largest genera of plant pathogenic fungi, with more than 3000 *Mycosphaerella* species (Aptroot, 2006), Goodwin *et al.* (2004) proposed *Mycosphaerella graminicola* as the working model for the Dothideomycetes. As a result, the genomes of both *M. graminicola* and *M. fijiensis* are currently being sequenced within the DOE-JGI Community Sequencing Program (see http://www.jgi.doe.gov/sequencing/cspseqplans2006.html).

The sexual cycle of the fungus plays an important role in BLSD epidemiology (Gauhl *et al.*, 2000; Hayden *et al.*, 2003). Apart from the generation of air-borne inoculum, sexual reproduction results in genetic variation and contributes to evolution. In heterothallic ascomycetes, such as *M. fijiensis* (Mourichon and Zapater, 1990), mating can only occur between strains with opposite mating types. These mating types are determined by highly dissimilar sequences, called idiomorphs, that are embedded in regions common to all isolates of a given species, and from which

the only conserved regions are designated as the alpha box (for *mat1-1*) and the HMG box (for *mat1-2*) (Turgeon and Yoder, 2000). In addition, the structure of the *mat* idiomorphs has aided in understanding the evolution of heterothallic and homothallic species (Pöggeler, 1999; Turgeon, 1998; Yun *et al.*, 1999).

As *Mycosphaerella* leaf spot disease in banana is caused by a complex of at least three species (M. Arzanlou, personal communication), knowledge of the *mat* genes sequences in *M. fijiensis* is a starting point to a better understanding of the relevance of reproduction and recombination, in relation to the epidemiology of these important pathogens and the interaction with other species.

We presumed substantial synteny between *M. graminicola* and *M. fijiensis* as a basis to isolate the *mat* genes of the latter. Indeed, a PCR-based strategy using the DNA lyase gene, which flanks the idiomorphs in *M. graminicola*, allowed cloning the *mat1-2* idiomorph. In turn, the flanks of the *mat1-2* idiomorph were used to clone the *mat1-1* idiomorph by long-range PCR. Comparative analyses showed that both idiomorphs contain a highly unusual inversion not previously observed in idiomorphs in other ascomycetes.

RESULTS

PCR amplification of the HGM box and flanking genes

To amplify the HMG box from *mat1-2* isolates, primers reported for *M. graminicola* (Waalwijk *et al.*, 2002) and *Septoria passerinii* (Goodwin *et al.*, 2003) were unsuccessfully assayed. However, the degenerate primer pair KIKRP-F + SEKKR-R (-F for forward and -R for reverse) (Table 1) produced amplicons with the expected size of around 300 bp in some isolates of *M. fijiensis.* Therefore, these bands were cloned and sequenced. TBLASTx analysis revealed homology with the HMG box from *M. graminicola* (E = 6e-10) and *S. passerinii* (E = 1e-09), showing 71.9% identity in a predicted 82-amino-acid sequence. A multiple alignment of the predicted amino acid sequences is presented in Fig. 1a.

In addition, DNA lyase and *sla2* homologues in *M. fijiensis* were both amplified by degenerate PCR. A 1200-bp amplicon showed homology to the putative SLA2 protein (involved in cytoskeleton assembly, with no known function in mating) in *Aspergillus fumigatus* (E = 6e-85, 87% identity; GenPept accession no. EAL92953), *Magnaporthe grisea* (E = 6e-85, 78% identity; GenPept accession no. EAL92953) and *Neurospora crassa* (E = 6e-85, 73% identity; GenPept accession no. EAA35004).

The 850-bp amplicon showed homology to DNA lyase from several other fungi. Further on, a DNA walking strategy was employed using the DNA lyase as initial anchor. This strategy resulted in a 9817-bp genomic sequence containing the complete DNA lyase gene, a gene encoding the anaphase promoting complex (APC) and the *mat1-2* idiomorph (Fig. 2), which was deposited in GenBank with accession number DQ787016.

The DNA lyase sequence obtained is 1868 bp long; the predicted gene has a single exon and 622 amino acids. Local alignments showed up to 70% identity in a 106-amino-acid stretch (E = 2e-158; 82% similarities) and up to 75% identity on a 65-amino-acid portion (E = 1e-95, 78% similarities) with the DNA lyase sequence from *M. graminicola*. ClustalW global multiple alignments showed an overall identity of 40% with *M. graminicola* and *Cordyceps militaris.*

Tblastx analysis on the *mat1*-2 idiomorph showed only homology with the HMG box region of *mat1-2* from *M. graminicola*, with expected value E = 6e-69, on a 40-amino-acid stretch (67% identities and 77% similarities). The next hit was *mat1-2* from *S. passerinii,* with an expected value of E = 4e-27 (70% identities, 77% similarities). No other region with similarity was found within the *mat1-2* idiomorph.

Table 1 Primers used to amplify regions of the mating type idiomorphs, the DNA lyase gene and the *sla2* gene.

Primer	Sequence	Based on
KIKRP-F	5′-AAGATCAAGCGYCCAAAG-3′	*mat1-2* from *M. graminicola* and *S. passerinii*
SEKKR-R	5′-ATGCGRCGTTTCTTCTCSG-3′	*mat1-2* from *M. graminicola* and *S. passerinii*
DNAlydegF3	5′-CGCGCCGATCAGAGCNGARGARGG-3′	DNA lyase from *M. graminicola*, *N. crassa*, *F. graminearum*, *M. grisea* and *A. nidulans*
DNAlydegR3	5′-CATTCTTCTTCGTGTCCCARYRNGTRT-3′	DNA lyase from *M. graminicola*, *N. crassa*, *F. graminearum*, *M. grisea* and *A. nidulans*
DNAlyaseMfF	5′-TCTCACGTCGGTCCAGATG-3′	DNA lyase from *M. fijiensis*
DNAlyaseMfR	5′-TGAGGTTTCGTATCCGATGG-3′	DNA lyase from *M. fijiensis*
SLA2degF1	5′-CATCAACGACCCCAAYGARGGNTAYGA-3′	*sla2* from *A. nidulans*, *N. crassa* and *M. grisea*
SLA2degR1	5′-CCCGCTCVCGGATCATRTCNGC-3′	*sla2* from *A. nidulans*, *N. crassa* and *M. grisea*
E56	5′-CCCTCTGGACCAGGATACC-3′	Flanking region upstream of *M. fijiensis* idiomorphs
E66	5′-TCGCAAAATGAGTTGAAACG-3′	Flanking regions downstream of *M. fijiensis* idiomorphs
flan4739-F	5′-GCGGTTTTGGAGCGGTCAGG-3′	Flanking regions upstream of *M. fijiensis* idiomorphs
inver5656-R	5′-GAAGCTCTGGGTATCTCAGCACAGG-3′	Inverted sequence within idiomorphs (upstream *mat1-2*)
inver8486-F	5′-GCACCTCAGGGAGGCATTGG-3′	Inverted sequence within idiomorphs (downstream *mat1-2*)
flan9352-R	5′-TGATGCATCCTGCCGAGACC-3′	Flanking regions downstream of *M. fijiensis* idiomorphs

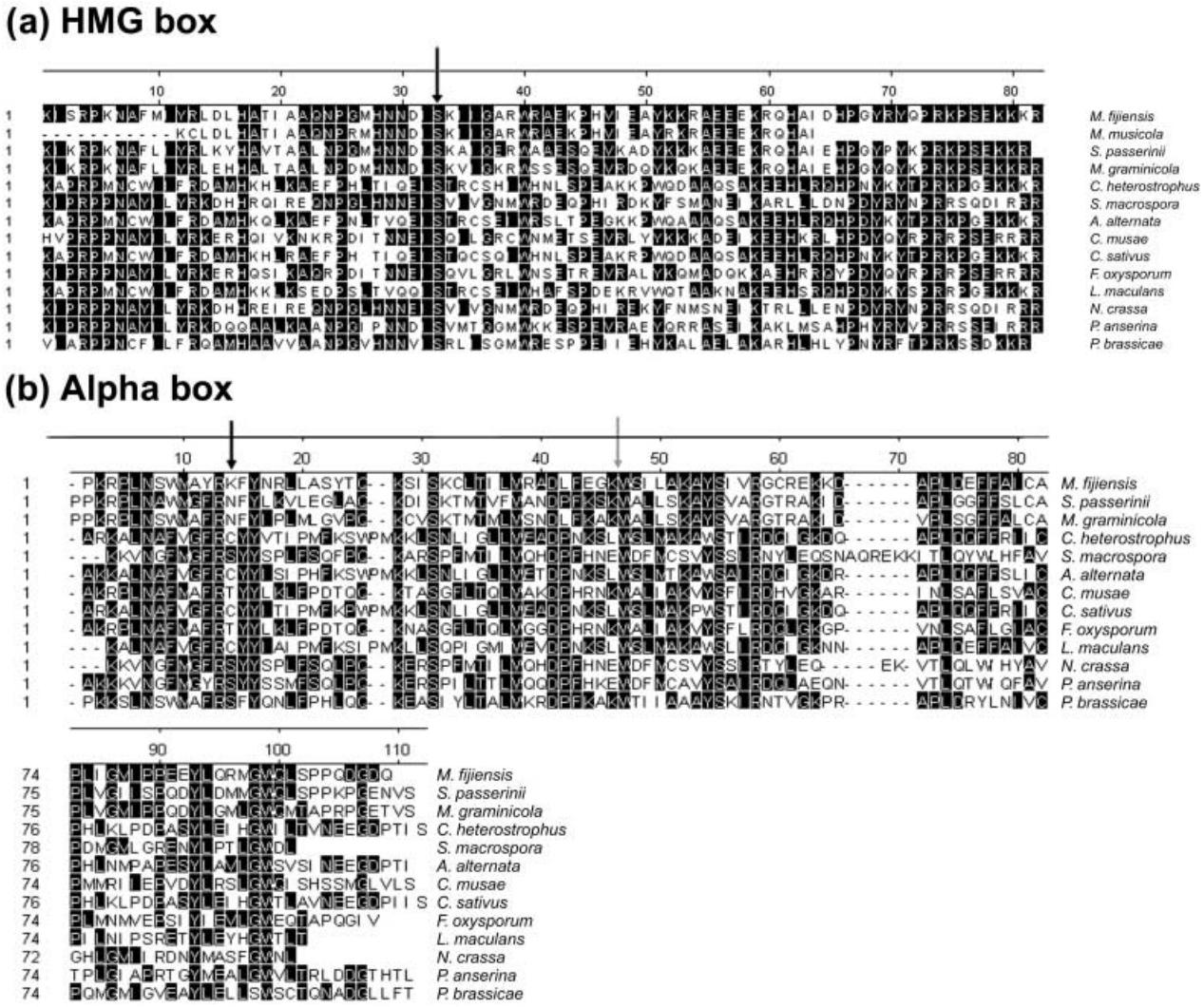


Fig. 1 Multiple sequence alignments of predicted amino acids from (a) the HMG box and (b) the alpha box of *Mycosphaerella fijiensis*, *Mycosphaerella musicola*, *Septoria passerinii*, *Cochliobolus heterostrophus*, *Sordaria macrospora*, *Alternaria alternata*, *Colletotrichum musae*, *Cochliobolus sativus*, *Fusarium oxysporum*, *Leptosphaeria maculans*, *Neurospora crassa*, *Podospora anserina* and *Pyrenopeziza brassicae*. Black arrows indicate the positions of the conserved introns; the grey arrow indicates another intron present only in the *Mycosphaerellacea fijiensis* alpha box. Shaded amino acids are in agreement with the consensus.

Long-range PCR isolation of the *mat1-1* idiomorph

Long-range PCRs using primers anchored on the flanking regions of the *mat1-2* idiomorph (primers E56 and E66, Table 1) resulted in a 5.8-kb fragment corresponding to *mat1-2* in some isolates, and a slightly smaller fragment (5.2 kb) in others. This 5.2-kb fragment was identified as containing the *mat1-1* idiomorph. The sequence was deposited in the GenBank database with accession number DQ787015. Tblastx analysis of the *mat1-1* sequence of *M. fijiensis* showed homology with *mat1-1* of *M. graminicola* and *S. passerinii* as first hits (E = 3e-75 and E = 5e-68, respectively). MegAlign and ClustalW multiple alignment of the amino-acid sequence showed a particularly conserved region, which corresponds to the alpha box of *mat1-1* (Fig. 1b).

Characterization of the idiomorphs

Primers that were developed on the flanking regions enabled the amplification of the complete idiomorphs (see Fig. 2). To define the flanking regions of the idiomorphs, the nucleotide sequences of *mat1-1* and *mat1-2* of *M. fijiensis* were aligned in ClustalW. The sequence similarity between the flanking regions, which had 95.2–98.4% identity, differed significantly from that of the idiomorphs (Fig. 3). We predicted the upstream ends of the idiomorphs *mat1-1* and *mat1-2* in positions 1004 bp and 4848 bp, respectively, and the downstream ends were determined on nucleotide 4877 of the *mat1-1* sequence, and nucleotide 9254 of the *mat1-2* sequence. Hence, the length of the *mat1-1* idiomorph is 3873 bp and that of *mat1-2* is 4406 bp.

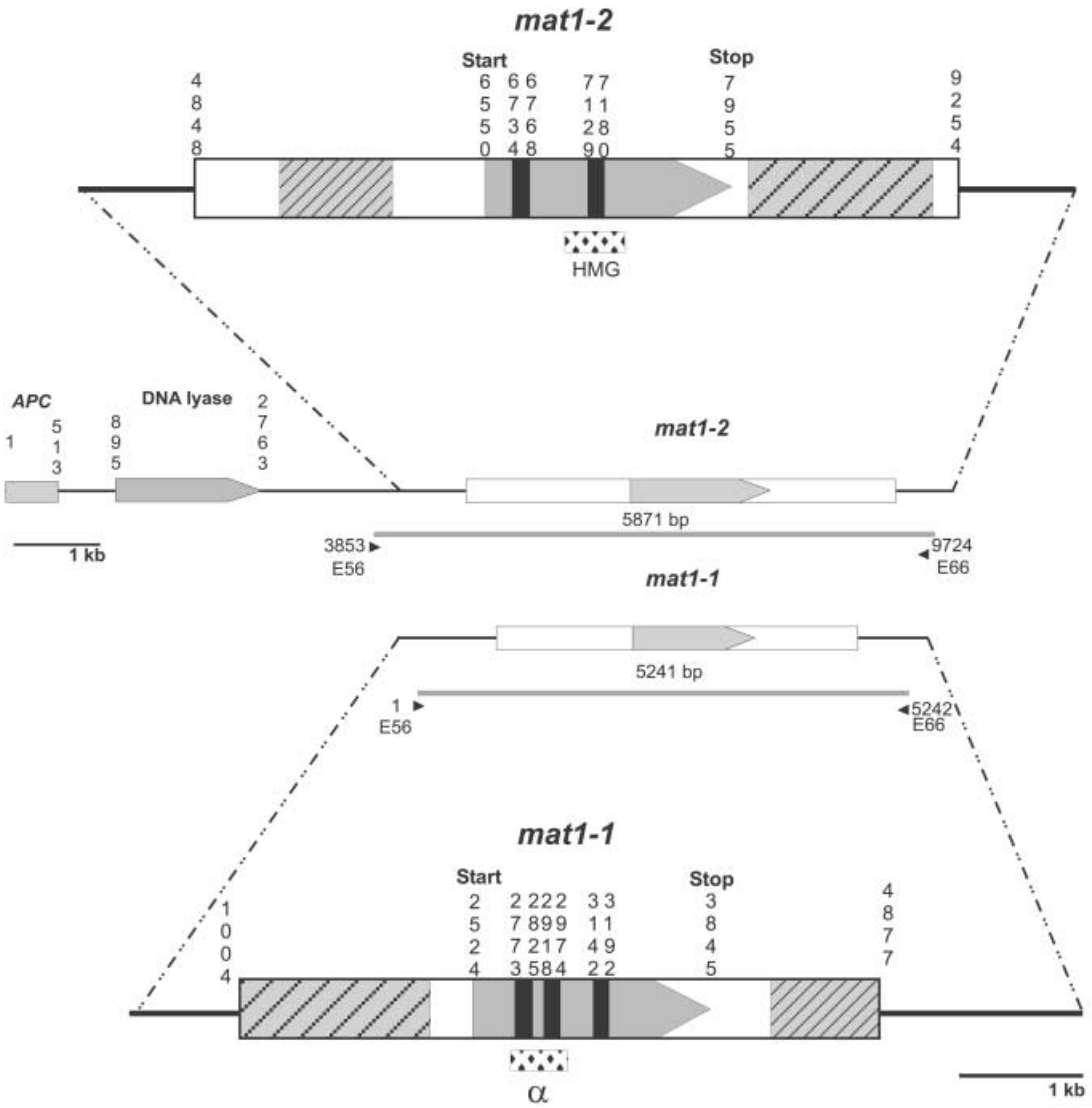


Fig. 2 Diagram of the *mat* locus organization in *Mycosphaerella fijiensis*. The boxes labelled *mat1-1* and *mat1-2* represent the idiomorphs; the adjacent genes are labelled *APC* (Anaphase Promoting Complex) and DNA lyase. Numbers indicate positions in the sequences of *mat1-1* (GenBank DQ787015) and *mat1-2* (GenBank DQ787016) from *M. fijiensis*. Numbered triangles represent primers used in long-range PCR. In expanded diagrams, the arrow-shaped boxes represent the predicted genes interrupted by introns (black boxes); the dashed boxes are inverted regions with high percentage identity between the plus (+) strand of one idiomorph and the minus (–) strand of the other.

Surprisingly, the alignment showed two regions with a high percentage of identity but in reverse orientation compared with the flanking sequences. These regions of inverted homology were located close to the ends of both idiomorphs (Fig. 3). In *M. fijiensis* a portion of 1127 bp in the 5′ end of the *mat1-1* idiomorph shows 77% identity (in nucleotide sequence) to a portion of the 3′ end of the *mat1-2* idiomorph, but in reverse orientation. The same is observed on the 3′ end of *mat1-1*, in which a 697-bp portion is highly similar (89% identity) to the 5′ end of *mat1-2* in the minus chain. Interestingly, this phenomenon was not observed when the idiomorphs of *M. graminicola*, *S. passerinii*, *Leptosphaeria maculans* or *Phaeosphaeria nodorum* were aligned among each other (data not shown). To rule out the possibility of assembly artefacts, we confirmed the presence of these inverted regions by PCR using primers on the flanking regions and others designed on the sequence of one of the idiomorphs. This approach allowed us to observe the amplification in one of the mating types and absence of an amplicon in the opposite mating type, and the reciprocal situation was observed when combining two reverse or two forward primers to produce amplicons from the opposite mating type (Fig. 4).

ORF finding and gene prediction

The predicted *M. fijiensis mat1-1* gene has three introns (Fig. 2); the conserved alpha region is included in a predicted protein of

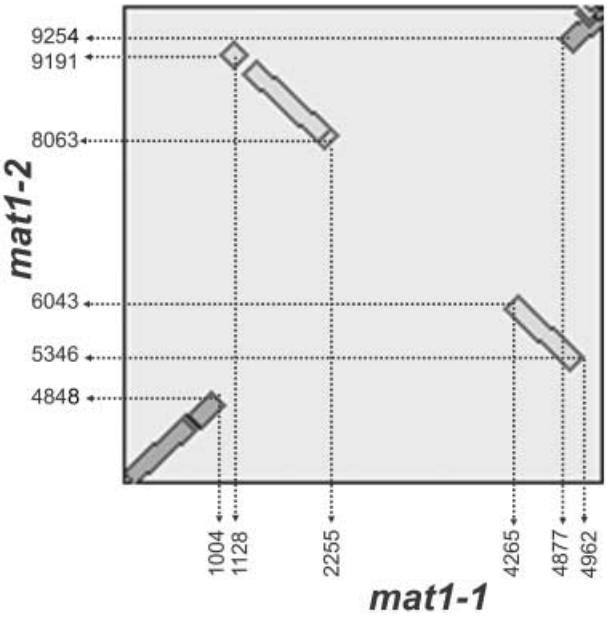


Fig. 3 Dot-plot graph of both idiomorphs of *M. fijiensis* generated with Blast2: *x*-axis, sequence of *mat1-1*; *y*-axis, sequence of *mat1-2*. The flanking regions draw two lines from bottom left to top right. The high-identity regions within the idiomorphs appear from top left to bottom right (between + and – strands). The numbers correspond to nucleotide positions on each sequence.

388 amino acids. In comparison, in *M. graminicola* the predicted MAT1-1 protein has 297 amino acids (Waalwijk *et al.*, 2002); the orthologous gene in *S. passerinii* (Goodwin *et al.*, 2003) has two introns as well, and the predicted protein has 310 residues. It is important to notice that two of the predicted introns are located in the alpha box. The first (2774–2824) is located in a conserved position shared by all ascomycetes (Fig. 1b). The second intron (in position 2919–2973) is unique to *Mycosphaerellacea fijiensis* and has not been reported in other fungi (Figs 1b and 2). The conserved sequence of the alpha box would be interrupted if this intron was not excised. The last intron is located in position 3143–3191 (Fig. 2). The excision of the last intron was confirmed by comparison with the *mat1-1* cDNA sequence (S. Zhong, personal communication). The predicted coding region of *mat1-1* of *M. fijiensis* is 1167 bp, which is in good agreement with those reported for *M. graminicola* and S. *passerinii.*

By contrast, the *M. fijiensis mat1-2* predicted gene has two introns and the protein has 441 amino acids. The intron positions are 6735–6767 bp and 7130–7179 bp (Figs 1a and 2).

Comparative analysis

When comparing the idiomorphs of *M. fijiensis* with those from *M. graminicola* and *S. passerinii*, using Blast2, the alignments showed that *mat1-1* sequences are closer to the diagonal than *mat1-2* (data not shown). This indicates a generally higher similarity between the *mat1-1* sequences, which is remarkable as it has been observed in ascomycetes that the HMG box shows a stronger conservation than the alpha box (Arie *et al.*, 1997). Nevertheless, in a short 39-amino-acid stretch of the HMG box, the comparison with *M. graminicola* was as high as 66% identity and 73% similarity, whilst *mat1-1* showed a lower percentage of similarity but in a longer stretch of 157 amino acids (36% identity, 46% similarity). ClustalW analysis of the conserved alpha and HMG amino acid sequences of several ascomycetes resulted in two phylogenetic trees (Fig. 5).

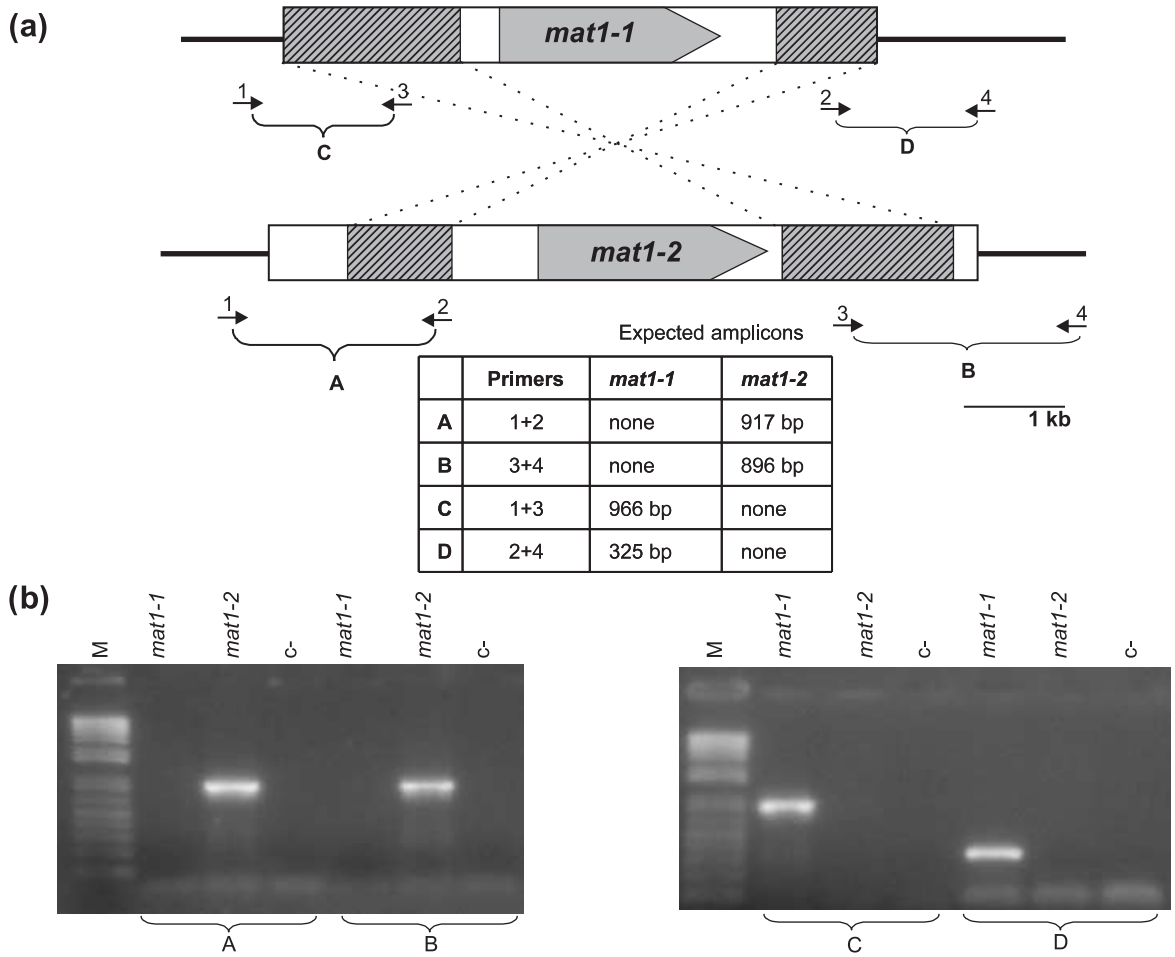
DISCUSSION

Amplification of the HMG box

We successfully used postulated synteny between *M. graminicola* and *M. fijiensis* to clone the *mat1-1* and *mat1-2* idiomorphs from the banana black leaf streak pathogen *M. fijiensis*. The HMG box of the *mat1-2* locus from several distantly related fungi has been obtained by PCR (Arie *et al.*, 1997), but reported primers did not amplify the HMG box in *M. fijiensis*. The HMG box has been used as an indicator to differentiate *Fusarium* species (O'Donnell *et al.*, 2004), and it has been proposed as a candidate to construct reliable phylogenetic trees (Yun *et al.*, 2000). Based on internal transcribed spacer (ITS) sequences, *M. graminicola* and *M. fijiensis* appeared to be closely related (Goodwin *et al.*, 2001). However, primers reported for *M. graminicola* and *S. passerinni* also failed to amplify the HMG box of *M. fijiensis*. Our sequence analyses of the *mat* orthologues and the flanking regions suggest that *M. fijiensis* and *M. graminicola* seem to be more distantly related than previously expected.

We did expect similarity between the two banana pathogens *M. fijiensis* and *M. musicola*, and indeed primers based on the *M. fijiensis* HMG box sequence successfully produced an amplicon in *M. musicola*. tBLASTx analysis of the *M. musicola* amplicon sequences (212 bp) showed homology with *mat1-2* from *S. passerinii* (E = 4e-04, 68% identities). Overall nucleotide sequences of the HMG box of *M. fijiensis* and *M. musicola* showed a high percentage of identities (85%) on pair-wise alignments. Predicted amino-acid sequence alignment of the HMG box of *M. fijiensis* and the fragment of the HMG box of *M. musicola* (53 amino acids) showed 92% identity (Fig. 1a). The similarities found between these banana pathogens are in agreement with the findings obtained with ITS analysis (Goodwin and Zismann, 2001), according to which *M. fijiensis* and *M. musicola* are closely related (*Pseudocercospora* anamorphs) as are *S. passerinii* and *M. graminicola* (*Septoria* anamorphs).

As seen in all fungi reported so far, a serine is found in the conserved intron position of the HMG box, which is present as well in *M. fijiensis* (Fig. 1a).



	Primers	*mat1-1*	*mat1-2*
A	1+2	none	917 bp
B	3+4	none	896 bp
C	1+3	966 bp	none
D	2+4	325 bp	none

Fig. 4 PCR confirmation of the inverted regions found within the idiomorphs of *M. fijiensis*. (a) Location of the primers and predicted PCR results if inversions occur (inserted table). Solid boxes represent the idiomorphs, the dashed boxes are inverted regions with high percentage identity between the plus (+) strand of one idiomorph and the minus (–) strand of the other; numbered arrows represent primers; primer 1 is flan4739, primer 2 is inver5656, primer 3 is inver 8486 and primer 4 is flan9352, which were used in four different combinations (A, B, C and D). (b) Agarose gel showing PCR results with primer pairs A, B, C and D. Primers 1 and 4 work as 'forward' and 'reverse' primers respectively for both idiomorphs, while primer 2 is 'reverse' for *mat1-2* but 'forward' for the inverted sequence in *mat1-1*; and primer 3 is 'forward' for *mat1-2* but 'reverse' for the inverted sequence in *mat1-1*. Therefore, if the inversions are present, in *mat1-2* the amplification only occurs if forward and reverse primers are combined (e.g. 1+2, 3+4); while in *mat1-1* the amplification only occurs if two forward or two reverse primers are combined (1+3, 2+4). DNAs employed were from strains 86 (mating type *mat1-1*) and 89 (mating type *mat1-2*); marker (M) is 1 kb Plus (Invitrogen®).

Characterization of the idiomorphs

We hypothesize that the inverted sequences in the *M. fijiensis* idiomorphs might have originated from an inversion event in a recent ancestor, given that independent blast analyses of these regions did not result in significant homologies in the databases, including the *M. graminicola* genome. The only exception was a hit with a small fragment of *mat1-1* of *S. passerinii* (E = 2e-04, 58%, 20/34 identities), on a non-coding region of the upstream end of the idiomorph. To our knowledge, the only two reports in which large portions of sequences of one idiomorph are found in the opposite is in *Cordyceps takaomontana* (anamorph: *Paecilomyces tenuipes*), where a *mat1-1-1* pseudogene is found in the *mat1-2* idiomorph (Yokoyama *et al.*, 2003); and *Aspergillus fumigatus* (Paoletti *et al.*, 2005), where a fragment of *mat1-2-1* is found within the flanking region of the *mat1-1* idiomorph. In addition, small fragments of common sequences were found in *Cochliobolus* spp., which were called 'islands of identity' of 8–9 bp, and may indicate recombination spots (Yun *et al.*, 1999).

It is not known how such dissimilar sequences of the idiomorphs can occupy the same locus in the genome, but it is thought that they were initially identical, but diverged through successive rearrangements and deletion/insertion events (Turgeon, 1998). It has also been suggested that the small fragments of identical sequence that have been found in both idiomorphs are probably remnants, explaining their common origin (Coppin *et al.*, 1997).

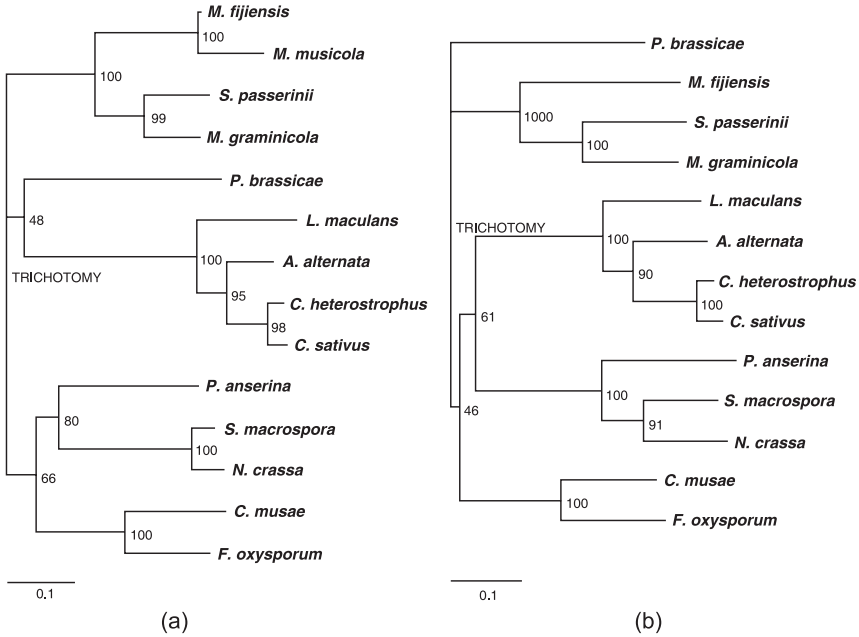


Fig. 5 Neighbour-joining phylograms based on multiple alignment of the predicted amino acid sequences from (a) the HMG box and (b) the alpha box, of several ascomycetes (same species presented in Fig. 1). Bootstrap support for each node is expressed as a percentage (of 1000 replications).

In *Stemphylium* spp. it was found that an inversion-fusion event gave rise to selfing species from outcrossing ancestors, resulting in a *mat* locus with both idiomorphs, one of which was inverted with respect to the heterothallic ancestor (Inderbitzin *et al.*, 2005). Conversely, the inverted regions found in *M. fijiensis* correspond to non-coding sequences. Whether a homothallic ancestor gave rise to *M. fijiensis* as a heterothallic species with this particular feature, as proposed for *Aspergillus* spp. (Paoletti *et al.*, 2005), is a question that can be addressed with the cloning of the mating type genes of other close *Mycosphaerella* relatives. These inverted regions represent an important finding, which can be the basis for further evolutionary studies on homo- and heterothallic species that are related to *M. fijiensis*.

Comparative analysis

As seen in *M. graminicola* (Waalwijk *et al.*, 2002) and other fungi (such as *Rhynchosporium secalis* and *Cochliobolus heterostrophus*) (Foster and Fitt, 2004; Turgeon *et al.*, 1993), the boundaries of the idiomorphs in *M. fijiensis* are defined by highly similar flanking regions, but in other species (e.g. *Phaeosphaeria nodorum* and *Neurospora crassa*) there is only a gradual transition from the flanking region to the idiomorphs (Bennett *et al.*, 2003; Randall and Metzenberg, 1998).

M. fijiensis idiomorphs (3873 bp for *mat1-1* and 4406 bp for *mat1-2*) are longer than the orthologues in *M. graminicola* (2839 bp for *mat1-1*, AF440399, and 2772 bp for *mat1-2*, AF440398), *S. passerinii* (3048 bp for *mat1-1*, AF483193, and 2897 bp for *mat1-2*, AF483194) and *Xanthoria polycarpa* in which the *mat1-1* idiomorph is 3270 bp (AJ884599) and *mat1-2* is 3150 bp (AJ884598). The *M. fijiensis* idiomorphs seem to be more similar in size to more distantly related species such as *P. nodorum*, in which *mat1-1* is 4282 bp (AY212018) and *mat1-2* is 4505 bp (AY212019); *Fusarium oxysporum* (*mat1-1* is 4618 bp, AB011379, and *mat1-2* is 3849 bp, AB011378) or *R. secalis* (*mat1-1* is 4049 bp, EMBL: AJ537511 and *mat1-2* is 3153 bp, EMBL: AJ549759). Conversely, other single-gene idiomorphs can be much shorter, for example *C. heterostrophus* idiomorph *mat1-1* is 1297 bp (AF029913) and *mat1-2* is 1171 bp (AF027687).

Synteny among ascomycetes

We used synteny to isolate the idiomorphs of *M. fijiensis* by DNA walking anchored on a flanking gene. An extraordinary level of synteny has recently been reported in the mating type region of distantly related fungi such as *Aspergillus nidulans*, *M. grisea*, *N. crassa*, *Fusarium proliferatum* and *Gibberella zeae* (Waalwijk *et al.*, 2004). It has been proposed that the conservation in gene order in the vicinity of the *mat* locus is due to suppressed recombination that is caused by the dissimilarity of the idiomorph sequences. The *mat* locus of several ascomycete species has been characterized in detail. In yeast species such as *Saccharomyces cerevisiae*, *S. castellii*, *Candida glabrata* and *Kluyveromyces delphensis*, the genes encoding BUD5 and HO endonuclease, which are required for mating type switching, are located in the region of the idiomorphs; in other species such as *S. kluyveri*, *K. lactis*, *Pichia angusta* and *Yarrowia lipolytica*, the *mat* locus is located next to the *sla2* gene (Butler *et al.*, 2004; Debuchy and Turgeon, 2006). This gene is also located near the *mat* locus in filamentous fungi such as *N. crassa* (AABX01000036) and

several species of *Fusarium* (Waalwijk *et al.*, 2004). The DNA lyase and APC genes flank the *mat* idiomorphs in *M. graminicola* (Waalwijk *et al.*, 2002). In *L. maculans* an orthologue of the DNA lyase gene was also found in the same location (Cozijnsen and Howlett, 2003). In fact, data mining performed during this work showed that orthologues of the DNA lyase gene are found near the *mat* locus in *Glomerella cingulata* (AY357890), *Sordaria macrospora* (Y10616), *P. tenuipes* (AB084921) and *C. militaris* (AB084257). Other examples of fungi in which both the DNA lyase and the *sla2* genes flank the idiomorphs are *Y. lipolytica* (CR382129), *Xanthoria parietina* and *X. polycarpa* (AJ884600 and AJ884598).

However, in *M. fijiensis*, it is unlikely that the *sla2* gene is located near the *mat* locus. The 1200-bp *sla2* amplicon that we obtained in our study was used as a probe against a *M. fijiensis* BAC library (90 kb mean insert size) at high stringency conditions (D.K. Guillén-Maldonado *et al.*, unpublished data). The *sla2* homologue was present in at least six different BAC clones (data not shown), but hybridizations with different heterologous and homologous *mat* probes suggested that *sla2* and the *mat* idiomorphs are not physically linked in *M. fijiensis*, which is supported by the draft sequence of the *M. graminicola* genome (DOE-JGI Community Sequencing Program).

The use of the idiomorphs for the study of populations and evolution

The synteny observed in gene order and intron position, in contrast to the high sequence divergence within the idiomorphs, reflects the importance of the *mat* locus in determining a species barrier. Because of its polymorphism, the *mat* locus can be used as a marker for population studies, and it is currently being used for genetic mapping (G. Manzo-Sánchez *et al.*, unpublished data).

Population genetic studies support the hypothesis that in *M. fijiensis* sexual reproduction is random and frequent (Carlier *et al.*, 1996; Hayden *et al.*, 2003; Rivas *et al.*, 2004). The identification of the *mat* genes will provide further evidence for this observation, because populations with a high occurrence of sexual reproduction would have strains of opposite mating type distributed in a 1 : 1 ratio (Zhan *et al.*, 2002).

The highly similar inverted regions that we identified within the idiomorphs of *M. fijiensis*, and the additional intron found within the alpha box are unique features, which will be useful in evolution studies. Ongoing studies have shown that both are present in the *mat* genes from *M. musicola*, suggesting that these events occurred before speciation (L. Conde-Ferráez, unpublished results). Future work focusing on these characteristics would give additional information about the evolution and ecology of the genus *Mycosphaerella*. As the *Mycosphaerella* pathogens of banana constitute a complex of species that have coexisted and interacted on their common host, the analysis of their mating type loci would give insights for a better understanding on their relationship and evolution.

EXPERIMENTAL PROCEDURES

Fungal isolates and DNA extraction

Monoascoporic strains of *M. fijiensis* were obtained from diverse sources (Table 2) and were grown at 26 °C in potato dextrose broth (PDB) with continuous shaking (100 r.p.m.) under continuous light. DNA was extracted from mycelium collected after filtration by grinding under liquid nitrogen according to the protocol described by Johanson (1997). We used DNAs from *M. graminicola* and *S. passerinii* isolates of known mating type as comparisons in degenerate PCRs. The list of DNAs used in this study is summarized in Table 2.

PCR and DNA walking strategies

We aligned the *mat* genes and flanking sequences of *M. graminicola* and its close relative *S. passerinii*, and the DNA lyase and the *sla2*

Table 2 DNAs used in this work. Grand Naine is a cultivar of *Musa acuminata*, which is highly susceptible to *M. fijiensis*.

Isolate	Species	Host	Mating type	Geographical origin	Reference
IPO323	*M. graminicola*	Wheat	*mat1-1*	The Netherlands	Kema *et al.* (2002)
IPO94269	*M. graminicola*	Wheat	*mat1-2*	The Netherlands	Kema *et al.* (2002)
P64	*S. passerinii*	Barley	*mat1-1*	USA	Goodwin *et al.* (2003)
P76	*S. passerinii*	Barley	*mat1-2*	USA	Goodwin *et al.* (2003)
86	*M. fijiensis*	Grand Naine	A*	Cameroon	Mourichon and Zapater (1990)
89	*M. fijiensis*	Grand Naine	a*	Cameroon	Mourichon and Zapater (1990)
139a	*M. fijiensis*	Grand Naine	a*	Colombia	Mourichon and Zapater (1990)
138	*M. fijiensis*	Grand Naine	A*	Colombia	Mourichon and Zapater (1990)
1231	*M. fijiensis*	Grand Naine	*mat1-2*	Tabasco, Mexico	This work
1233	*M. fijiensis*	Grand Naine	*mat1-1*	Tabasco, Mexico	This work

*The mating type tester strains have been traditionally designated as 'A' and 'a'; in this work it was determined that 'A' mating type corresponds to *mat1-1*, while 'a' corresponds to *mat1-2*.

genes from several other ascomycetes, to develop degenerate primers using the Codehop program (http://nar.oxfordjournals.org/cgi/content/full/31/13/3763). Specific primers were developed using the Primer3 program (Rozen and Skaletsky, 2000). Only those primers that yielded amplicons are detailed in Table 1.

PCRs were performed in an Eppendorf thermal cycler. General cycling conditions were 94 °C for 1 min, 30 cycles of 94 °C for 1 min, annealing at 55 °C for 1 min, and 72 °C for 1 min, followed by a final extension of 7 min at 72 °C. For degenerate PCR, with primers based on the DNA lyase, *sla2* and the idiomorphs from other fungi, a touchdown PCR was used, from 60 to 55 °C for the annealing temperature (–1 °C per cycle), with 30 additional cycles at an annealing temperature of 55 °C.

To obtain the *mat1-2* idiomorph, a DNA walking strategy was performed with the DNAwalking SpeedUp™ Premix Kit (Seegene, Seoul, Republic of Korea), using the DNA lyase gene as anchor according to the specifications given in the DNA walking manual. The full length of the *mat1-1* idiomorph was obtained by long-range PCR with primers E56 and E66 (Table 1), which were designed based on the *mat1-2* sequence, and are located in a region outside the idiomorph. The annealing temperature was 55 °C for 40 s; extension was at 72 °C for 3.5 min. The resulting 5.5-kb PCR product was sonicated, followed by treatment with T4 DNA polymerase (Promega Benelux b.v., Leiden, The Netherlands) to obtain blunt ends and treated with kinase (T4 polynucleotide kinase, NEB) for 30 s. The fragments were then ligated into the pUC19 vector linearized with *Sma*I (Promega). Clones with insert sizes between 300 and 1200 bp were selected to be sequenced. Other amplicons were either cloned into the pCR2.1-TOPO vector (Invitrogen®), according to the manufacturer's instructions or were purified and sequenced directly. Sequencing was performed using Big Dye® Terminator technology (Applied Biosystems, Foster City, USA).

PCR confirmation of inverted sequences within the idiomorphs

Primers flan4739-F, inver5656-R, inver8486-F and flan9352-R (Table 1) were designed on the predicted inverted regions and on the flanking regions of the idiomorphs; the orientations (forward or reverse) correspond to the sequence of *mat1-2*. These primers were used in four different combinations (A, B, C and D), which are specified in the table inserted in Fig. 4. Annealing was at 62 °C for 40 s, and extension at 72 °C for 40 s.

Bioinformatic analyses

Sequences were assembled and edited in the SeqMan and EditSeq programs (DNAstar, Lasergene™) and analysed using Blast (http://www.ncbi.nlm.nih.gov). Sequence alignments were performed by using ClustalW (http://www.ebi.ac.uk/clustalw/), Blast2-sequences (http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi) and MegAlign (Lasergene DNAstar™). Neighbour-joining trees were constructed based on alignments of the conserved alpha and HMG domains, respectively, using the tree-drawing application in Clustal X. Bootstrap analysis (1000 replicates) was performed to evaluate the degree of support for each group in the tree.

Identification of open reading frames (ORFs) and gene predictions were performed using FGENESH and FGENESH+ software (Softberry™, http://www.softberry.com/berry.phtml) with the codon usage of *M. grisea*. The results were compared with those obtained with the GenScan (http://genes.mit.edu/GENSCAN.html) and GenomeScan (http://genes.mit.edu/genomescan.html) programs, using as homologue models the MAT proteins of *M. graminicola* and *S. passerinii*. Conserved intron boundaries (GT/AG) and branching signals (RCURAY) most commonly found in other fungi (Kupfer *et al.*, 2004) were also identified in the predicted genes.

ACKNOWLEDGEMENTS

We thank I. de Vries, O. Mendes, T. van der Lee (PRI, The Netherlands) and M. Arzanlou (CBS, The Netherlands) for technical support; J. Carlier (CIRAD, France), S. Goodwin (Purdue University, USA), G. Manzo (CICY, México) for providing fungal strains and DNAs; and S. Zhong (Department of Plant and Environmental Protection Sciences, University of Hawaii, USA) for sharing a cDNA sequence from the *mat1-1* gene. L. Cámara Ferráez is acknowledged for valuable support with figure preparation and S. Peraza-Echeverría for support with preparation of Fig. 5. L.C.-F. is the recipient of scholarship CONACYT (70133) and this work was partially supported by project CONACYT 37602-B; the work was developed in the CBS, CICY and PRI facilities.

REFERENCES

Aptroot, A. (2006) *Mycosphaerella* and its anamorphs: 2. Conspectus *Mycosphaerella*. *CBS Biodiv. Series*, **5**, in press.

Arias, P., Dankers, C., Liu, P. and Pilkauskas, P. (2003) *The World Banana Economy 1985–2002*. FAO Commodity studies. Rome: Food and Agriculture Organization of the United Nations.

Arie, T., Christiansen, S.K., Yoder. and Turgeon, B.G. (1997) Efficient cloning of ascomycete mating type genes by PCR amplification of the conserved *MAT* HMG box. *Fungal Genet. Biol.* **21**, 118–130.

Bennett, R.S., Yun, S.-H., Lee, T.Y., Turgeon, B.G., Arseniuk, E., Cunfer, B.M. and Bergstrom, G.C. (2003) Identity and conservation of mating type genes in geographically diverse isolates of *Phaeosphaeria nodorum*. *Fungal Genet. Biol.* **40**, 25–37.

Butler, G., Kenny, C., Fagan, A., Kurischko, C., Gaillardin, C. and Wolfe, K.H. (2004) Evolution of the *mat* locus and its Ho endonuclease in yeast species. *Proc. Natl Acad. Sci. USA*, **101**, 1632–1637.

Carlier, J., Lebrun, M.H., Zapater, M., Dubois, C. and Mourichon, X. (1996) Genetic structure of the global population of banana black leaf streak fungus, *Mycosphaerella fijiensis*. *Mol. Ecol.* **5**, 499–510.

Coppin, E., Debuchy, R., Arnaise, S. and Picard, M. (1997) Mating types

and sexual development in filamentous ascomycetes. *Microbiol. Mol. Biol. R.* **61**, 411–428.

Cozijnsen, A.J. and Howlett, B.J. (2003) Characterisation of the mating type locus of the plant pathogenic ascomycete *Leptosphaeria maculans*. *Curr. Genet.* **43**, 351–357.

Debuchy, R. and Turgeon, B.G. (2006) Mating-type structure, evolution and function in euascomycetes. In: *The Mycota I. Growth, Differentiation and Sexuality* (Kües, U. and Fischer, R., eds), pp. 293–323. Berlin/Heidelberg: Springer Verlag.

Foster, S.J. and Fitt, B.D. (2004) Isolation and characterisation of the mating-type (MAT) locus from *Rhynchosporium secalis*. *Curr. Genet.* **44**, 277–286.

Gauhl, F., Pasberg-Gauhl, C. and Jones, D.R. (2000) Black leaf streak. Disease cycle and epidemiology. In: *Diseases of Banana, Abaca and Enset* (Jones, D.R., ed.), pp. 56–62. Wallingford, UK: CAB International.

Goodwin, S.B., Dunkle, L.D. and Zismann, V.L. (2001) Phylogenetic analysis of *Cercospora* and *Mycosphaerella* based on the internal transcribed spacer region of ribosomal DNA. *Phytopathology*, **91**, 648–658.

Goodwin, S.B., Waalwijk, C. and Kema, G.H.J. (2004) Genetics of Mycosphaerella graminicola: a model for the Dothideales. In: *Applied Mycology and Biotechnology. An International Series, Vol. 4, Fungal Genomics* (Arora, D.K. and Khachatourians, G., eds), Chap.13, pp. 315–330. Amsterdam: Elsevier B.V.

Goodwin, S.B., Waalwijk, C., Kema, G.H.J., Cavaletto, J.R. and Zhang, G. (2003) Cloning and analysis of the mating type idiomorphs from the barley pathogen *Septoria passerinii*. *Mol. Gen. Genom.* **269**, 1–12.

Goodwin, S.B. and Zismann, V.L. (2001) Phylogenetic analyses of the ITS region of ribosomal DNA reveal that *Septoria passerinii* from barley is closely related to the wheat pathogen *Mycosphaerella graminicola*. *Mycologia*, **93**, 934–946.

Hayden, H.L., Carlier, J. and Aitken, E.A.B. (2003) Genetic structure of *Mycosphaerella fijiensis* populations from Australia, Papua New Guinea and the Pacific Islands. *Plant Pathol.* **52**, 703–712.

Inderbitzin, P., Harkness, J., Turgeon, B.G. and Berbee, M. (2005) Lateral transfer of mating system in *Stemphylium*. *Proc. Natl Acad. Sci. USA*, **102**, 11390–11395.

Johanson, A. (1997) *Detection of Sigatoka Leaf Spot Pathogens of Banana by the Polymerase Chain Reaction.* Chatham, UK: Natural Resources Institute.

Kema, G.H.J., Goodwin, S.B., Hamza, S., Verstappen, E.C.P., Cavaletto, J.R., Van der Lee, T.A.J., de Weerdt, M., Bonants, P.J.M. and Waalwijk, C. (2002) A combined amplified fragment length polymorphism and randomly amplified polymorphism DNA genetic linkage map of *Mycosphaerella graminicola*, the *Septoria tritici* leaf blotch pathogen of wheat. *Genetics*, **161**, 1497–1505.

Kupfer, D.M., Drabenstot, S.D., Buchanan, K.L., Lai, H., Zhu, H., Dyer, D.S., Roe, B.A. and Murphy, J.W. (2004) Introns and splicing elements of five diverse fungi. *Eukaryot. Cell*, **3**, 1088–1100.

Marín, D.H., Romero, M.A., Guzmán, M. and Sutton, T.B. (2003) Black Sigatoka: an increasing threat to banana cultivation. *Plant Dis*, **87**, 208–222.

Mobambo, K.N., Gauhl, F., Vuylsteke, D., Ortiz, C., Pasberg-Gauhl, C. and Swennen, R. (1993) Yield loss in plantain from black Sigatoka leaf spot and field performance of resistant hybrids. *Field Crops Res.* **35**, 35–42.

Mourichon, X. and Zapater, M.F. (1990) Obtention *in vitro* du stade *Mycosphaerella fijiensis*, forme parfaite de *Cercospora fijiensis*. *Fruits*, **45**, 553–557.

O'Donnell, K., Ward, T.J., Geiser, D.M., Kistler, C.H. and Aoki, T. (2004) Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal Genet. Biol.* **41**, 600–623.

Paoletti, M., Rydholm, C., Schwier, E.U., Anderson, M.J., Szakacs, G., Lutzoni, F., Debeaupuis, J.P., Latgé, J.P., Denning, D.W. and Dyer, P.S. (2005) Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Curr. Biol.* **15**, 1242–1248.

Pöggeler, S. (1999) Phylogenetic relationships between mating-type sequences from homothallic and heterothallic ascomycetes. *Curr. Genet.* **36**, 222–231.

Randall, T.A. and Metzenberg, R.L. (1998) The mating type locus of *Neurospora crassa*: identification of an adjacent gene and characterization of transcripts surrounding the idiomorphs. *Mol. Gen. Genet.* **259**, 615–621.

Rivas, G.G., Zapater, M.F., Abadie, C. and Carlier, J. (2004) Founder effects and stochastic dispersal at the continental scale of the fungal pathogen of bananas *Mycosphaerella fijiensis*. *Mol. Ecol.* **13**, 471–482.

Romero, R.A. and Sutton, T.B. (1997) Sensitivity of *Mycosphaerella fijiensis* causal agent of Black Sigatoka of banana to propiconazole. *Phytopathology*, **87**, 96–100.

Rozen, S. and Skaletsky, H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (Krawetz, S. and Misener, S., eds), pp. 365–386. Totowa, NJ: Humana Press.

Sierotzki, H., Parisi, S., Steinfeld, U., Tenzer, I., Poirey, S. and Gisi, U. (2000) Mode of resistance to respiration inhibitors at the cytochrome bc1 enzyme complex of *Mycosphaerella fijiensis* field isolates. *Pest. Manage. Sci.* **56**, 833–841.

Turgeon, B.G. (1998) Application of mating type gene technology to problems in fungal biology. *Annu. Rev. Phytopathol.* **36**, 115–137.

Turgeon, B.G., Bohlmann, H., Ciuffetti, L.M., Christiansen, S.K., Yang, G., Schäfer, 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. and Yoder, O.C. (2000) Proposed nomenclature for mating type genes of filamentous ascomycetes. *Fung. Genet. Biol.* **31**, 1–5.

Waalwijk, C., Mendes, O., Verstappen, E., de Waard, M. and Kema, G.H.J. (2002) Isolation and characterization of the mating-type idiomorphs from the wheat septoria leaf blotch fungus *Mycosphaerella graminicola*. *Fungal Genet. Biol.* **35**, 277–286.

Waalwijk, C., van der Lee, T., de Vries, I., Hesselink, T., Arts, J. and Kema, G.H.J. (2004) Synteny in toxigenic *Fusarium* species: the fumonisin gene cluster and the mating type region as examples. *Eur. J. Plant Pathol.* **110**, 533–544.

Yokoyama, E., Yamagishi, K. and Hara, A. (2003) Structures of the mating-type loci of *Cordyceps takaomontana*. *Appl. Environ. Microbiol.* **69**, 5019–5022.

Yun, S.-H., Arie, T., Kanwko, I., Yoder, O.C. and Turgeon, B.G. (2000) Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella/Fusarium* species. *Fungal Genet. Biol.* **31**, 7–20.

Yun, S., Berbee, M., 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**, 5592–2297.

Zhan, J., Kema, G.H.J., Waalwijk, C. and McDonald, B.A. (2002) Distribution of mating type alleles in the wheat pathogen *Mycosphaerella graminicola* over spatial scales from lesions to continents. *Fungal Genet. Biol.* **36**, 128–136.