

1 **Intraspecific comparison and annotation of two complete mitochondrial**
2 **genome sequences from the plant pathogenic fungus *Mycosphaerella***
3 ***graminicola***

4
5 **Stefano F. F. Torriani • Stephen B. Goodwin • Gert H. J. Kema • Jasmyn L. Pangilinan**
6 **• Bruce A. McDonald**

7
8
9 S. F. F. Torriani • B. A. McDonald ()
10 Plant Pathology Group, Institute of Integrative Biology
11 ETH Zurich/LFW, Universitätstrasse 2
12 CH-8092 Zurich, Switzerland
13 e-mail: bruce.mcdonald@agrl.ethz.ch
14 Telephone: +41(0) 44 632 3847, Fax: +41(0) 44 632 1572

15
16 S. B. Goodwin
17 USDA-ARS, Crop Production and Pest Control Research Unit,
18 Department of Botany and Plant Pathology, 915 West State Street,
19 Purdue University, West Lafayette, IN 47907-2054, USA

20
21 G. H. J. Kema
22 Plant Research International B.V., P.O. Box 16,
23 6700 AA Wageningen, The Netherlands

24
25 J. L. Pangilinan
26 Department of Energy Joint Genome Institute
27 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

28 **Abstract**

29 The mitochondrial genomes of two isolates of the wheat pathogen *Mycosphaerella*
30 *graminicola* were sequenced completely and compared to identify polymorphic regions. This
31 organism is of interest because it is phylogenetically distant from other fungi with sequenced
32 mitochondrial genomes and it has shown discordant patterns of nuclear and mitochondrial
33 diversity. The mitochondrial genome of *M. graminicola* is a circular molecule of
34 approximately 43,960 bp containing the typical genes coding for 14 proteins related to
35 oxidative phosphorylation, one RNA polymerase, two rRNA genes and a set of 27 tRNAs.
36 The mitochondrial DNA of *M. graminicola* lacks the gene encoding the putative ribosomal
37 protein (*rps5*-like), commonly found in fungal mitochondrial genomes. Most of the tRNA
38 genes were clustered with a gene order conserved with many other ascomycetes. A sample of
39 thirty-five additional strains representing the known global mt diversity was partially
40 sequenced to measure overall mitochondrial variability within the species. Little variation
41 was found, confirming previous RFLP-based findings of low mitochondrial diversity. The
42 mitochondrial sequence of *M. graminicola* is the first reported from the family
43 Mycosphaerellaceae or the order Capnodiales. The sequence also provides a tool to better
44 understand the development of fungicide resistance and the conflicting pattern of high
45 nuclear and low mitochondrial diversity in global populations of this fungus.

46

47 *Keywords:* Comparative genomics • Genome organization • Microsatellites • Mitochondrial
48 genome (mtDNA) • *Septoria tritici*

49

50

51 **1. Introduction**

52 *Mycosphaerella graminicola* (anamorph *Septoria tritici*) is the causal agent of *Septoria*
53 *tritici* blotch of wheat and other poaceous hosts, and occurs worldwide across a wide range of

54 climates (Eyal, 1999). The life cycle of *M. graminicola* includes both sexual and asexual
55 stages. The sexual stage permits genetic recombination and produces airborne ascospores
56 with the potential to be dispersed over several kilometers (Sanderson, 1972), whereas the
57 asexual phase (*S. tritici*) produces pycnidiospores disseminated over limited distances from
58 plant to plant via rain splash (Bannon and Cooke, 1998).

59 In most fungi studied to date there is concordance between genetic variation in the
60 mitochondrial (mt) and nuclear genomes (Sommerhalder et al., 2007; Zhan et al., 2004), with
61 some fungi having high levels of mt and nuclear diversity (Kudla et al., 2002; Liu et al.,
62 1996) and others having low mt and nuclear genetic variability (Kurdyla et al., 1995; Xia et
63 al., 2000). However, the pattern of variability in *M. graminicola* is different; a comparison of
64 RFLP markers in nuclear and mt genomes showed a pattern of high nuclear and low mt
65 diversity in populations around the world (Zhan et al., 2003). Over 1300 nuclear RFLP
66 genotypes were found among 1673 isolates, with an average of 18 alleles per nuclear RFLP
67 locus. In contrast, only seven mtDNA haplotypes were found globally, with the two most
68 common representing 93% of the world population. The high nuclear diversity is thought to
69 be the consequence of high gene flow (Boeger et al., 1993), coupled with large effective
70 population sizes (Zhan and McDonald, 2004) and recurring sexual reproduction (Chen and
71 McDonald, 1996; Kema et al., 1996; Hunter et al., 1999). Zhan et al. (2004) suggested a
72 selective sweep to explain the low diversity found in the mtDNA. Selective sweeps may be
73 common in mt genomes because all of the genes are linked in one molecule so that selection
74 on one gene can affect the frequency of all genes through hitchhiking.

75 Mt genomes have proven to be highly useful for research in evolutionary biology and
76 systematics because of their uniparental inheritance, the near absence of genetic
77 recombination, and uniform genetic backgrounds (Chen and Hebert, 1999). The evolution of
78 mtDNAs has been characterized by extensive loss and translocation of genes to the nucleus

79 (Adams et al., 2000) since their origin by endosymbiosis of a bacterial ancestor (John and
80 Whatley, 1975). The result of this process is that most mt proteins are encoded by nuclear
81 genes whose products are imported into the mitochondrion by translocase complexes, leaving
82 relatively few mt proteins that are synthesized directly within the organelle (Brennicke et al.,
83 1993; Hartl et al., 1989).

84 Mt genomes are characterized by high A + T content, lack of methylation, conservation in
85 gene function, and high copy number (Campbell et al., 1999), and they can evolve at their
86 own rate relative to the nuclear genomes of the organisms in which they occur (Ballard and
87 Whitlock, 2004). The size and topology of the mt genome, the number and nature of the
88 proteins it encodes, and even the genetic code itself can vary greatly among species (Gray et
89 al., 1999). Fungal mtDNAs are generally an order of magnitude smaller than those of plants
90 but larger than animal mtDNAs (Burger et al., 2003) and usually contain 14 genes encoding
91 hydrophobic subunits of respiratory chain complexes, as well as genes for the large (*rnl*) and
92 small (*rns*) ribosomal subunits and a set of tRNAs (Gray et al., 1999). The coding percent
93 ranges between 40 and 60% in the Pezizomycotina. Among fungi, mt genomes vary widely
94 in size, from approximately 18 to 109 kb (NCBI database). The variability of mt genome size
95 among species is strongly influenced by differences in length and organization of intergenic
96 regions, as well as by differences in intron content (from 0 to 30) and size (ranging from 0.15
97 and 4 kb). Burger et al. (2003) showed that there is no correlation between mtDNA size and
98 gene content.

99 The taxonomic placement of *Mycosphaerella* within the class Dothideomycetes until
100 recently was uncertain, and it usually was placed near *Dothidea* in the Dothideales (Kirk et
101 al., 2001; Goodwin et al., 2004). However, recent analyses of a multigene phylogeny showed
102 that *Mycosphaerella* belongs in the Capnodiales, a sister group to the Dothideales and
103 Myriangiiales (Schoch et al., 2006). Though the mt genome of *Stagonospora nodorum*,

104 another member of the Dothideomycetes, was recently published (Hane et al., 2007), no mt
105 genomes have been published from *Mycosphaerella* or any species in the Capnodiales,
106 Dothideales, or Myriangiales.

107 The goals of this research were to obtain and annotate the first complete mitochondrial
108 genome sequence from the *Mycosphaerella* branch of the fungal evolutionary tree, and to test
109 a previous hypothesis of low mitochondrial diversity within global populations of *M.*
110 *graminicola*. Complete sequences of the mtDNA genomes from two isolates of *M.*
111 *graminicola* (one from North America and one from Europe) plus sequences at three
112 mitochondrial loci for 35 additional isolates representing most of the known global diversity
113 were compared, first to quantify the overall mtDNA sequence diversity in *M. graminicola*
114 and, second, to compare it with earlier findings of low diversity based on RFLP analysis. An
115 interspecific analysis of the tRNA genes flanking *ml* of species in the Pezizomycotina
116 revealed a consensus in tRNA gene content and order.

117

118 **2. Materials and methods**

119 *2.1. Fungal strains, DNA extraction, and library construction*

120 Strain IPO323 was isolated from a naturally infected leaf of the soft white wheat cultivar
121 Arminda collected in Brabant, the Netherlands during 1981 (Kema and Van Silfhout, 1997).
122 Fungal mycelia were produced on liquid shake cultures, harvested, stored and prepared for
123 DNA extraction as described in Kema et al. (2002). Fungal spores and mycelia were ground
124 with a Hybaid Ribolyser (model FP120HY-230) for 10 s at 2500 rpm with two tungsten
125 carbide beads, and total genomic DNA was extracted using the Promega Wizard Magnetic
126 DNA Purification System for Food as described by the manufacturer except with only 50 mg
127 of lyophilised fungal material and 500 µl of lysis buffer. Plasmid libraries with insert sizes of

128 3 and 8 kb were created at the U.S. Department of Energy's Joint Genome Institute (JGI) and
129 sequenced to 4× genomic sequence coverage (~150,000 clones each).

130 Strain STBB1 was isolated from a wheat field 5 km southwest of College Station, Texas,
131 USA, during 1989. The entire mt genome of this isolate was purified from total DNA by
132 cesium chloride (CsCl) ultracentrifugation as described by Garber and Yoder (1983), with a
133 CsCl density of 1.6 g/ml. A library was constructed by digesting the purified mtDNA to
134 completion using the restriction enzyme *Hind*III, ligating the fragments into the plasmid
135 vector pUC18 and cloning in *Escherichia coli* strain DH5α.

136

137 2.2. DNA sequencing and assembly

138 Shotgun sequencing of the nuclear and mt genomes of isolate IPO323 was through the
139 Community Sequencing Program of the JGI (www.jgi.doe.gov/CSP/) by analysis of libraries
140 with insert sizes averaging 3, 8 and 40 kb. The mt genome was assembled from ~7,680
141 sequencing reads from 10 plates of the 3-kb library using phrap (<http://www.phrap.org/>) with
142 its standard parameters. This corresponds to roughly 5-6 Mb of sequence. Approximately
143 5.5% of the reads (~260 kb) represented mtDNA so the initial sequence was assembled at a
144 depth of about 6×. The average depth of coverage for the entire project was 8.9× and was
145 released publicly (<http://genome.jgi-psf.org/Mycgr1/Mycgr1.home.html>) during November
146 2006.

147 The mtDNA library obtained from isolate STBB1 was sequenced using the BigDye™
148 Terminator v3.0 Cycle Sequencing kit and the primer walking strategy. The sequencing
149 reactions were in a total volume of 10 µl using 20-40 ng of plasmid DNA, 10 pmol of
150 primers and 2 µl of BigDye reaction mix, previously diluted 1:4. The cycling profile was 10 s
151 denaturation at 95°C, 5 s annealing at 50°C and 4 min extension at 60°C for 100 cycles. The
152 sequencing reactions were purified through Sephadex G-50 DNA Grade F (Amersham

153 Biosciences, Switzerland) before being loaded into an ABI 3100 automated sequencer
154 (Applied Biosystems). The sequences were aligned and analyzed with the Sequencher
155 version 4.2 software package (Gene Codes Corporation, Ann Arbor, MI) using the genetic
156 code of Pezizomycotina that diverged from the standard nuclear code for the codon TGA,
157 which was read as Trp and not as Stop. Sequencing of the isolate STBB1 mt library generated
158 approximately 75% of the entire mtDNA genome. Gaps in the STBB1 sequence were filled
159 by aligning the sequenced *Hind*III fragments to the complete mtDNA sequence of isolate
160 IPO323 and designing pairs of primers to amplify the missing regions in STBB1. The
161 amplicons were sequenced as described above to obtain the entire mt genome of isolate
162 STBB1.

163

164 2.3. Sequence annotation

165 The mtDNA sequence of *M. graminicola* was screened for similarity with those from
166 other organisms in the NCBI database using the BlastN tool. Sequences showing matches
167 with protein-coding genes of other organisms were subsequently compared using the BlastX
168 tool (Altschul et al., 1990). The mt sequences of strains IPO323 and STBB1 were aligned
169 using the Sequencher program and screened manually for polymorphisms including
170 transitions, transversions, insertions and deletions (indels). The genes coding for ribosomal
171 RNAs were determined by comparison with sequences from other fungi. The tRNAs were
172 defined by tRNAscan-SE v1.21 (Lowe and Eddy, 1997) and by comparison with the NCBI
173 database. Expression of mt genes was tested by blast searches against databases of EST
174 sequences (Goodwin et al., 2007; Kema et al., 2003; Soanes and Talbot, 2006). Repetitive
175 elements, including minisatellites, simple-sequence repeats (SSRs) and mononucleotide
176 repeats were identified using the online program Perfect Microsatellite Repeat Finder
177 (<http://sgdp.iop.kcl.ac.uk/nikammar/repeatfinder.html>).

178

179 *2.4. Intraspecific comparison and haplotype network*

180 Three mtDNA loci, named *Mg1*, *Mg2* and *Mg3*, were used to assess the overall mt
181 diversity within the species. These loci were located in different regions of the mtDNA and
182 were chosen because they displayed different degrees of polymorphism in the comparison
183 between STBB1 and IPO323. *Mg1* was located within *orf1* and had no polymorphism. *Mg2*
184 included a portion of *orf4* and had several polymorphisms including single-nucleotide
185 polymorphisms (SNPs), indels and homopolymers of various lengths. *Mg3* included the
186 region with *tRNA-Gly*, *tRNA-Asp*, *tRNA-Ser* and *tRNA-Trp* and had two polymorphic
187 microsatellites and one homopolymer. Thirty-five isolates (Table 1) belonging to four RFLP
188 haplotypes (Zhan et al., 2003) and originating from five continents were amplified using
189 primers: *Mg1F* (5'-CCG GTC CCT CTA ATA GTT CTG G-3') and *Mg1R* (5'-TAA TCC
190 GCC ATT ACT TCT CAG G-3'); *Mg2F* (5'-GGT TCC AAT GGG TTT AAT GCT A-3')
191 and *Mg2R* (5'- TGG GTG TAG CTA GAA ACC CTT C-3'); *Mg3F* (5'-AAG CTA CGC
192 CTA TGG CTA ACA C-3') and *Mg3R* (5'-AGG TAA GAC GCA CGC ATT TC-3'). Each
193 PCR reaction contained 5-10 ng of DNA in a 20- μ l reaction volume containing 10 pmol of
194 each primer, 100 μ M of each nucleotide, 2 μ l of 10x PCR buffer (1x PCR buffer: 10 mM
195 KCl, 10 mM (NH₄)₂SO₄, 20 mM Tris-HCl, 2 mM MgSO₄, 0.1% Triton X-100 [pH 8.8]) and
196 1 U of *Taq* DNA Polymerase (New England Biolabs). The PCR amplifications were carried
197 out under the following conditions: initial denaturation at 96°C for 2 min, followed by 35
198 cycles of 96°C for 1 min, 56°C for 1 min, and 72°C for 1 min, with a final extension at 72°C
199 for 5 min. The PCR products were sequenced using the primers *Mg1F*, *Mg2F* and *Mg3R*,
200 generating a total of 1339 bp (338 bp for *Mg1*, 333 bp for *Mg2* and 668 bp for *Mg3*).
201 Sequencing reactions were performed as described previously for STBB1. The program
202 SNAP Workbench (Price and Carbone, 2005) was used to collapse the sequences into

203 haplotypes and DnaSP (Rozas et al., 2003) was used to test for recombination within and
204 among the tested loci. The software package TCS version 1.21 (Clement et al., 2000) was
205 used to infer intraspecific evolution of the *M. graminicola* mtDNA. This program applies a
206 statistical parsimony method to infer unrooted cladograms based on Templeton's 95%
207 parsimony connection limit (Templeton et al., 1992).

208

209 **3. Results**

210 *3.1. Gene content and genome organization*

211 The mt genome of *M. graminicola* is a circular molecule of approximately 43,960 bp
212 containing 15 protein-coding genes, the large (*rnl*) and small (*rns*) ribosomal subunits, 27
213 tRNAs and eight putative open reading frames (*orfs1-8*) of unknown function (Fig.1 and
214 Table 2). The protein-coding genes included three ATP synthase subunits (*atp6*, *atp8*, and
215 *atp9*), the three cytochrome oxidase subunits I, II, and III (*cox1-3*), cytochrome b (*cytb*),
216 seven nicotinamide adenine dinucleotide ubiquinone oxidoreductase subunits (*nad1-6*,
217 *nad4L*) and a DNA-directed RNA polymerase (*RNA-Pol*). These genes were transcribed in
218 two contiguous segments of opposite direction (Fig. 1).

219 A putative ribosomal protein (rps5-like) commonly found within *rnl* of ascomycetes was
220 missing (Fig. 1, Table 2). To test whether this gene could have been transferred to the nuclear
221 genome, *blastp* and *tblastn* searches were performed on the 8.9× draft genomic sequence of
222 *M. graminicola*. The *blastp* searches identified no matching proteins among the list of
223 annotated genes. However, the *tblastn* searches identified matches at better than e^{-5} on
224 scaffold 5 to rps5-like proteins from *Phaeosphaeria nodorum* (e^{-9}) and *Penicillium marneffeii*
225 (e^{-7}), but not to those from the Sordariomycetes *Hypocrea jecorina* or *Verticillium dahliae*.
226 Therefore, this gene most likely occurs in the nuclear rather than the mitochondrial genome
227 of *M. graminicola*.

228 The eight putative *orfs* of unknown function are predicted to produce proteins containing
229 from 126 to 481 amino acids. The 3' terminus of *orf3* overlapped with *orf2* for 52
230 nucleotides. The exact function of these putative proteins remains to be determined, but the
231 TMHMM2 method (Krogh et al., 2001) predicted *orf2* to encode a non-membrane protein,
232 whereas the other *orfs* were predicted to encode proteins having from one (*orf5* and *orf8*) to
233 ten (*orf7*) transmembrane domains. Expressed sequence tag (EST) databases (Goodwin et al.,
234 2007; Kema et al., 2003; Soanes and Talbot, 2006) provided evidence for the transcription of
235 *orf5*, *orf6* and *orf8*.

236 Putative protein-coding genes covered 51.8% of the genome (including 15.9% composed
237 of putative *orfs*), while 4.5 and 11.5% corresponded to tRNA genes and both *rnl* and *rns*,
238 respectively. These values were similar to those reported for other ascomycetes (Table 2).
239 Overall, the *M. graminicola* mtDNA was 34.5% A, 33.5% T, 16.3% G and 15.7% C. MtDNA
240 AT-content was 68% with coding and non-coding parts of the genome having, on average,
241 the same AT-percentage.

242

243 3.2. Codon usage and tRNA genes

244 As expected given the transmembrane location of most mt proteins, the three most
245 frequent codons were TTA (377 counts), ATA (371 counts) and TTT (270 counts) encoding
246 Leu, Ile and Phe, respectively. These amino acids have hydrophobic side chains commonly
247 found in transmembrane helices. These three codons accounted for 19.3% of all codons in the
248 mt genome. One codon was not used at all (CGA, Arg) and eight codons (CGC, TGG, TGC,
249 CGG, CTC, GGC, GTC, CCG) were under represented, being used from one to ten times
250 each. All 15 protein-coding genes started with the canonical translation initiation codon
251 ATG. The preferred stop codon was TAA, present in 11 protein-coding genes; the alternative
252 stop codon was TAG. Codon usage of the *orfs* was similar to that of the protein-coding loci.

253 The 27 tRNAs encoded by the mt genome of *M. graminicola* could carry all 20 amino
254 acids (Fig. 1). Two tRNA isoacceptors were identified for serine and leucine, three for
255 arginine and four for methionine. Among the 27 tRNAs, only *tRNA-Val* occurred singly. The
256 remaining 26 tRNA genes were grouped into five clusters, composed of 12, 5, 4, 3 and 2
257 tRNA genes (Fig. 1). As in other filamentous fungi, several tRNA genes flanked the *rnl* gene
258 (Table 3, Tambor et al., 2006). In *M. graminicola*, these tRNA genes had an order similar to
259 that of Eurotiomycetes and generally followed a conserved pattern found in other fungi
260 (Table 3, Ghikas et al., 2006). Surprisingly, *M. graminicola* did not possess the TEM-tRNA
261 genes at the beginning of the 3' tRNA gene consensus, in contrast to both Eurotiomycetes
262 and Sordariomycetes, suggesting an independent rearrangement in this species. The
263 secondary structures of *tRNA-Phe* and *tRNA-Thr* diverged from the expected cloverleaf form
264 as they contained nine instead of the canonical seven nucleotides in the anticodon loop.

265

266 3.3 Repetitive elements and comparative genomics

267 One 27-mer minisatellite repeated three times (located between the *nad4* and *nad4L*
268 genes), 186 SSRs and 51 mononucleotide repeats larger than seven nucleotides (mainly
269 located in non-coding regions), were found in the mt genome of *M. graminicola*.

270 The total nucleotide diversity between IPO323 and STBB1 was 0.16%. The two *M.*
271 *graminicola* isolates differed by only 23 base substitutions, including fourteen transversions
272 and nine transitions. These changes represented 0.05% of the entire mt genome. Twenty-two
273 additional mutations were found between IPO323 and STBB1: 18 were mononucleotide
274 repeats of different lengths (9 poly-A and 9 poly-T), two were tetra- (AAAT) or penta-
275 nucleotide microsatellite repeats (ATTTA), one was a frameshift mutation and the last was a
276 17-base deletion (Fig. 2). The nucleotide diversity among the global sample of 35 isolates
277 was 35% greater than that between only IPO323 and STBB1 for the same three

278 mitochondrial loci (*Mg1*, *Mg2* and *Mg3*). *Mg2* was the most variable locus, having 3 SNPs, a
279 17-bp indel, a polymorphic microsatellite with 2 alleles, and a mononucleotide repeat with 3
280 alleles. *Mg1* had the fewest mutations, with two SNPs that were exclusive to isolates
281 collected from durum wheat (*Triticum turgidum*). *Mg3* had a mononucleotide repeat with 2
282 alleles and 2 microsatellites, respectively with 2 and 4 alleles. All microsatellite alleles were
283 due to differences in the number of repeats. The concatenated sequences of *Mg1*, *Mg2*, and
284 *Mg3* from all 37 isolates identified 14 haplotypes. If all mutations other than SNPs were
285 excluded from the analysis, only three haplotypes were found (Table 1). If the increase of
286 35% in nucleotide diversity detected for the *Mg* loci is extrapolated to the total genome, it
287 results in a value of 0.22% for mitochondrial nucleotide diversity in a global sample of 37
288 isolates representing most of the known mt variants.

289 A haplotype network was inferred from all three *Mg* loci using the concatenated
290 alignments (Fig. 3). The haplotype network did not show a clear pattern of geographical
291 association, although isolates from North America were in the top half and all of those at the
292 bottom were from Europe. Some frequent haplotypes such as H5, H6, and H7 included
293 isolates of mixed origin, while others (H1, H11 and H13) were geographically limited. The
294 *M. graminicola* haplotypes originating from durum wheat (*Triticum turgidum* ssp. *durum*)
295 were distinguished from those originating from bread wheat (*T. aestivum*) by three SNPs. The
296 two sequenced haplotypes (H5 and H9) represented different parts of the network. No
297 evidence for recombination was found in the mtDNA of *M. graminicola* using the DnaSP
298 program.

299

300 **4. Discussion**

301 The mt genomes of two strains of the plant pathogenic fungus *M. graminicola* originating
302 from different continents (Europe and North America) were sequenced completely, annotated

303 and compared to identify polymorphisms. Both isolates had mt genomes belonging to RFLP
304 haplotype 2 (Zhan et al., 2003; Table 1). The mtDNA of *M. graminicola* was circular and
305 A+T biased like those of most other fungi (Table 2).

306 These two *M. graminicola* sequences represent the first complete mt genomes of any
307 species in the genus *Mycosphaerella* or from the branch of the fungal evolutionary tree that
308 includes the Capnodiales, Dothideales, or Myriangiales (Schoch et al. 2006). *Mycosphaerella*
309 and its related asexual genera (e.g., *Cercospora*, *Septoria*) comprise one of the largest and
310 most economically important groups of pathogenic fungi (Goodwin et al., 2001) with several
311 thousand species infecting virtually every major family of plants (Corlett, 1991). Species of
312 *Mycosphaerella* are not closely related to model fungi or those with completely sequenced mt
313 genomes, so represent a previously unsampled branch of the fungal evolutionary tree.

314 The mtDNA of *M. graminicola* contains genes for 14 inner mt membrane proteins
315 involved in electron transport and coupled oxidative phosphorylation, as well as *rnl*, *rns* and
316 *RNA-Pol* genes (Fig. 1). Except for presence of the *RNA-Pol* gene and absence of a gene
317 encoding a putative ribosomal protein (*rps5*-like), this is the standard set of mtDNA-encoded
318 genes found in other fungi. The *rps5*-like gene is found commonly in mt genomes of different
319 fungal species and it was postulated that mtDNA-encoded *rps5* was present in the common
320 ancestor of fungal and animal mtDNAs (Bullerwell et al., 2000). As *M. graminicola* is one of
321 the few ascomycetes known to be lacking *rps5*, the absence of this gene could indicate an
322 independent loss in this species. A possible homolog of this gene was identified on scaffold 5
323 of the 8.9× draft genomic sequence of *M. graminicola*, so it may have been transferred to the
324 nuclear genome rather than having been lost.

325 Genes in the *M. graminicola* mtDNA had no introns, a finding that contrasts with other
326 fungal mtDNAs that possess large introns containing intron-encoded proteins, as found in
327 *Podospira anserina* (Cummings et al., 1990) and *Penicillium marneffeii* (Woo et al., 2003).

328 Eight *orfs*, with no obvious homology to any other sequenced genes present in the GenBank
329 database, were found in the mt genome of *M. graminicola*. The functions of these putative
330 genes remain unclear, although some of them may represent highly diverged versions of
331 known mtDNA-encoded genes, no longer recognizable by identity searches (Gray et al.,
332 1998). EST databases provided evidence for transcription of *orf5*, *orf6*, and *orf8*, indicating
333 that they may be expressed. Interestingly, these three *orfs* were the only ones of the eight that
334 were located adjacent to tRNA genes, so possibly they may be transcribed along with the
335 tRNAs but not translated.

336 All tRNA secondary structures had the expected cloverleaf form, but particularly
337 interesting were *tRNA-Thr* (UGU as anticodon) and *tRNA-Phe* (GAA as anticodon) because
338 they had nine nucleotides in the anticodon loop instead of the canonical seven. This rare
339 tRNA structure was described previously in *Metarhizium anisopliae* for *tRNA-Thr* and *tRNA-*
340 *Glu* (Ghikas et al., 2006), and in *Verticillium dahliae* for *tRNA-Thr*, *tRNA-Glu*, *tRNA-Arg* and
341 *tRNA-Ser* (Pantou et al., 2006).

342 Nuclear genomes, including that of *M. graminicola* (Goodwin et al., 2007), possess SSRs
343 that are known to be highly variable in terms of motif repeat number and distribution (Katti et
344 al., 2001; Toth et al., 2000). This study presents a similar picture for the mt genome of *M.*
345 *graminicola*. SSRs and mononucleotide repeats may play a significant role in the regulation
346 and evolution of the entire molecule. In nuclear genomes it was demonstrated that these
347 highly variable tracts, if placed in promoter regions, could influence transcriptional activity
348 (Kashi et al., 1997) and could play an important role in creating and maintaining quantitative
349 genetic variation (Kashi et al., 1997; Tautz et al., 1986). In the mtDNA of *M. graminicola*,
350 mononucleotide repeats became less common in coding regions as their length increased.
351 Because most long mononucleotide repeats are located 5'-upstream of ATG start codons
352 (Fig. 2), we hypothesize that they might play a role in regulating transcription. These tracts

353 could be protein binding signals and, more precisely, upstream promoter elements, as
354 demonstrated previously in nuclear genomes (Kashi et al., 1997).

355 The intraspecific mt diversity was first assessed by comparing the total genome
356 sequences of two isolates (STBB1 and IPO323), giving a nucleotide diversity of 0.16%. In
357 order to assess species-wide variation, another 35 isolates were chosen, originating from five
358 continents and belonging to four of the seven known RFLP haplotypes (Table 1). Using these
359 additional isolates, the total mtDNA variation in *M. graminicola* was estimated to range from
360 0.16 to 0.22%, falling within the lower range of published intraspecific nucleotide diversities.
361 The nucleotide diversity would decrease to 0.12% if the 17-bp indel was excluded. This 17-
362 bp indel appears to be a recent mutation that occurred during the 1970s (Torriani SFF,
363 unpublished), suggesting that the *M. graminicola* mtDNA may be increasing in diversity
364 following the hypothesized selective sweep (Zhan et al., 2003). Other examples of low
365 intraspecific mtDNA nucleotide diversity based on complete mtDNA sequences were 0.2%
366 for the olive fly *Bactrocera oleae* (Nardi et al., 2003) and 0.36% for *Drosophila simulans*
367 (Ballard, 2000).

368 These results support earlier findings of low mt diversity in *M. graminicola* obtained by
369 RFLP analysis (Zhan et al., 2003). While the haplotypic diversity based on sequences was
370 higher than that found using RFLPs, the total nucleotide diversity remains the lowest reported
371 to date in fungi. The greater number of haplotypes found through sequencing reflects the
372 higher resolution of this method, especially the ability to resolve small indels that are missed
373 by RFLP analysis (Fig. 3). In fact, if indels were removed from the sequence analysis and
374 only SNPs were considered, only three mt haplotypes were found, but they did not always
375 correspond with the RFLP data (Table 1). For example, isolates with RFLP haplotypes 1 and
376 2 were the most polymorphic and could have SNP haplotypes 1 or 3. Isolates with RFLP
377 haplotype 3 always had SNP haplotype 1. It was interesting that isolates of *M. graminicola*

378 adapted to durum wheat had unique RFLP and SNP haplotypes 4 and 2, respectively (Table
379 1). The nonrandom association between mitochondrial RFLP haplotypes and host species,
380 presumably caused by natural selection operating on the mt genome, was noted previously in
381 *M. graminicola* (Zhan et al., 2004) and other fungi (Demanche et al., 2001; Gomes et al.,
382 2000). The intraspecific haplotype network (Fig. 3) that included all mutational events also
383 distinguished between haplotypes originating from bread wheat and durum wheat.

384 The contrasting genetic diversity among mt and nuclear genomes in *M. graminicola*
385 (Zhan et al., 2003, 2004) raises intriguing questions about the mechanisms leading to this
386 phenomenon. At least two hypotheses can be proposed to account for the observed low levels
387 of mt variation, including a lower mutation rate in the mt genome or a selective sweep. A
388 comparison among the three yeast species *Saccharomyces cerevisiae*, *Kluyveromyces lactis*,
389 and *Candida glabrata* showed that the frequency of nucleotide changes is higher in nuclear
390 than in mt genomes (Clark-Walker, 1991), which is the opposite of mammals where nuclear
391 genes evolve slower than mt genes (Saccone et al., 2000). On the other hand, the low level of
392 polymorphism in the *M. graminicola* mtDNAs may have been generated through fixation of
393 an advantageous mt mutation during a selective sweep. The selection of a favored mt
394 haplotype leading to low levels of polymorphism was suggested in the oomycete
395 *Phytophthora infestans* (Gavino and Fry, 2002).

396 The analysis of intraspecific diversity in the largely conserved mt genome of *M.*
397 *graminicola* provides the basis for developing new tools essential to clarify the conflicting
398 patterns of nuclear and mt diversity and to understand its cause. For example, the
399 polymorphic microsatellites differed in numbers of 4- or 5-base repeats, making them
400 amenable to agarose gel assays. These polymorphisms have already been used to analyze
401 paternity in crosses (Ware, 2006) and are now being used to analyze the evolution of
402 resistance to strobilurin fungicides in *M. graminicola* (Torriani and McDonald, unpublished).

403

404 **Acknowledgments**

405 We gratefully acknowledge Marcello Zala for technical support and Charles Crane for
406 bioinformatics analyses. DNA sequencing of isolate IPO323 was performed at the
407 Department of Energy's Joint Genome Institute through the Community Sequencing Program
408 (www.jgi.doe.gov/csp/). All sequencing data are public and may be accessed through
409 <http://www.jgi.doe.gov/Mgraminicola>. This project was supported by the Swiss National
410 Science Foundation (grant 3100A0-104145) and by USDA CRIS project 3602-22000-013-
411 00D.

412

413 **References**

- 414 Adams, K.L., Daley, D.O., Qiu, Y.L., Whelan, J., Palmer, J.D., 2000. Repeated, recent and
415 diverse transfer of a mitochondrial gene to the nucleus in flowering plants. *Nature* 408,
416 354-357.
- 417 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment
418 search tool. *J. Mol. Biol.* 215, 403-410.
- 419 Ballard, J.W.O., 2000. Comparative genomics of mitochondrial DNA in *Drosophila*
420 *simulans*. *J. Mol. Evol.* 51, 64-75.
- 421 Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria.
422 *Mol. Ecol.* 13, 729-744.
- 423 Bannon, F.J., Cooke, B.M., 1998. Studies on dispersal of *Septoria tritici* pycnidiospores in
424 wheat-clover intercrops. *Plant. Pathol.* 47, 49-56.
- 425 Boeger, J.M., Chen, R.S., McDonald, B.A., 1993. Gene flow between geographic populations
426 of *Mycosphaerella graminicola* (anamorph *Septoria tritici*) detected with restriction
427 fragment length polymorphism markers. *Phytopathology* 83, 1148-1154.

- 428 Brennicke, A., Grohmann, L., Hiesel, R., Knoop, V., Schuster, W., 1993. The mitochondrial
429 genome on its way to the nucleus: different stages of gene transfer in higher plants. *FEBS*
430 *Lett.* 325, 140-145.
- 431 Bullerwell, C.E., Burger, G., Lang, F., 2000. A novel motif for identifying Rps3 homologs in
432 fungal mitochondrial genomes. *Trends Biochem. Sci.* 25, 363-365.
- 433 Burger, G., Gray, M.W., Lang, B.F., 2003. Mitochondrial genomes: anything goes. *TRENDS*
434 *Genet.* 19, 709-716.
- 435 Campbell, A., Mrazek, J., Karlin, S., 1999. Genome signature comparisons among
436 prokaryote, plasmid, and mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 96, 9184-9189.
- 437 Chen, R.S., McDonald, B.A., 1996. Sexual reproduction plays a major role in the genetic
438 structure of populations of the fungus *Mycosphaerella graminicola*. *Genetics* 142, 1119-
439 1127.
- 440 Chen, J.Z., Hebert, P.D.N., 1999. Intra-individual sequence diversity and hierarchical
441 approach to the study of mitochondrial DNA mutations. *Mutat. Res.* 434, 205-217.
- 442 Clark-Walker, G.D., 1991. Contrasting mutation rates in mitochondrial and nuclear genes of
443 yeasts versus mammals. *Curr. Genet.* 20, 195-198.
- 444 Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene
445 genealogies. *Mol. Ecol.* 9, 1657-1659.
- 446 Corlett, M., 1991. An annotated list of the published names in *Mycosphaerella* and
447 *Sphaerella*. *Mycol. Mem.* 18, 1-328.
- 448 Cummings, D.J., McNally, K.L., Domenico, J.M., Matsuura, E.T., 1990. The complete DNA
449 sequence of the mitochondrial genome of *Podospora anserina*. *Curr. Genet.* 17, 375-402.
- 450 Demanche, C., Berthelemy, M., Petit, T., Polack, B., Wakefield, A.E., Dei-Cas, E., Guillot,
451 J., 2001. Phylogeny of *Pneumocystis carinii* from 18 primate species confirms host
452 specificity and suggests coevolution. *J. Clin. Microbiol.* 39, 2126-2133.

- 453 Eyal, Z., 1999. The Septoria/Stagonospora blotch disease of wheat: past, present and future.
454 In: van Ginkel, M., McNab, A., Krupinsky, J. (Eds.), Septoria and Stagonospora Diseases
455 of Cereals: A Compilation Research, CIMMYT, Mexico. Eur. J. Plant. Pathol. 105, 629-
456 641.
- 457 Garber, R.C., Yoder, O.C., 1983. Isolation of DNA from filamentous fungi and separation
458 into nuclear, mitochondrial, ribosomal, and plasmid components. Anal. Biochem. 135,
459 416-422.
- 460 Gavino, P.D., Fry, W.E., 2002. Diversity in and evidence for selection on the mitochondrial
461 genome of *Phytophthora infestans*. Mycologia 94, 781-793.
- 462 Ghikas, D.V., Kouvelis, V.N., Typas, M.A., 2006. The complete mitochondrial genome of
463 the entomopathogenic fungus *Metarhizium anisopliae* var. *anisopliae*: gene order and trn
464 gene clusters reveal a common evolutionary course for all Sordariomycetes, while
465 intergenic regions show variation. Arch. Microbiol. 185, 393-401.
- 466 Gomes, E.A., de Abreu, L.M., Borges, A.C., de Araujo, E.F., 2000. ITS sequences and
467 mitochondrial DNA polymorphism in *Pisolithus* isolates. Mycol. Res. 104, 911-918.
- 468 Goodwin, S.B., Dunkle, L.D., Zismann, V.L., 2001. Phylogenetic analysis of *Cercospora* and
469 *Mycosphaerella* based on the internal transcribed spacer region of ribosomal DNA.
470 Phytopathology 91, 648-658.
- 471 Goodwin, S.B., van der Lee, T.A.J., Cavaletto, J.R., te Lintel-Hekkert, B., Crane, C.F.,
472 Kema, G.H.J., 2007. Identification and genetic mapping of highly polymorphic
473 microsatellite loci from an EST database of the septoria tritici blotch pathogen
474 *Mycosphaerella graminicola*. Fungal. Genet. Biol. 44, 398-414.
- 475 Goodwin, S.B., Waalwijk, C., Kema, G.H.J., 2004. Genetics and genomics of
476 *Mycosphaerella graminicola*: a model for the Dothideales. In: Arora DK, Khachatourians

- 477 GG (Eds.), Applied Mycology & Biotechnology. Volume 4. Fungal Genomics Elsevier
478 Science B.V., Amsterdam, pp. 315-330.
- 479 Gray, M.W., Burger, G., Lang, B.F., 1999. Mitochondrial evolution. Science 283, 1476-1481.
- 480 Gray, M.W., Lang, B.F., Cedergren, R., Golding, B., Lemieux, C., Sankoff, D., Turmel, M.,
481 Brossard, N., Delage, E., Littlejohn, T.G., Plante, I., Rioux, P., Saint-Louis, D., Zhy, Y.,
482 Burger, G., 1998. Genome structure and gene content in protist mitochondrial DNAs.
483 Nucleic Acids Res. 26, 865-887.
- 484 Hane, J., Lowe, R., Solomon, P., Tan, K.C., Schoch, C.L., Spatafora, J.W., Crous, P.W.,
485 Kodira, C., Birren, B., Torriani, S.S.F., McDonald, B.A., Oliver, R.P. 2007.
486 Dothideomycete-plant interaction illuminated by genome sequencing and EST analysis of
487 the wheat pathogen *Stagonospora nodorum*. The Plant Cell, published on-line November
488 16, 2007; 10.1105/tpc.107.052829.
- 489 Hartl, F.U., Pfanner, N., Nicholson, D.W., Neupert, W. 1989. Mitochondrial protein import.
490 Biochim. Biophys. Acta 988, 1-45.
- 491 Hunter, T., Coker, R.R., Royle, D.J. 1999. The teleomorph stage, *Mycosphaerella*
492 *graminicola*, in epidemics of septoria tritici blotch on winter wheat in the UK. Plant
493 Pathol. 48, 51-57.
- 494 John, P., Whatley, F.R., 1975. *Paracoccus denitrificans* and the evolutionary origin of
495 mitochondria. Nature 254, 495-498.
- 496 Kashi, Y., King, D., Soller, M., 1997. Simple sequence repeats as a source of quantitative
497 genetic variation. Trends Genet. 13, 74-78.
- 498 Katti, M.V., Ranjekar, P.K., Gupta, V.S., 2001. Differential distribution of simple sequence
499 repeats in eukaryotic genome sequences. Mol. Biol. Evol. 18, 1161-1167.
- 500 Kema, G.H.J., Goodwin, S.B., Hamza, S., Verstappen, E.C.P., Cavaletto, J.R., Van der Lee,
501 T.A.J., Hagenaar-de Weerd, M., Bonants, P.J.M., Waalwijk, C., 2002. A combined

- 502 AFLP and RAPD genetic linkage map of *Mycosphaerella graminicola*, the septoria tritici
503 leaf blotch pathogen of wheat. Genetics 161, 1497-1505.
- 504 Kema, G.H.J., van Silfhout, C.H. 1997. Genetic variation for the virulence and resistance in
505 the wheat *Mycosphaerella graminicola* pathosystem .3. Comparative seedling and adult
506 plant experiments. Phytopathology 87, 266-272.
- 507 Kema, G.H.J., Verstappen, E.C.P., Todorova, M., Waalwijk, C., 1996. Successful crosses and
508 molecular tetrad and progeny analyses demonstrate heterothallism in *Mycosphaerella*
509 *graminicola*. Curr. Genet. 30, 251-258.
- 510 Kema, G.H.J., Verstappen, E., van der Lee, T., Mendes, O., Sandbrink, H., Klein-Lankhorst,
511 R., Zwiers, L., Csukai, M., Baker, K., Waalwijk, C., 2003. Gene hunting in
512 *Mycosphaerella graminicola*. Proceedings of the 22nd Fungal Genetics Conference,
513 Asilomar, California, USA Page 252.
- 514 Kirk, P.M., Cannon, P.F., David, J.C., Stalpers, J.A., 2001. Ainsworth & Bisby's Dictionary
515 of the Fungi. 9th Ed., CAB International, Wallingford, UK.
- 516 Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L.L., 2001. Predicting
517 transmembrane protein topology with a hidden Markov model: Application to complete
518 genomes. J. Mol. Biol. 305, 567-580.
- 519 Kudla, J., Albertazzi, F.J., Blazevic, D., Hermann, M., Bock, R., 2002. Loss of the
520 mitochondrial cox2 intron 1 in a family of monocotyledonous plants and utilization of the
521 mitochondrial intron sequence for the construction of a nuclear intron. Mol. Genet.
522 Genomics 267, 223-230.
- 523 Kurdyla, T.M., Guthrie, P.A.I., McDonald, B.A., Appel, D.N. 1995. RFLPs in
524 mitochondrial and nuclear DNA indicate low levels of genetic diversity in the oak wilt
525 pathogen *Ceratocystis fagacearum*. Curr. Genet. 27, 373-378.

- 526 Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: A program for improved detection of transfer
527 RNA genes in genomic sequence. *Nucleic. Acids Res.* 25, 955-964.
- 528 Liu, Y.C., Cortesi, P., Double, M.L., MacDonald, W.L., Milgroom, M.G., 1996. Diversity
529 and multilocus genetic structure in populations of *Cryphonectria parasitica*.
530 *Phytopathology* 86, 1344-1351.
- 531 Nardi, F., Carapelli, A., Dallai, R., Frati, F., 2003. The mitochondrial genome of the olive fly
532 *Bactrocera oleae*: two haplotypes from distant geographical locations. *Insect Mol. Biol.*
533 12, 605-611.
- 534 Pantou, M.P., Kouvelis, V.N., Typas, M.A., 2006. The complete mitochondrial genome of
535 the vascular wilt fungus *Verticillium dahliae*: a novel gene order for *Verticillium* and a
536 diagnostic tool for species identification. *Curr. Genet.* 50, 125-136.
- 537 Price, E.W., Carbone, I., 2005. SNAP: workbench management tool for evolutionary
538 population genetic analysis. *Bioinformatics* 21, 402-404.
- 539 Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA
540 polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496-
541 2497.
- 542 Saccone, C., Gissi, C., Lanave, C., Larizza, A., Pesole, G., Reyes, A., 2000. Evolution of the
543 mitochondrial genetic system: an overview. *Gene* 261, 153-159.
- 544 Sanderson, F.R., 1972. A *Mycosphaerella* species as the ascogenous state of *Septoria tritici*
545 Rob. and Desm. *New Zeal. J. Bot.* 10, 707-710.
- 546 Schoch, C.L., Shoemaker, R.A., Seifert, K.A., Hambleton, S., Spatafora, J.W., Crous, P.W.,
547 2006. A multigene phylogeny of the Dothideomycetes using four nuclear loci. *Mycologia*
548 98, 1041-1052.
- 549 Soanes, D.M., Talbot, N.J., 2006. Comparative genomic analysis of phytopathogenic fungi
550 using expressed sequence tag (EST) collections. *Mol. Plant Pathol.* 7, 61-70.

- 551 Sommerhalder, R.J., McDonald, B.A., Zhan, J., 2007. Concordant evolution of mitochondrial
552 and nuclear genomes in the wheat pathogen *Phaeosphaeria nodorum*. Fungal Genet. Biol.
553 44, 764-772.
- 554 Tambor, J.H.M., Guedes, R.F., Nobrega, M.P., Nobrega, F.G., 2006. The complete DNA
555 sequence of the mitochondrial genome of the dermatophyte fungus *Epidermophyton*
556 *floccosum*. Curr. Genet. 49, 302-308.
- 557 Tautz, D., Trick, M., Dover, G., 1986. Cryptic simplicity in DNA is a major source of genetic
558 variation. Nature 322, 652-656.
- 559 Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of the phenotypic
560 associations with haplotypes inferred from restriction endonuclease mapping and DNA
561 sequence data. III. Cladogram estimation. Genetics 132, 619-633.
- 562 Toth, G., Gaspari, Z., Jurka, J., 2000. Microsatellites in different eukaryotic genomes: survey
563 and analysis. Genome Res. 10, 967-981.
- 564 Ware, S.B., 2006. Aspects of sexual reproduction in *Mycosphaerella* species on wheat and
565 barley: genetic studies on specificity, mapping, and fungicide resistance. Ph.D. thesis,
566 Wageningen University, the Netherlands
- 567 Woo, P.C.Y., Zhen, H.J., Cai, J.J., Yu, J., Lau, S.K.P., Wang, J., Teng, J.L.L., Wong, S.S.Y.,
568 Tse, R.H., Chen, R., Yang, H.M., Liu, B., Yuen, K.Y., 2003. The mitochondrial genome
569 of the thermal dimorphic fungus *Penicillium marneffeii* is more closely related to those of
570 molds than yeasts. FEBS LETT. 555, 469-477.
- 571 Xia, J.Q., Correll, J.C., Lee, F.N., Ross, W.J., Rhoads, D.D., 2000. Regional population
572 diversity of *Pyricularia grisea* in Arkansas and influence of host selection. Plant Dis. 84,
573 877-884.
- 574 Zhan, J., Kema, G.H.J., McDonald, B.A., 2004. Evidence for natural selection in the
575 mitochondrial genome of *Mycosphaerella graminicola*. Phytopathology 94, 261-267.

- 576 Zhan, J., McDonald, B.A., 2004. The interaction among evolutionary forces in the pathogenic
577 fungus *Mycosphaerella graminicola*. Fungal Genet. Biol. 41, 590-599.
- 578 Zhan, J., Pettway, R.E., McDonald, B.A., 2003. The global genetic structure of the wheat
579 pathogen *Mycosphaerella graminicola* is characterized by high nuclear diversity, low
580 mitochondrial diversity, regular recombination and gene flow. Fungal Genet. Biol. 38,
581 286-297.

582 Table 1

583 *Mycosphaerella graminicola* isolates included in the analysis of mtDNA variation.

Isolate	Haplotype			Host	Year	Location	Source
	RFLP ^a	Sequence ^b	SNP ^c				
AU49	1	1	1	bread wheat	1993	Australia	B. Ballantyne
AU54	1	1	1	bread wheat	1993	Australia	B. Ballantyne
AU58	1	1	1	bread wheat	1993	Australia	B. Ballantyne
CH9B12A	1	4	1	bread wheat	1999	Switzerland	B.A. McDonald
AU59	1	11	1	bread wheat	1993	Australia	B. Ballantyne
OR402	1	12	1	bread wheat	1990	Oregon	J. Boeger, B.A. McDonald, M. Schmitt
CH9B12C	1	7	3	bread wheat	1999	Switzerland	B.A. McDonald
IN11	1	7	3	bread wheat	1993	Indiana	G. Shaner
IN12	1	7	3	bread wheat	1993	Indiana	G. Shaner
IN13	1	7	3	bread wheat	1993	Indiana	G. Shaner
OR389	1	7	3	bread wheat	1990	Oregon	J. Boeger, B.A. McDonald, M. Schmitt
OR409	1	7	3	bread wheat	1990	Oregon	J. Boeger, B.A. McDonald, M. Schmitt
CH9C1A	1	8	3	bread wheat	1999	Switzerland	B.A. McDonald
IN1	1	10	3	bread wheat	1993	Indiana	G. Shaner
IN9	1	10	3	bread wheat	1993	Indiana	G. Shaner
CH9B5A	2	5	1	bread wheat	1999	Switzerland	B.A. McDonald
CH9B9A	2	5	1	bread wheat	1999	Switzerland	B.A. McDonald
GEA2a.2	2	5	1	bread wheat	1992	Germany	R. Huang, G Koch
IPO323	2	5	1	bread wheat	1981	Netherlands	G.H.J. Kema
CH9B7B	2	6	1	bread wheat	1999	Switzerland	B.A. McDonald
GEE2b.2	2	6	1	bread wheat	1992	Germany	R. Huang, G Koch
GEE3a.2	2	6	1	bread wheat	1992	Germany	R. Huang, G Koch
U17	2	6	1	bread wheat	1981	Netherlands	G.H.J. Kema
GEE1a.2	2	14	1	bread wheat	1992	Germany	R. Huang, G Koch
OR428	2	7	3	bread wheat	1990	Oregon	J. Boeger, B.A. McDonald, M. Schmitt
STBB1	2	9	3	bread wheat	1989	Texas	B.A. McDonald
AU57	3	11	1	bread wheat	1993	Australia	B. Ballantyne
AU70	3	11	1	bread wheat	1993	Australia	B. Ballantyne
AU72	3	11	1	bread wheat	1993	Australia	B. Ballantyne
MX156	3	13	1	bread wheat	1993	Mexico	L. Gilchrist
MX160	3	13	1	bread wheat	1993	Mexico	L. Gilchrist
MX163	3	13	1	bread wheat	1993	Mexico	L. Gilchrist
MX167	3	13	1	bread wheat	1993	Mexico	L. Gilchrist
MX169	3	13	1	bread wheat	1993	Mexico	L. Gilchrist
U2	4	2	2	durum wheat	1991	Syria	G.H.J. Kema
U7	4	2	2	durum wheat	1991	Tunisia	G.H.J. Kema
U6	4	3	2	durum wheat	1991	Tunisia	G.H.J. Kema

584 ^aRFLP haplotypes are groups of isolates having identical RFLP patterns following Zhan et al. (2003)

585 ^bSequence haplotypes are groups of isolates having identical concatenated sequences for mitochondrial loci *Mg1*,

586 *Mg2* and *Mg3*

587 °SNP haplotypes are groups of isolates having the identical concatenated sequence for *Mg1*, *Mg2* and *Mg3* after
588 removing all indels

589 Table 2

590 A comparison of the principal features of some completely sequenced fungal mt genomes^a

Species	Size (kb)	A + T content	Coding genes ^b	Orfs	Percent coding ^c	RNAs ^d	Accession number
<i>Aspergillus niger</i>	31.1	74%	14	2	47%*	27	<u>DQ207726</u>
<i>Aspergillus tubingensis</i>	33.6	74%	14	2	43%*	27	<u>DQ217399</u>
<i>Penicillium marneffeii</i>	35.5	76%	15	10	63%*	30	<u>AY347307</u>
<i>Epidermophyton floccosum</i>	30.9	77%	15	5	67%*	27	<u>AY916130</u>
<i>Mycosphaerella graminicola</i>	43.9	68%	15	8	52%*	29	<u>EU090238</u>
<i>Lecanicillium muscarium</i>	24.5	73%	15	0	58%	27	<u>AF487277</u>
<i>Verticillium dahliae</i>	27.2	73%	15	0	53%	27	<u>DQ351941</u>
<i>Fusarium oxysporum</i>	34.5	69%	15	1	44%	27	<u>AY945289</u>
<i>Metarhizium anisopliae</i>	24.7	72%	15	0	59%	26	<u>AY884128</u>

591 ^a All fungi in this list have mt genomes with circular topologies

592 ^b If present the fifteenth gene is *rps5* except for *M. graminicola* that has a *RNA-pol*

593 ^c Asterisks mark the genomes in which *orfs* were considered as coding genes in the calculation of percent

594 coding

595 ^d All fungi in this list have two genes encoding for the large and small ribosomal subunit

596 Table 3

597 Comparison of tRNA^a gene clusters flanking the *rnl* gene in several ascomycetes^b

Species	Class	5'-upstream region ^c	<i>rnl</i>	3'-downstream region ^c	Accession number
<i>A. niger</i>	Eurotiomycetes	KGDS ¹ WIS ² P	<i>rnl</i>	TEVM ¹ M ² L ¹ AFL ² QM ³ H	<u>DQ207726</u>
<i>A. tubingensis</i>	Eurotiomycetes	KGDS ¹ WIS ² P	<i>rnl</i>	TEVM ¹ M ² L ¹ AF ² QLM ³ H	<u>DQ217399</u>
<i>P. marneffei</i>	Eurotiomycetes	RKG ₁ G ² DS ¹ WIS ² P	<i>rnl</i>	TEVM ¹ M ² L ¹ AFL ² QM ³ H	<u>AY347307</u>
<i>E. floccosum</i>	Eurotiomycetes	KGDS ¹ IWS ² P	<i>rnl</i>	TEVM ¹ M ² L ¹ AFL ² QM ³ H	<u>AY916130</u>
<i>M. graminicola</i>	Dothideomycetes	GDS ¹ WIS ² P	<i>rnl</i>	M ¹ L ¹ EAF ² L ² YQM ² HRM ³	<u>EU090238</u>
<i>L. muscarium</i>	Sordariomycetes	GVISW*P	<i>rnl</i>	TE ¹ M ¹ M ² L ¹ E ² FKL ² QHM ³	<u>AF487277</u>
<i>V. dahliae</i>	Sordariomycetes	KGDS*VW*R*P ¹ *P ²	<i>rnl</i>	TE ¹ M ¹ M ² L ¹ AFL ² QHM ³	<u>DQ351941</u>
<i>F. oxysporum</i>	Sordariomycetes	VISWP	<i>rnl</i>	TEM ¹ M ² L ¹ AFKL ² QHM ³	<u>AY945289</u>
<i>M. anisopliae</i>	Sordariomycetes	YDS ¹ N*G*LIS ² W	<i>rnl</i>	TE M ¹ M ² L ¹ AFKL ² QHM ³	<u>AY884128</u>

598 ^a The underlined tRNA genes showed rearrangement if compared to the consensus (bold)

599 ^b The tRNA gene order of listed organisms is based on Genbank sequences

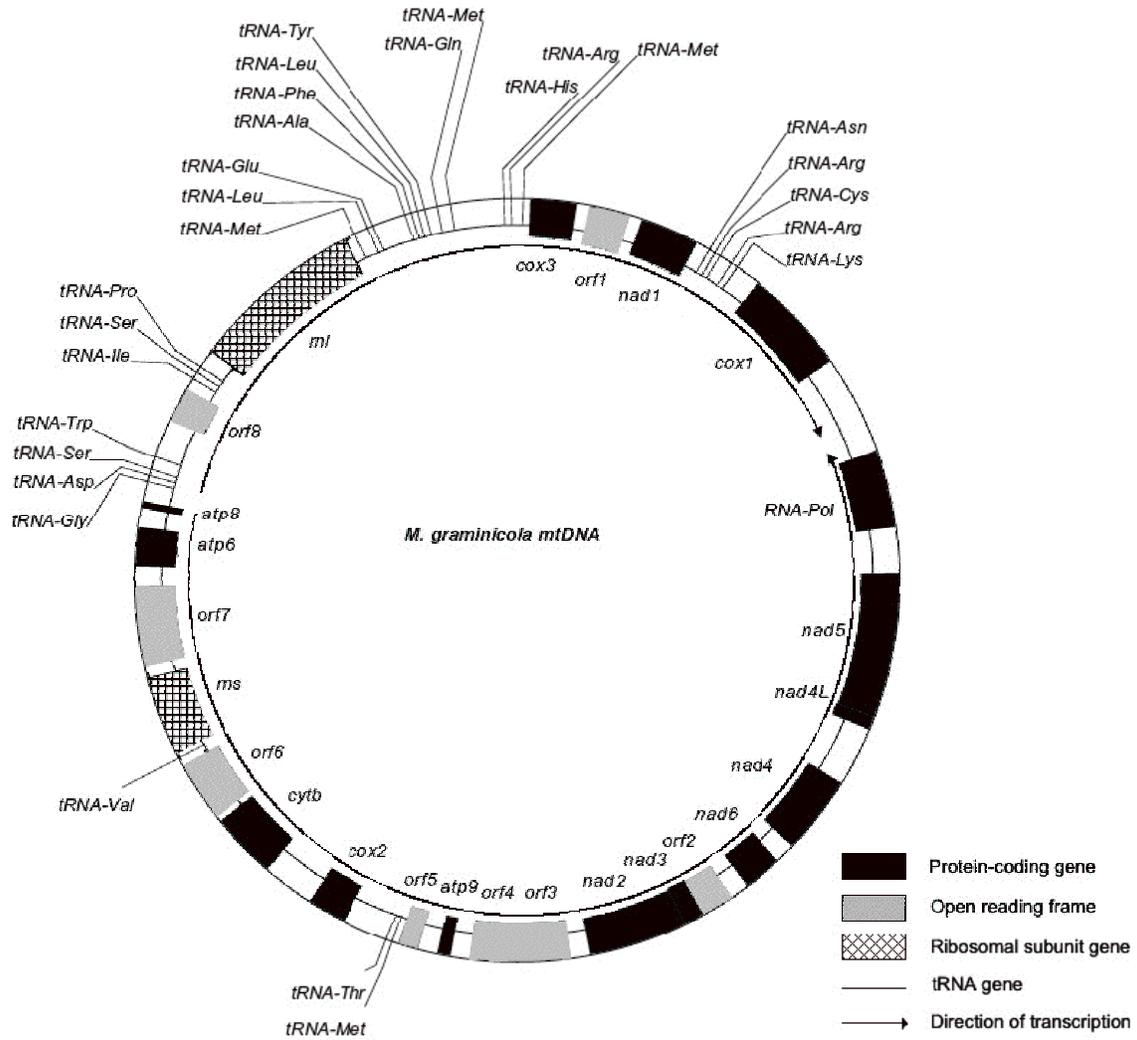
600 ^c Capital letters refer to tRNA genes for: R=arginine, K=lysine, G=glycine, D=aspartic acid, S=serine,

601 W=tryptophan, I=isoleucine, P=proline, T=threonine, E=glutamic acid, V=valine, L=leucine, A=alanine,

602 F=phenylalanine, Q=glutamine, H=histidine, Y=tyrosine, N=asparagine

603 * Asterisk indicates where functional genes interrupt the tRNA genes sequence

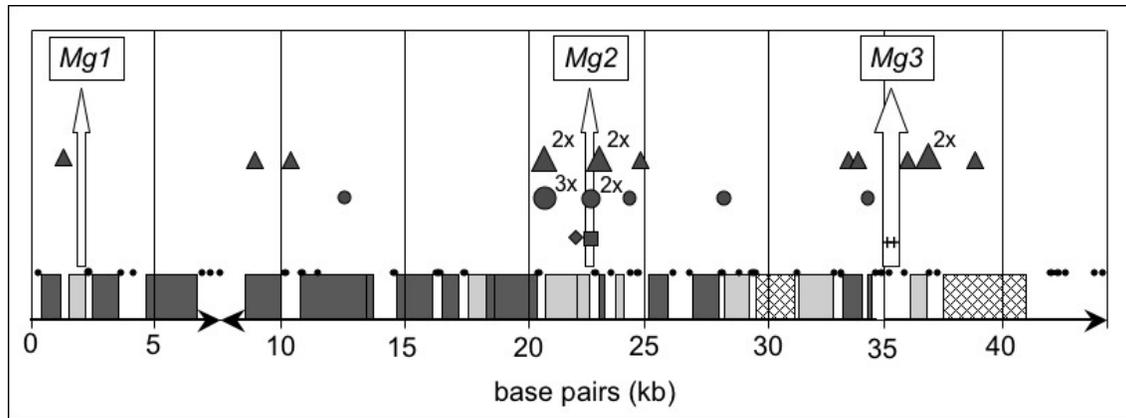
604 ^{1,2,3} The numbers indicate the presence of more tRNA genes for the same amino acid in the consensus sequences



605

606 Fig. 1. Circular map of the mitochondrial genome of *Mycosphaerella graminicola*. Black
607 blocks, grey blocks, hatched blocks and bars show, respectively, protein-coding, orfs, rRNA
608 and tRNA genes. Arrows indicate the direction of transcription.

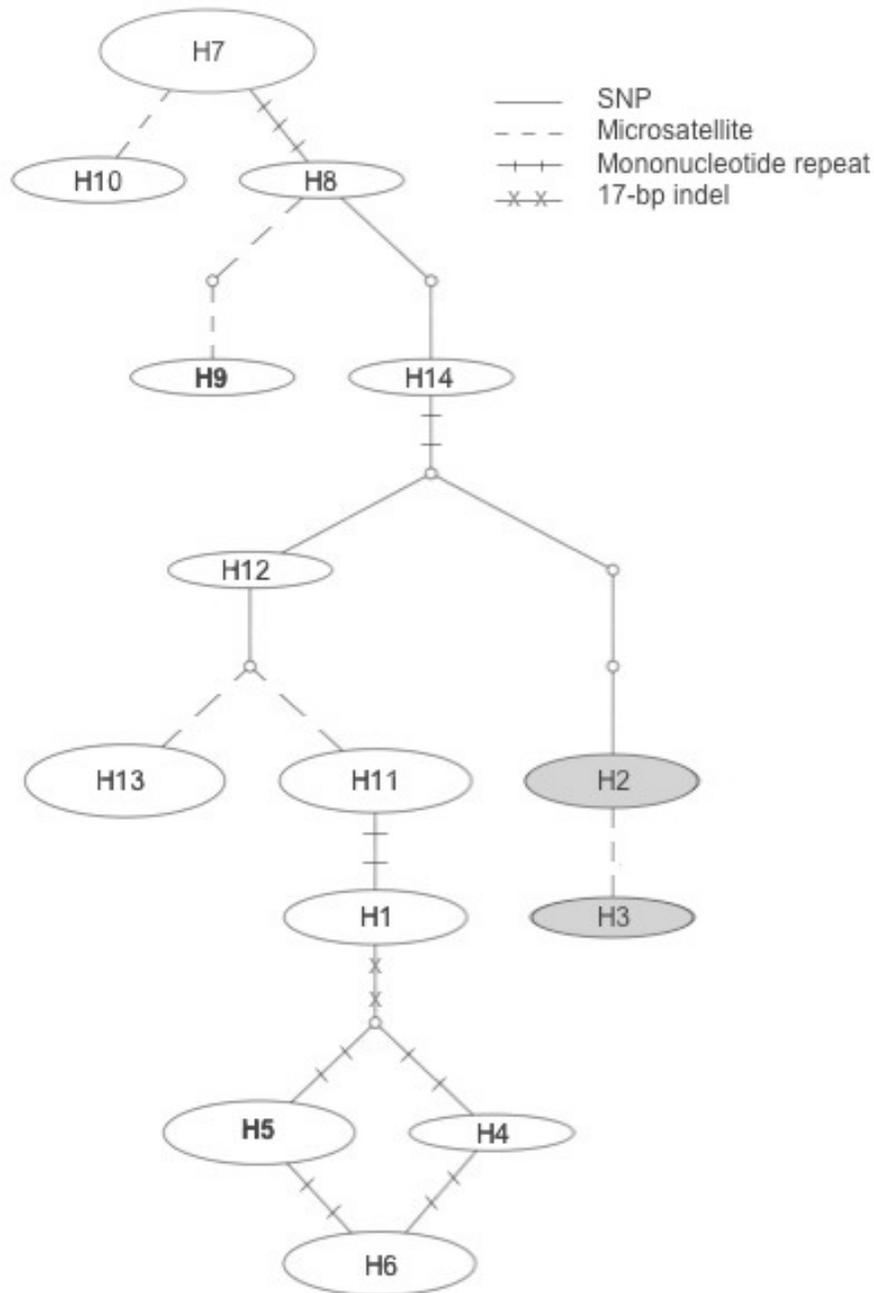
609



610

611 Fig. 2. A linearized map showing polymorphisms between the mt genomes of
612 *Mycosphaerella graminicola* isolates STBB1 and IPO323. Protein coding genes are presented
613 as dark-grey blocks, orfs as light-grey blocks and ribosomal subunit as hatched blocks. The
614 polymorphisms occurring between the mt genomes of isolates STBB1 and IPO323 are
615 transversions (triangles), transitions (circles), microsatellites (plus signs), a frameshift
616 (diamond), and a 17-bp-long indel (square). Arrows under genes indicate the direction of
617 transcription and the small black dots close to the genes show mononucleotide repeats longer
618 than 7 bases. The tRNAs are not presented to simplify the figure. *Mg1*, *Mg2* and *Mg3* were
619 the three regions used to assess the total intraspecific mt diversity.

620



621

622 Fig. 3. Phylogenetic relationships among *Mycosphaerella graminicola* mtDNA haplotypes
623 inferred using a parsimony haplotype network. Haplotypes are presented as ovals with sizes
624 proportional to the haplotype frequencies. Haplotypes containing isolates originating from
625 durum wheat (*Triticum turgidum* ssp. *durum*) are grey. Open circles are hypothetical missing

626 intermediate haplotypes. The completely sequenced haplotypes, H5 (IPO323) and H9
627 (STBB1), are indicated with bold text.

628

629

630 *This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and*
631 *Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under*
632 *contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-AC52-*
633 *07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396.*

634

635