Two Polyketide Synthase-Encoding Genes Are Required for Biosynthesis of the Polyketide Virulence Factor, T-toxin, by Cochliobolus heterostrophus

Scott E. Baker,1 Scott Kroken,1 Patrik Inderbitzin,2 Thipa Asvarak,2 Bi-Yu Li,1 Liang Shi,1 O. C. Yoder,1 and B. Gillian Turgeon1,2

1Torrey Mesa Research Institute (former)/Syngenta, 3115 Merryfield Row, San Diego, CA 92121 U.S.A.; 2Department of Plant Pathology, 334 Plant Science Bldg., Cornell University, Ithaca, NY 14853 U.S.A.


Cochliobolus heterostrophus race T, causal agent of southern corn leaf blight, requires T-toxin (a family of C35 to C49 polyketides) for high virulence on T-cytoplasm maize. Production of T-toxin is controlled by two unlinked loci, Tox1A and Tox1B, carried on 1.2 Mb of DNA not found in race O, a mildly virulent form of the fungus that does not produce T-toxin, or in any other Cochliobolus spp. or closely related fungus. PKS1, a polyketide synthase (PKS)-encoding gene at Tox1A, and DEC1, a decarboxylase-encoding gene at Tox1B, are necessary for T-toxin production. Although there is evidence that additional genes are required for T-toxin production, efforts to clone them have been frustrated because the genes are located in highly repeated, A+T-rich DNA. To overcome this difficulty, ligation specificity-based expression analysis display (LEAD), a comparative amplified fragment length polymorphism/gel fractionation/capillary sequencing procedure, was applied to cDNAs from a near-isogenic pair of race T (Tox1') and race O (Tox1') strains. This led to discovery of PKS2, a second PKS-encoding gene that maps at Tox1A and is required for both T-toxin biosynthesis and high virulence to maize. Thus, the carbon chain of each T-toxin family member likely is assembled by action of two PKSs, which produce two polyketides, one of which may act as the starter unit for biosynthesis of the mature T-toxin molecule.

Additional keywords: ascomycete, Didymella zae-maydis, Mycosphaerella zeae-maydis, Phyllosticta maydis.

Cochliobolus heterostrophus, the filamentous ascomycete that causes southern corn leaf blight, exists in two forms: race T, which produces a family of linear, predominantly C41 (described variously as C35 to C47 or C35 to C49) polyketides called T-toxin (Fig. 1) (Kono et al. 1980, 1981, 1983), and race O, which does not produce T-toxin. Both races are pathogenic to maize; however, race T is highly virulent to maize containing Texas male sterile (T) cytoplasm because this cytoplasm is uniquely sensitive to T-toxin; race O is weakly virulent on maize in general (Turgeon and Lu 2000; Yoder 1980; Yoder et al. 1997). Previous work (Kodama et al. 1999; Tzeng et al. 1989, 1992; Tzeng 1990) revealed that the genome of race T contains approximately 1.2 Mb of DNA not found in race O and that the Tox1 locus, which determines ability to produce T-toxin, maps to this genomic region. Conventional genetic methods in which segregation of pathogenicity types and host-specific toxin production were examined in progenies of crosses between naturally occurring race T and O isolates led to the prediction that Tox1 is a single locus (Yoder and Gracen 1975). However, molecular and genetic analyses of T-toxin-deficient race T mutants have shown that genetic control of T-toxin production maps at two loci (Tox1A and Tox1B) on two different race T chromosomes (designated 12;6 and 6;12) which are reciprocally translocated in race T with respect to their counterparts in race O (chromosomes 6 and 12) (Fig. 2) (Kodama et al. 1999; Tzeng et al. 1989, 1992). Tox1A, Tox1B, and the 1.2 Mb of DNA unique to race T are all genetically inseparable from each other and from the breakpoints of the translocated chromosomes. T-toxin production and high virulence depend on genes at both Tox1A and Tox1B. Two such genes have been described to date: PKS1 (Yang et al. 1996), which encodes a polyketide synthase (PKS) and maps to Tox1A on chromosome 12;6, and DEC1 (Rose et al. 2002), which encodes a decarboxylase and maps to Tox1B on chromosome 6;12. Another identified gene, RED1, encodes a reductase and is adjacent to DEC1 on chromosome 6;12 but is not required for T-toxin production (Asvarak 2003; Rose et al. 2002). Complementation of either a deletion (approximately 700 kb) covering PKS1 at Tox1A with a functional copy of PKS1 or a deletion (approximately 75 kb) covering DEC1 at Tox1B with a functional copy of DEC1 does not restore ability to produce T-toxin (Zhu 1999). Based on this, we concluded there must be as yet unidentified genes at both Tox1A and Tox1B required for T-toxin biosynthesis.

Previous attempts to identify additional genes at Tox1 using positional cloning procedures failed, likely due to the fact that both Tox1A and Tox1B contain A+T-rich (approximately 70%), repeated DNA (Kodama et al. 1999; Tzeng et al. 1992).
In addition, it has not been possible to detect genes at \textit{Tox1} that have been cloned already (PKS1, DEC1, and RED1) in any of several dozen libraries (plasmid, lambda, bacterial artificial chromosome, and yeast artificial chromosome) prepared from race T genomic DNA. To circumvent these problems, we used a cDNA-based strategy (Li et al. 2002) aimed at recovering genes expressed in race T but not in race O. This led to the identification of a second PKS-encoding gene, PKS2, which is unique to race T, maps at \textit{Tox1A}, and, like PKS1, is required for both \textit{T}-toxin biosynthesis and high virulence to maize.

**RESULTS**

Sequence analysis of ligation specificity-based expression analysis display fragments.

Amplified cDNA fragments differentially expressed between race O and race T were sequenced, either directly after extraction from gels or after extraction and cloning. In all, 47 cDNA fragments were sequenced. All were blasted against the \textit{C. heterostrophus} genome sequence contigs (Blastn; the proprietary genome was sequenced by Celera for the Torrey Mesa Research Institute [TMRI]/Syngenta) and against sequences in GenBank (tBlastx) (Altschul et al. 1990). Twelve of the sequences had similarity to PKSs and we focused on these (Table 1; Fig. 3). Figure 4 is a cDNA fingerprinting profile showing an example of a clear difference between triplicate race T and race O samples.

**Identification of PKS2.**

Of the 12 fragments with similarity to PKS-encoding genes, three corresponded to \textit{C. heterostrophus} \textit{PKS1} (Yang et al. 1996) (Table 1; Fig. 3). The other nine fragments corresponded to a single \textit{C. heterostrophus} genomic contig that carried an open reading frame (ORF) designated \textit{PKS2}. No other PKS fragments were identified. The \textit{PKS2} ORF was incomplete in the sequence assembly; therefore, gaps were filled by sequencing polymerase chain reaction (PCR) products amplified using primers corresponding to known sequences flanking the gaps. In addition, a near-full-length cDNA was obtained and sequenced to verify intron positions. The predicted PKS2 protein contains 2,144 amino acids and has domains corresponding to a ketosynthase (KS), an acyltransferase (AT), a dehydratase (DH), an enoyl reductase (ER), a ketoacyl reductase (KR), and an acyl carrier protein (ACP). There is no methyl transferase (MeT), in contrast to PKS1, which has a degenerate MeT. There are three introns (58, 56, and 88 bp), the second of which is in the same position as the third intron (of four) in \textit{PKS1}. Previous work (Kroken et al. 2003) suggested that the product of PKS2 is a linear, variously reduced polyketide, as is the product of PKS1.

**PKS2 is not present in race O.**

\textit{PKS2} was not identified from the genome of race O strain C5 by ligation specificity-based expression analysis display (LEAD) cDNA fingerprinting. To test the possibility that race O has an unexpressed version of PKS2, PCR amplification was performed with race T (strain C4) and race O (strain C5) gen-

---

**Table 1. Selected cDNA fingerprinting results**

<table>
<thead>
<tr>
<th>Restriction enzyme</th>
<th>Band no.</th>
<th>Size (bp)</th>
<th>Chromosome</th>
<th>Assembly contig</th>
<th>BLAST top hit</th>
</tr>
</thead>
<tbody>
<tr>
<td>EcoRI</td>
<td>E12*</td>
<td>264</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>KpnI</td>
<td>9</td>
<td>245</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>EcoRI</td>
<td>15</td>
<td>181</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>ApeLI</td>
<td>A28</td>
<td>333</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>KpnI</td>
<td>10</td>
<td>286</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>PstI</td>
<td>33</td>
<td>174</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>BglII</td>
<td>G1*</td>
<td>186</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>BamHI</td>
<td>B5</td>
<td>86</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>NcoI</td>
<td>N2*</td>
<td>130</td>
<td>12:6</td>
<td>4FP</td>
<td>ChPKS1</td>
</tr>
<tr>
<td>SphI</td>
<td>43</td>
<td>118</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>SphII</td>
<td>34*</td>
<td>191</td>
<td>12:6</td>
<td>4LU</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>BamHI</td>
<td>B1</td>
<td>192</td>
<td>12:6</td>
<td>4FP</td>
<td>ChPKS2</td>
</tr>
<tr>
<td>SacI</td>
<td>31</td>
<td>360</td>
<td>12:6</td>
<td>4LU</td>
<td>lambda crystallin-like</td>
</tr>
<tr>
<td>ApeLI</td>
<td>3</td>
<td>188</td>
<td>6:12</td>
<td>3PL</td>
<td>ChRED1</td>
</tr>
</tbody>
</table>

*a Shown is a subset ligation specificity-based expression analysis display (LEAD) fragments that were expressed clearly in race T but not in race O samples. Note that LEAD identified positive controls \textit{PKS1} and \textit{RED1}, both shown previously to be present in race T but not race O (Rose et al. 2002). LEAD PKS fragments are listed in the order (5' to 3') that they map to \textit{ChPKS1} and \textit{ChPKS2} (Fig. 3).

*b TMRI internal identification number. Bands marked by an asterisk (*) were cloned.

*c Tzeng et al. (1992); Kodama et al. (1999).

*d TMRI internal supercontig or scaffold identification number.

*e Blastn query of the genome assembly, tBlastx query of GenBank.

---

**Fig. 1.** T-toxin structure. Structures of the two most abundant representatives (C_{39} and C_{41}) of the T-toxin family of molecules. Family members are long-chain (C_{35} to C_{47} or C_{49} [Kono et al. 1980, 1982, 1983]), linear, saturated hydrocarbons with an odd number of carbon atoms and a pattern of functional groups that is repeated four times. Asterisk indicates position C{24} which may harbor a keto (as shown) or a hydroxyl group; biological activity is the same.
nomic DNAs as templates and primers corresponding to specific regions of \textit{PKS2} located outside the conserved PKS enzymatic domains in order to avoid nonspecific amplification of other \textit{PKS}-encoding genes. A fragment of predicted size was amplified from race T but not race O DNA (Fig. 5; compare lanes 4 and 5 with lanes 8 and 9). As a positive control, primers corresponding to an unknown gene recovered by LEAD that was present in both races, but more highly expressed in race T than in race O, were shown to amplify a product of predicted size from both races (Fig. 5; compare lanes 2 and 3 with lanes 6 and 7).

**Effect of \textit{PKS2} deletion on T-toxin production.**
To assess the role of \textit{PKS2} in T-toxin production, a construct designed to delete the putative KS domain (Fig. 3) was generated and used to transform race T. Genomic DNA preparations from three purified hygromycin B–resistant transformants (C4.910.1B, C4.910.2B, and C4.910.3B) were used as templates in PCR reactions with primers INT5 and INT3. All three transformants sustained homologous integration events resulting in deletion of the target DNA (Fig. 6). The microbial assay for T-toxin showed that none of the \textit{pks2} transformants produced detectable T-toxin, whereas T-toxin production by wild-type race T strain C4 (\textit{PKS2}) was normal (Fig. 7).

**Virulence of \textit{pks2} strains.**
All \textit{pks2} strains produced small lesions on T-cytoplasm maize similar to those of race O, indicating that inability to make T-toxin is associated with loss of high virulence of race T on T-cytoplasm maize (Fig. 8). This reduced virulence phenotype is similar to those observed for \textit{pks1} and \textit{dec1} mutant strains (Rose et al. 2002; Yang et al. 1996).

**Mapping \textit{PKS2}.**
Crosses were made between \textit{pks2} strain C4.910 and \textit{pks1} strain 1326-R-5, which carries a deletion of \textit{PKS1} at Tox1A, and between \textit{pks2} strain C4.910 and \textit{dec1} strain 1213-R-64, which carries a mutation in \textit{DEC1} at Tox1B; all mutations are tagged with \textit{hygB} (Table 2). Of 164 progeny collected from the first cross (\textit{pks2} X \textit{pks1}), none was able to produce T-toxin, indicating that \textit{PKS2} is tightly linked to \textit{PKS1} and, therefore, maps at Tox1A. Progeny (\(n = 356\)) from the second cross (\textit{pks2} X \textit{dec1}) segregated 271:85 \textit{hygB}R:\textit{hygB}S, which approximates a 3:1 ratio and indicates two unlinked genes in a haploid organism. Of the total 356, 55 random progeny tested for ability to produce T-toxin segregated 3:1 (42:13 \textit{hygB}R;\textit{tox}–:\textit{hygB}S;Tox+), confirming that the \textit{PKS2} deletion cosegregates with lack of T-toxin production and that there is no linkage between the Tox1B mutant (\textit{dec1};\textit{hygB}R) and the Tox1A mutant (\textit{pks2};\textit{hygB}R).

**Fig. 2.** Diagrammatic overview of \textit{Tox1} organization. Electrophoretic karyotyping of chromosomes of highly backcrossed race T by race O laboratory strains (Kodama et al. 1999; Tzeng et al. 1989) indicates that most chromosomes are the same size in the two isogenic races, except that race O has a pair (chromosome 6 and 12) that is missing in race T. Race T has a pair (chromosome 6;12 and 12;6) that is reciprocally translocated with respect to the race O pair and, in addition, is approximately 1.2 Mb larger than the sum of 6 and 12 in race O. This extra DNA cosegregates with both breakpoints of the reciprocal translocation and is highly A+T-rich and repeated. Known \textit{Tox1} genes (\textit{PKS1}, \textit{DEC1}, and \textit{RED1}) (Kodama et al. 1999; Rose et al. 2002) map to loci (Tox1A and Tox1B) at each of the two breakpoints.

**Fig. 3.** Diagram of a generic polyketide synthase (PKS) protein showing locations of ligation specificity-based expression analysis display ‘hits’ (arrows). Protein enzymatic domains are KS = ketosynthase, AT = acyltransferase, DH = dehydratase, ER = enoyl reductase, KR = ketoacyl reductase, and ACP = acyl carrier protein. Gray arrows = hits to 	extit{Cochliobolus heterostrophus} \textit{PKS1} and black = hits to \textit{C. heterostrophus} \textit{PKS2}.
PKS2 and PKS1 are unique to race T, whereas the remaining 23 PKS genes are common to both races.

Kroken and associates (2003) described 25 PKS-encoding genes in the *C. heterostrophus* race T genome sequence. Except for reactions with PKS1- and PKS2-specific primers, products were amplified from both C4 (race T) and C5 (race O) genomic DNA, when DNAs from both races were tested for the presence of each PKS (Fig. 9). Thus, none of the other 23 PKS genes in the *C. heterostrophus* genome is a candidate for a role in T-toxin biosynthesis, because each is common to both race T and race O. PCR products were of predicted size and were the same size in both races (Fig. 9).

Phylogenetic relationship of PKS2 to other PKS genes.

The amino acid sequences of the KS and AT domain from PKS2 and from all fungal PKSs available in GenBank were used to construct a phylogenetic tree illustrating the relationship of PKS1 and PKS2 to each other and to known fungal PKSs (Fig. 10A and B). Both PKS1 and PKS2 are found in a large clade of PKSs (77% bootstrap support) in which all of the characterized gene products make linear and variously reduced polyketide intermediates. This clade includes the majority of PKSs that make fungal toxins, and is subdivided into three subclades. One of these (fungal reducing PKS clade I, 62% bootstrap support) contains *ChPKS1* (Fig. 10A, thick arrow) and its ortholog *MzmPKS1* (Fig. 10A, double-headed arrow). The latter is responsible for *Phyllosticta maydis* toxin (PM-toxin) production (Yun 1998; Yun et al. 1998) by *Didymella zaeae-maydis* (Mukunya & Boothroyd) von Arx, formerly *Mycosphaerella zaeae-maydis* Mukunya & Boothroyd (Mukunya and Boothroyd 1973; von Arx 1987), and also known as *P. maydis* Arny & Nelson (Arny and Nelson 1971). In addition, this clade contains *ChPKS2* (Fig. 10A, thick arrow, above *ChPKS1*), and two orthologous genes whose PKS enzymes make the diketide component of lovastatin (*Aspergillus terreus* *lovF*) and compactin/citrinin (*Penicillium citrinum* *mlcB*), respectively (thin arrows above *ChPKS2*). *ChPKS1* and *ChPKS2* are not closely related within this clade, because they are not united by any branch that received significant (>50%) bootstrap support. Similarly, neither *ChPKS1* nor *ChPKS2* is closely related to the only other characterized PKS in this subclade, *pksN*, which makes the polyketide component of alternapyrone in *Alternaria solani* (accession no. BAD83684) (Fig. 10A, white arrow).

**DISCUSSION**

PKS2 is the third gene identified in *C. heterostrophus* race T that is required for T-toxin production and high virulence to maize. All three *Tox* genes (*PKS1*, *PKS2*, and *DEC1*) map at the *Tox1* locus, which is composed of two large segments of DNA (totaling approximately 1.2 Mb) that map to two chromosomes in race T (Turgeon and Lu 2000). *PKS1* and *PKS2* map to *Tox1A*, whereas *DEC1* maps to *Tox1B* (Fig. 2).

The first identification of a gene at *Tox1* (*PKS1*) was by REMI mutagenesis (Yang et al. 1996), whereas *DEC1* was found among cDNA fragments present in wild-type race T, but missing in a *Tox–* mutant carrying a 75-kb deletion at *Tox1B* (Rose et al. 2002). Here, PKS2 was identified by taking advantage of the fact that the two genes heretofore known to be involved in T-toxin production are missing from race O, an observation that predicts that genes expressed in race T, but not in near-isogenic race O, are candidate *Tox1* genes. The LEAD procedure (Li et al. 2002), which is based on differential gene

---

**Fig. 4.** Screen shot image of ligation specificity-based expression analysis display cDNA fingerprint profile of race T strain C4 (*Tox+*) versus race O strain C5 (*Tox–*) samples. Top row is the combined profiles of six independent samples. Next three rows are three race T samples and the bottom two are race O samples (third race O sample is not shown). Arrow indicates a peak in race T samples that is missing in race O samples.
expression between two samples (in this case, between mRNA of races T and O), efficiently identified PKS2 as a candidate Tox1 gene and also identified the previously known PKS1 and RED1 genes (Table 1).

PKS2 is distinct from PKS1 in its sequence, domain structure, and phylogenetic placement (Fig. 10A), suggesting different functions of the two gene products in the assembly of the T-toxin molecule. The two PKS proteins are only 32% identical and 50% similar at the amino acid level in regions of alignment, and PKS2 is shorter than PKS1 (2,144 vs. 2,528 amino acids), partly because it lacks a degenerate MeT domain present in PKS1. PKS1 is similar to other PKSs that have residual MeT domains (Kroken et al. 2003) with poorly conserved motifs that likely are nonfunctional. This domain would not be predicted as a requirement for T-toxin biosynthesis because the molecule is not methylated. A methylated intermediate that subsequently is demethylated is unlikely in the biosynthetic pathway; there is no precedent for this unparsimonious step in the biochemical pathway of any fungal polyketide derivative, and we have not yet found a demethylase at the Tox1 locus. To characterize the role of the degenerate MeT domain in T-toxin biosynthesis, site-directed mutagenesis could be used to knock out functionality while preserving the overall size and topology of this domain in the PKS protein.

Both PKS1 and PKS2 fall in the major fungal clade of PKSs involved in biosynthesis of linear and variously reduced polyketides but are not closely related to each other and, thus, do not represent a recent duplication of one or the other of these PKS genes. In contrast to PKS2, which has no known ortholog, PKS1 is orthologous to \( M_{\text{zm}} \text{PKS1} \) (Fig. 10A, thick arrow) of \( D.\ zeae\text{-}maydis \) (Kroken et al. 2003; Yun et al. 1998), also a pathogen of maize (Turgeon and Lu 2000; Yoder et al. 1997). \( M_{\text{zm}} \text{PKS1} \) makes a family of linear polyketides (PM-toxin), members of which are structurally similar to T-toxin and have the same biological activity (Kono et al. 1983; Turgeon and Lu 2000; Yoder 1973, 1980; Yoder and Mukunya 1972). It currently is unknown whether or not \( D.\ zeae\text{-}maydis \) has an ortholog of PKS2 or of any of the other genes found at the \( C.\ heterostrophus \) Tox1 locus. We do know that the immediate neighborhood around \( M_{\text{zm}} \text{PKS1} \) is different from that of \( Ch \text{PKS1} \) (Yun 1998).

The discovery that PKS2 is required, along with PKS1, for T-toxin biosynthesis makes it likely that T-toxin is the combined product of two interacting PKSs. There is precedent for a mechanism that links two polyketide chains, namely in the two orthologous biosynthetic pathways that produce lovastatin and compactin, in which diketide side chains made by the PKSs LDKS or MLCB, respectively, are attached to nonaketides, fashioned iteratively by the PKSs LNKS or MLCA, respectively (Fig. 10A, thin arrows) (Abel et al. 2002; Kennedy et al. 1999). In lovastatin biosynthesis, the monomodular PKS (LNKS [LOVB], encoded by \( \text{lovB} \)) produces the nonaketide in an iterative, highly controlled way, differentially utilizing reducing domains at each step. A diketide produced by LDKS (LOVF, encoded by \( \text{lovF} \)) is then attached to the modified LNKS product by means of an ester bond.

**Fig. 5.** PKS2 is unique to race T. Polymerase chain reaction (PCR) amplification using primers specific to a sequence corresponding to PKS2 and race T strain C4 (lanes 4 and 5) or race O strain C5 DNA (lanes 8 and 9) as template yielded a product from race T only (lanes 4 and 5, no product in lanes 8 and 9). PCR amplification, performed using primers specific to a sequence corresponding to a ligation specificity-based expression analysis display fragment that was present in both races but more highly expressed in race T than in race O, and the same template DNAs yielded a product of the same size from both race T (lanes 2 and 3) and race O (lanes 6 and 7). Lanes 1 and 10 are the 1-kb plus marker (Invitrogen).

**Fig. 6.** Proof of PKS2 ketsynasate region deletion. Primers corresponding to sequences outside of the \( 5’ \) and \( 3’ \) genomic flanks on the deletion construct were designed and used in polymerase chain reactions with DNA preparations from wild-type strain C4 (lane 2) and three different candidate \( pks2 \) single-spore isolates (lanes 3 to 5) as templates. Products were obtained in all cases; however, amplicons (approximately 2.9 kb) in lanes 3 to 5 showed an increased intensity in size compared with the wild type (approximately 1.1 kb), indicative of replacement of a portion (286 bp) of the PKS2 gene with the hygromycin B cassette (approximately 2.1 kb). Lane 1 is the 1-kb plus marker (Invitrogen).

**Fig. 7.** Microbial assay of T-toxin production. Plate was spread with T-toxin sensitive \( E.\ coli \) cells (Ciuffetti et al. 1992). Zone of clearing around the positive race T mycelial plug control (arrow) indicates production of T-toxin. Absence of a halo around the negative race O plug control (double arrow) indicates no T-toxin production. Absence of halos around duplicates of the three candidate \( pks2 \) strains, analyzed as described in Figure 6, indicates that there is no detectable T-toxin (bottom two rows). Thus, PKS2 is required for T-toxin production. Mycelial plug in top right corner of plate is from a mutant strain that is reduced, but not null, in T-toxin production compared with the wild type (Yang et al. 1994).
Hutchinson and associates (2000), commenting on lovastatin biosynthesis, noted that “Many other nonaromatic fungal polyketides are made by PKSs that probably resemble LNKS and LDKS closely. Only two such genes have been reported to date; one is involved in T-toxin (Yang et al. 1996) [biosynthesis]….” This reference was to PKS1, the only PKS involved in T-toxin production identified at that point. The suggestion that the variously reduced β-keto groups of T-toxin result from highly controlled and coordinated differential use, in each iterative cycle, of the various reducing domains in PKS1 (or PKS2 or both), as with LNKS in lovastatin biosynthesis, is likely apt. Furthermore, the lovastatin model for building T-toxin requires two and only two PKSs, in line with the observation that only 2 (PKS1 and PKS2) of the 25 PKS-encoding genes found in the C. heterostrophus genome are unique to race T (Fig. 9). However, the proposed mechanism for joining the two T-toxin polyketides must be different from that required for the lovastatin pathway. With lovastatin, a dehydration, catalyzed by the transesterase lovD (Kennedy et al. 1999), results in a cyclized, branched polyketide. The T-toxin family members remain as linear polyketides; the major component is a C41 chain with four groups of three unreduced to partially reduced β-ketones, separated by one or two completely reduced β-ketones (Fig. 1). The structure of T-toxin suggests that the two polyketides produced by PKS1 and PKS2 are joined via a carbon-carbon bond, not by an ester bond as in lovastatin.

The following hypothesis for T-toxin synthesis is proposed. T-toxin synthesis is similar to lovastatin synthesis in that two PKSs are involved, both of which are monomodular and, given the differential processing of the β-keto groups in the final product, likely exhibit a high degree of control, per iterative cycle, over the use of the various reducing domains. However, the relative contribution of each PKS to T-toxin synthesis is unknown. As noted above, T-toxin lacks the ester group formed in lovastatin through the concatenation of the two polyketide chains, and the putative PKS1 and PKS2 gene products appear to be joined by a carbon-carbon bond, reminiscent of the mechanism of aflatoxin synthesis (Hitchman et al. 2001; Yabe and Nakajima 2004).

Synthesis of the aflatoxin backbone requires the interaction of a fatty acid synthetase (FAS) and a PKS, two proteins with the same domain structure, encoded by members of the same gene superfamily (Hopwood 1997). The FAS synthesizes a C6 starter unit for the PKS (Hitchman et al. 2001). PKS1 and PKS2 could interact in a similar way. One of the PKSs could synthesize a precursor, used subsequently by the other PKS as starter unit, to add additional carbons. This mechanism also is reminiscent of the action of modular bacterial PKSs, in which the ACP of the upstream protein subunit passes the nascent chain to the KS domain of the downstream subunit (Staunton and Weissman 2001). Variability in the length of the final carbon backbone C35–47 could be achieved by varying the number of condensation cycles, or use of different starter or extender units or might be due to decarboxylation of the penultimate product, catalyzed by DEC1, as proposed earlier (Rose et al. 2002).

The dual PKS requirement for T-toxin biosynthesis is exciting because, to our knowledge, a mechanism in which two polyketides are stitched together via a carbon-carbon bond and in which the polyketide produced by one of the PKSs is used as a starter for the second PKS to produce the final metabolite has not been described for filamentous fungi. Most characterized type I PKSs from filamentous fungi act iteratively (i.e., a single monomodular PKS is used repeatedly to build the full-length final molecule or set of molecules). Two exceptions, described above, are the pairs of PKSs involved in lovastatin and compactin biosynthesis. In these cases, however, the two polyketides are joined by an ester, not a carbon-carbon bond. Bacterial polyketides, on the other hand, usually are synthesized by noniterative type I PKS complexes, encoded contiguously as modules in the bacterial genome. The linear, modular arrangement of bacte-

Fig. 8. Virulence assay. Leaves from 2-week-old T-cytoplasm maize seedlings sprayed with 15 ml of conidial suspension (1,000 spores/ml) and photographed 4 days after inoculation. Inoculated with (from left): control uninoculated, race T (strain C4), race O (strain C5), pks1 strain 1326-R-5, pks2 strain C4.910.1A, and pks2 strain C4.910.2A. Note that all toxI mutants (pks2 and pks1) are reduced in virulence compared with their progenitor race T strain and show symptoms similar to race O.
rial PKS genes makes it possible to deduce the encoded PKSs, their enzymatic domain organizations, the nature of the growing polyketide that is passed along to the adjacent module, and, ultimately, the biochemical structure of the metabolite produced. How a growing polyketide chain is processed is not obvious with iterative PKSs; reducing domains, if present, may or may not be employed in any given cycle. In addition, although we now know that biosynthesis of T-toxin requires two PKSs, the corresponding PKS genes are not adjacent to each other, although they do appear to be linked at ToxIA.

Using a novel method (LEAD), we were able to identify and characterize a second gene, encoding a PKS, that is involved in biosynthesis of the family of polyketides known as T-toxin, when other approaches failed. The number of additional genes required for the biosynthesis of T-toxin is unknown, but there are at least two more. One is the lambda crystallin-like gene, LAM1, also identified by LEAD (Table 1) (Asvarak 2003), which encodes a putative 3-hydroxyacyl-CoA dehydrogenase, and maps near PKS2 at ToxIA (based on the C. heterostrophus genome assembly). The existence of at least one more gene necessary for T-toxin biosynthesis is deduced from the observation that a full-length functional clone of DEC1 cannot complement a 75-kb deletion including DEC1 at ToxIB (Zhu 1999).

Although all known Toxl genes map to either ToxIA or ToxIB, they are scattered in the highly A + T-rich DNA of this region. For example, PKS1 corresponds to the only recognizable ORF in more than 20 kb of sequence and PKS2 and LAM1 are the only ORFs in more than 30 kb of sequence. The C. heterostrophus sequence assembly confirmed the complexity of this region; contigs carrying known Toxl genes were the smallest in the assembly, attesting to the difficulty of assembling highly repeated DNA. In light of the fact that secondary metabolite gene clusters may be found frequently in difficult to cloning DNA. In light of the fact that secondary metabolite gene clusters may be found frequently in difficult to assemble highly repeated DNA. In light of the fact that secondary metabolite gene clusters may be found frequently in difficult to sequence, A + T-rich regions (Zhang et al. 2004) that may be recalcitrant to standard cloning procedures, we expect that the LEAD technique will be useful for the characterization of genes in other such biochemical pathways.

MATERIALS AND METHODS

Strains, media, crosses, and transformation.

Near-isogenic C. heterostrophus laboratory strains C4 (ToxlA-MAT1-2; American Type Culture Collection [ATCC] number 48331) and C5 (ToxlA-MAT1-1; ATCC number 48332) (Leach et al. 1982) were used for the LEAD procedure (Li et al. 2002). T-toxin-deficient mutants 1326-R-5 (pks1;MAT1-1;hygB) and 1213-R-64 (dec1; MAT1-1;hygB) (Rose et al. 2002) were used as ToxIA and ToxIB mutant tester strains, respectively, in crosses to pks2 mutants. All strains were grown on complete medium (CM) (Leach et al. 1982) with xylose (CMX) (Tseng et al. 1992). Growth conditions, mating, and transformation of C. heterostrophus were described previously (Leach et al. 1982; Turgeon et al. 1985, 1993).

LEAD procedure.

Isolation of C. heterostrophus RNA and the LEAD cDNA fingerprinting methodology were described previously (Li et al. 2002). Conidia (10⁶) from either strain C4 or C5 were grown in still culture at room temperature for 48 h in 50 ml of Fries medium (Pringle and Braun 1957). Total RNA was isolated from these samples using RNAzol (Ambion, Austin, TX U.S.A.), cDNA was synthesized and digested with BsaI, a four-base cutter that recognizes CCNNGG. A second digestion was performed with each of ApaLI, BamHI, BgII, EcoRI, NcoI, Apal, KpnI, PstI, SacI, and SphI. Each cDNA preparation was subdivided into 16 pools (corresponding to the degeneracy of BsaI) and ligated to a pair of adapters matching one of 16 possible BsaI sequences. Primers specific to each adapter (one of which was fluorescently labeled with NED-dye) were used to amplify fragments from each of the 16 pools, which were separated electrophoretically using an ABI PRISM 3700 DNA Analyzer. The computer program GeneScan (Li et al. 2002) recorded and quantified each fragment. Fragment profiles then were analyzed by LEADfinder software (Li

Table 2. Mapping of PKS2

<table>
<thead>
<tr>
<th>Parents</th>
<th>No. of ascospores</th>
<th>hygB² tox–</th>
<th>hygB² tox+</th>
<th>hygB¹ tox–</th>
<th>hygB¹ tox+</th>
</tr>
</thead>
<tbody>
<tr>
<td>pks2– X pks1–</td>
<td>164</td>
<td>164</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>pks2– X dec1–</td>
<td>356 (55)</td>
<td>271 (42)</td>
<td>0</td>
<td>0</td>
<td>85 (13)</td>
</tr>
</tbody>
</table>

a Parents: pks2– is represented by strain C4,910 (pks2;MAT1-2;hygB), pks1– is represented by strain 1326-R-5 (pks1;MAT1-1;hygB), and dec1– is represented by strain 1213-R-64 (dec1;MAT1-1;hygB).

b Numbers in parentheses = number of progeny tested for toxin production and number of Tox– or Tox+.
Fig. 10. A and B, Phylogeny of fungal polyketide synthase (PKS) genes inferred from amino acid sequence of their ketosynthase (KS) and acyl transferase (AT) domains. Miniature complete tree (separated into A and B groupings) is on upper left of each partial tree (A or B). The tree includes all full-length PKS sequences available in GenBank as of April 2005. Numbers below branches are bootstrap support values. Putative orthologs (subclades of PKSs supported by bootstrap values of 100%) and whose topology is consistent with presumed organismal phylogeny) are noted with branches drawn as heavy lines. The overall topology is similar to that reported by Kroken and associates (2003), with the exception that the number of major clades of reducing PKSs has declined from four to three. Topology shown is one of two most parsimonious trees. Rescaled consistency index = 0.243. A, On this page, large arrows point to Cochliobolus heterostrophus PKS1 and PKS2 (involved in T-toxin production); double arrow indicates Didymella zeae-maydis PKS1 (involved in Phyllosticta maydis PKS1 (involved in Phyllosticta maydis toxin [PM-toxin] production); white arrow indicates the Alternaria solani pksN gene responsible for alternapyrone production; small black arrows indicate PKSs lovF and mclB involved in diketide production by Aspergillus terreus (producing lovastatin) and Penicillium citrinum (producing compactin/citrinin), respectively; and the PKSs lovB and mecA involved in nonaketide production by A. terreus and P. citrinum, respectively.
et al. 2002), and peaks differentially expressed in C4 and C5 were identified and correlated with bands run on conventional gels. Fragments of interest were isolated from the gel and either sequenced directly or cloned into pCR4-TOPO using a TOPO PCR cloning kit (Invitrogen, Carlsbad, CA, U.S.A.). In cases where the fragment was cloned, six candidate colonies were picked and plasmids sequenced.

Sequence of the full-length PKS2 ORF in the genome and cloning of PKS2 cDNA.

PKS2 mapped to an approximately 30-kb scaffold or supercontig designated 4PL (consisting of seven contigs) in the C. heterostrophus genome assembly; annotation of the PKS2 ORF indicated that it was incomplete. Primers, designed to fill gaps of 1,234 bp between contigs 4 and 5 and 270 bp between contigs 5 and 6, were designed and used with race T DNA to amplify products that were sequenced and used to construct a full-length ORF. In addition, a 3′ rapid amplification of cDNA ends (RACE) kit was used to amplify a nearly full-length cDNA from total RNA (isolated as described above for the LEAD experiment) using primer 5NF (Invitrogen). The ampli-con was cloned into the pCR-XL-TOPO vector as specified by the manufacturer (Invitrogen). The plasmid insert was sequenced using a transposon insertion-based DNA strategy (EZ Tn5 Insertion Kit, Epicentre Biotechnologies, Madison, WI U.S.A.) and an ABI3700 Sequencer. Sequence data were compared with genome sequence. The PKS2 cDNA (accession no. AY495643) and the genomic DNA (accession no. DQ186598) are deposited in GenBank.

Deletion of PKS2.

A deletion construct was generated using the PCR splicing by overlap extension method (PCR SOEing) (Horton et al. 1990) using HotStarTaq. Briefly, genomic sequences flanking a 286-bp portion of the KS region of the PKS2 ORF were amplified from purified genomic DNA using primers 5FF/5RHyg and 3FF/3RHyg for the 5′ and 3′ flanks, respectively. Primers 5RHyg and 3RHyg were designed as hybrid sequences that each contained PKS2-specific sequences and sequences matching the hygromycin B resistance cassette. Each flank was fused to an ampiclon of the hygromycin cassette (from vector pUCATPH) (Lu et al. 1994) as template and these products were fused to each other by PCR using nested primers 5NF and 3NF, thus creating the deletion construct. Primer sequences were, for the 5′ flank, forward primer (5FF) 5′-TCTTCTCGTTACGCGACCA-3′, nested forward primer (5NF), 5′-TCTGTAATGTCGACCATGCCGCACCA-3′, and reverse primer (plus hygromycin extension) (5RHyg) 5′-CAATATCATCTTCGTGCATCCTTCGTGCATCCTTCGTGAC TCTATGGACAGTGCGGCTGTGTCTC-3′; for the 3′ flank, (3FF) forward primer 5′-AGCTTGCACTCCTTGCACTCAG-3′, nested forward primer (3NF) 5′-GCCATTCGTGCCGGCCGTGAT-3′, and reverse primer plus hygromycin extension (3RHyg) 5′-AATCCTTCTTTAAGGAGTCCGCCGCTCAGTTACGC ATACA-3′; and, to amplify the hygromycin cassette, primer1

**Fig. 10. Continued from preceding page.** B, Fungal nonreducing polyketide synthase clades, including the melanins. Branches not present in a strict consensus are shown as a dashed line (found in the melanin subclade).
(Hyg1) 5′-TAGAGTCACAGAAGATGATTG-3′ and hygromycin primer2 (Hyg2) 5′-GGATCCCTAGAAAGGATT-3′. Underlined sequences indicate sequence overlap with the hygromycin cassette.

DNA from transformants was prepared using a Qiagen Plant DNAeasy kit following instructions supplied by the manufacturer. For verification of targeted integration of the deletion construct PKS2, integration 5′ primer (INT5) 5′ TCTACCCCAA TCTCCGCTCAA 3′ and integration 3′ primer (INT3) 5′ TGG CTTAAATCAGTGCTGA 3′, located outside the genomic DNA on the deletion construct, were used. All transformants were purified by isolation of single conidia.

Assays for T-toxin and fungal virulence.

The microbial T-toxin assay was described previously (Ciuffetti et al. 1992). To evaluate virulence, 2-week-old seedlings of maize inbred W64A with T cytoplasm (five plants per 4-inch pot) were sprayed with 15 ml of conidial suspension (1,000 spores/ml) in Tween-20 (0.05%) as described previously (Robbertse et al. 2003). Leaves were photographed 4 days after inoculation.

Screening of the complete set of C. heterostrophus PKSs for their presence in race O and race T.

Sequence of all 25 predicted C. heterostrophus PKSs (Kroken et al. 2003) (PKS1 to PKS25, accession no. U68040, AY95643 to AY95666, respectively) were used to design PKS gene-specific primers. These primers and template genomic DNA from strain C4 (race T) and strain C5 (race O) were used in PCR reactions to determine whether each individual PKS found in race T also was present in race O. Primer sequence sets are available on request.

Phylogenetic analysis of PKS proteins.

Phylogenetic analyses were conducted as described (Kroken et al. 2003). Briefly, the predicted protein sequences of the ketosynthase and acyl transferase domains from each PKS were aligned in ClustalX (Thompson et al. 1997) and used to construct a gene genealogy, using parsimony in PAUP* (Swofford 2003). Briefly, the predicted protein sequences of the ketosynthase and acyl transferase domains from each PKS were compared with annotation in Kroken and associates (2003).

ACKNOWLEDGMENTS

B. G. Turgeon acknowledges the support of USDA/NRI grant #98-35303-6483 to O. C. Yoder and B. G. Turgeon. P. Inderbitzin acknowledges postdoctoral fellowships from the Natural Sciences and Engineering Research Council of Canada and the Swiss National Science Foundation. During the course of this work, S. Kroken was supported by a collaborative agreement between the University of California–Berkeley (N. L. Glass and J. W. Taylor, Department of Plant and Microbial Biology) and the former TMRI/Syngenta. The genome of C. heterostrophus was sequenced by Celera Genomics for TMRI/Syngenta.

LITERATURE CITED


**AUTHOR-RECOMMENDED INTERNET RESOURCES**

The Broad Institute’s Magnaporthe grisea and Fusarium graminearum databases: www.broad.mit.edu/annotation/fungi/magnaporthe and www.broad.mit.edu/annotation/fungi/fusarium