A novel class of gene controlling virulence in plant pathogenic ascomycete fungi

Shun-Wen Lu*, Scott Kroken†‡§, Bee-Na Lee‡, Barbara Robbertse‡, Alice C. L. Churchill¶, O. C. Yoder‡§‖, and B. Gillian Turgeon*‖**

*Department of Plant Pathology, Cornell University, Ithaca, NY 14853; †Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720; ‡Torrey Mesa Research Institute, San Diego, CA 92121; and ¶Boyce Thompson Institute for Plant Research, Ithaca, NY 14853

Communicated by Steven P. Briggs, Torrey Mesa Research Institute, San Diego, CA, March 10, 2003 (received for review December 15, 2002).

Insertional mutants of the fungal maize pathogen *Cochliobolus heterostrophus* were screened for altered virulence. One mutant had 60% reduction in lesion size relative to WT but no other detectable change in phenotype. Analysis of sequence at the insertion site revealed a gene (*CPS1*) encoding a protein with two AMP-binding domains. *CPS1* orthologs were detected in all *Cochliobolus* spp. examined, in several other classes of ascomycete fungi, and in animals but not in basidiomycete fungi, bacteria, or plants. Phylogenetic analysis suggested that *CPS1* represents a previously undescribed subset of adenylate-forming enzymes that have diverged from certain acyl-CoA ligases, which in bacteria are involved in biosynthesis of nonribosomal peptides or polyketide/peptide hybrids. Disruption of *CPS1* caused reduced virulence of both race T and race O of *C. heterostrophus* on maize, of *Cochliobolus victoriae* on oats, and of *Gibberella zeae* on wheat. These results suggest that *CPS1* functions as a general fungal virulence factor in plant pathogenic ascomycetes.

Fungi cause more plant disease than any other group of microbes; among fungi, the ascomycetes constitute the largest group of pathogens. Nevertheless, among all ascomycetes, plant pathogenicity is rare. Our goal is to understand the molecular distinction between pathogens and saprobes. One hypothesis is that there must be a common set of pathogenicity/virulence determinants, as well as a set of unique factors that, in sum, contribute to pathogenic ability. There are now plenty of examples of the latter, but few of the former beyond the obvious, such as signal transduction pathway components and cell-wall-degrading enzymes. Here, we describe a virulence factor from the maize pathogen, *Cochliobolus heterostrophus*, and show that it is common among ascomycetes.

The genus *Cochliobolus* contains more than 30 species, some of which are saprobes, whereas others are pathogens, usually of grasses (Poaceae; ref. 1). Phylogenetic analysis has revealed that *Cochliobolus* spp. fall into two distinct groups (2). Group 2 contains 18 species, none of which is notable as a pathogen. In contrast, all members of the genus that are known to cause serious crop diseases, e.g., *C. heterostrophus*, *Cochliobolus carbonum*, *Cochliobolus victoriae*, *Cochliobolus sativus*, and *Cochliobolus miyabeanus* are found among the 13 species in group 1. Low levels of nucleotide substitution in internal transcribed spacer (ITS) and *GPD1* sequences among members of group 1 indicate that the highly virulent species underwent sudden radiation (2). This observation suggests that a progenitor within the genus *Cochliobolus* gave rise, over a relatively short period, to a series of distinct biotypes, each distinguished by having unique pathogenic capability to a particular type of plant. This radiation appears to be associated with acquisition of the ability to produce host-specific toxins required for both pathogenesis and host-range determination (3).

Each host-specific toxin is necessary for the development of a particular disease, e.g., T-toxin, a linear polyketide, is required by *C. heterostrophus* for virulence to maize, and victorin, a cyclized pentapeptide, is required by *C. victoriae* for pathogenicity to oats. However, ability to produce toxin is not sufficient to explain pathogenic ability entirely. *C. heterostrophus*, for example, is found in the field as either of two forms: race T, which produces T-toxin and is highly virulent on maize with Texas male sterile (T) cytoplasm, and race O, which does not produce T-toxin but is pathogenic (albeit weakly) on maize with either T cytoplasm or normal (N) cytoplasm (1). On the basis of this observation, it can be hypothesized that the *Cochliobolus* group 1 progenitor was a weakly virulent pathogen possessing genes for general virulence and pathogenicity.

While seeking general pathogenesis factors by random mutagenesis, we identified a *C. heterostrophus* gene (*CPS1*) that is conserved among ascomycete fungi and animals but does not appear to be in basidiomycete fungi, bacteria, or plants. In this paper, we provide genetic evidence that *CPS1* is necessary for normal virulence of *C. heterostrophus* races T and O on maize, *C. victoriae* on oats, and *Gibberella zeae* (*Fusarium graminearum*), a distantly related ascomycete, on wheat; the latter causes wheat head scab, one of the most destructive cereal diseases in North America (4, 5). We also provide phylogenetic evidence that *CPS1* encodes an enzyme belonging to the AMP-binding superfamily and is phylogenetically most closely related to a group of acyl-CoA-ligases in bacteria that are at the N-terminal end of nonribosomal peptide synthetase (*NRPS*) or polyketide synthase (*PKS*) genes, whose products are nonribosomal peptides, polyketides, or polyketide/peptide hybrids. This evolutionary insight is the starting point for predicting function and for biochemical and structural investigation.††

Materials and Methods

Strains, Media, Crosses, and Transformation. Race T strain C4 (*Tox1*$^+$;*MAT1-2*) and race O strain C5 (*Tox1*$^-$;*MAT1-1*) are near-isogenic lines of *C. heterostrophus* (6). R.C4.2696 (*cps1*$^-$;*Tox1*$^+$;*MAT1-2*;*hygB*^{R}) is a C4-derived mutant generated by using restriction enzyme-mediated integration (REMI). Strains 1301-R-45 (*cps1*$^-$;*Tox1*$^-$;*MAT1-1*;*hygB*^{R}) and 1301-R-26 (*cps1*$^-$;*Tox1*$^+$;*MAT1-2*;*hygB*^{R}) are progeny of the cross R.C4.2696 × C5. *C. victoriae* victorin-producing isolate HvW is from a field collection (7). *G. zeae* strain Z3639 (8) is a field isolate from Kansas. Other fungal isolates used in this report are listed in Table 4, which is published as supporting information on the PNAS web site, www.pnas.org. *C. heterostrophus* transformation procedures, described previously (9), were also used for *C. victoriae* and *G. zeae*. Crossing and culture conditions, methods for testing carbon, and nitrogen utilization are described in

Abbreviations: REMI, restriction enzyme-mediated integration; NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession nos. AF332878 (ChCPS1) and AY267334 (G2CPS1)].

§Present address: Diversa Corporation, 4955 Directors Place, San Diego, CA 92121.

‖B.G.T. and O.C.Y. contributed equally to this work.

**To whom correspondence should be addressed. E-mail: bgt1@cornell.edu.

††Sequence data may be found as follows: for *C. albicans* at www-sequence.stanford.edu/group/candida; for *N. crassa* and *M. grisea* at www.genome.wi.mit.edu/annotation/fungi/neurospora and www-genome.wi.mit.edu/annotation/fungi/magnaporthe; for *A. fumigatus* at www.tigr.org; for *C. neoformans* at http://valefor.stanford.edu/group/Cneoformans/overview.html; and for *P. chrysosporium* at www.jgi.doe.gov.

Table 5, which is published as supporting information on the PNAS web site.

Recovery of DNA from the REMI Insertion Site. Plasmids and a detailed description of procedures used for cloning are described in Table 6 and *Supporting Materials and Methods*, which are published as supporting information on the PNAS web site. Cloning was initiated by recovering REMI vector pUCATPH plus flanking DNA from the C4 (race T) *Kpn*I insertion site, by using plasmid rescue (10). Recovered clone p214B7 was used to disrupt *CPS1* in the race T genome to prove that the DNA flanking the REMI insertion site was responsible for the reduced virulence phenotype. To extend sequence to the right of the REMI *Kpn*I insertion site, two successive targeted gene disruption experiments were done in near-isogenic strain C5 (race O) to eliminate possible effects of T-toxin, a known polyketide virulence factor produced by race T (1).

Disruption of Cv*CPS1* in *C. victoriae*. Plasmid p118B14 (Table 6) was linearized with *Bgl*II, transformed into victorin-producing isolate HvW, and transformants selected on 80 μg of hygromycin B/ml. Targeted integration of the plasmid into the *C. victoriae* genome was confirmed by DNA gel blot analysis. Victorin production was examined by HPLC (details available on request from B.G.T.).

Cloning of Gz*CPS1* from *G. zeae* and Targeted Gene Disruption. The *G. zeae* ortholog of Ch*CPS1* was cloned by PCR by using procedures described in *Supporting Materials and Methods* and Table 6. pFgC8-hygB (Table 6) was transformed into Z3936 protoplasts. Targeted disruption of *CPS1* was confirmed by DNA gel blot analysis.

DNA and RNA Manipulations. Fungal genomic DNAs were prepared according to a previously described procedure (11). Plasmid DNA preparations, restriction enzyme digestions, and preparation of DNA gel blots were performed following standard protocols. Methods for DNA–DNA hybridization, DNA sequencing, and cDNA cloning are described in *Supporting Materials and Methods*.

Bioassays. The microbial assay for T-toxin was described (12). Pathogenicity assays for *C. heterostrophus* on corn, *C. victoriae* on oats, and *G. zeae* on wheat are described in *Supporting Materials and Methods*.

Comparative Molecular Analyses. The amino acid sequence of *CPS1* was used to query publicly available and Syngenta proprietary (*G. zeae*, *Gibberella moniliformis*, and *Botrytis cinerea*) databases by TBLASTN and BLASTP (13). BLAST hits of e-10 or lower were aligned and then adjusted manually. Conserved functional domains were identified by RPS-BLAST searches (13) by using adenylation (AMP-binding, pfam00501), phosphopantetheine attachment site (pfam00550), and condensation (pfam00668) consensus sequences. CLUSTALW (14) and T-coffee (15) were used to generate amino acid sequence alignments of the AMP binding domains of the *CPS1* orthologs and other representatives of the adenylation domain superfamily (16). The two alignments were trimmed at the N- and C-terminal ends to include only well-aligned blocks of amino acids. Trimmed alignments were compared for phylogenetic signal by using PAUP 4.0b8 (17) with maximum parsimony (100 replicates with random additions of sequences) and bootstrap analysis (334 pseudoreplicates). Gaps were treated as a "21st amino acid." The T-coffee alignment resulted in a longer most-parsimonious tree with more steps than the tree that resulted from the CLUSTALW alignment. Additionally, the T-coffee-based tree had a higher consistency index and stronger bootstrap support and was therefore used to estimate phylogenetic relationships. The placement of *CPS1* in the phylogeny was used to predict its possible function (18).

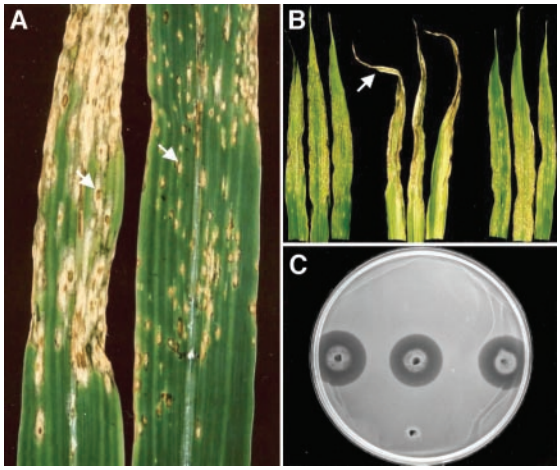

Fig. 1. REMI mutant R.C4.2696 has reduced virulence on corn. (*A*) Virulence test on N-cytoplasm corn. The mutant R.C4.2696 (*Right*) produced smaller lesions (compare those at arrows) than WT strain C4 (*Left*). (*B*) T-cytoplasm corn inoculated with (*Left* to *Right*) mutant R.C4.2696 (*Tox1*$^+$;*cps1*$^-$), WT race T (strain C4, *Tox1*$^+$;*CPS1*$^+$), and a *Tox1*$^+$;*cps1*$^-$ progeny (1301-R-26); symptoms (three leaves per strain) recorded at 5 days. Leaves infected by WT collapsed (arrow); those infected by *Tox1*$^+$;*cps1*$^-$ mutants remained alive despite sustaining small lesions. (*C*) Microbial assay for T-toxin production. Plate containing T-toxin-sensitive *Escherichia coli* cells was inoculated with agar blocks bearing mycelia of the three strains shown in *B* (same order); (*Lower*) a race O (*Tox1*$^-$) control. No significant differences in T-toxin production (indicated by halo sizes) among the three strains were evident.

Results

Characterization of a Weakly Virulent Mutant. REMI mutant R.C4.2696 caused lesions about 60% smaller than those of WT on both N- and T-cytoplasm maize (Fig. 1 *A* and *B*; for quantitative data, see Fig. 8, which is published as supporting information on the PNAS web site). This phenotype was not associated with the *Tox1* locus (Fig. 1*C*) or a reduced fitness trait linked to *Tox1* (19). Note that high virulence of race T on T-cytoplasm maize, which requires production of T-toxin, also depends on the factor altered in mutant R.C4.2696 (Fig. 1 *B* and *C*), i.e., mutant strains produced WT levels of T-toxin but were less virulent than WT. Mutant and WT were identical in growth rate, pigmentation, spore germination, appressorium formation (Fig. 2 *A* and *B*), and conidiation (not shown). Growth tests on various carbon or nitrogen sources showed no differences between the mutant and WT (Table 5). Progeny of crosses between mutant and WT race T or race O segregated 1:1 for parental

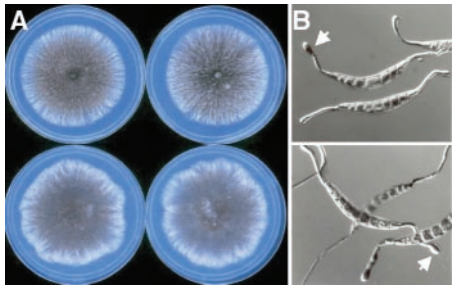

Fig. 2. Mutant R.C4.2696 has normal growth (*A*) and appressorium formation (*B*). (*A*) Plates containing complete medium with xylose (*Upper*) or minimal medium (*Lower*) were inoculated with mutant (*Right*) or WT (*Left*) and incubated in the light at 22°C for 7 days. (*B*) Mutant conidia (*Lower*) germinated and produced appressoria (arrows) like WT (*Upper*). Drops of conidial suspension (10^5/ml in water) were placed on a glass slide in a humid chamber and incubated for 6 h at 32°C.

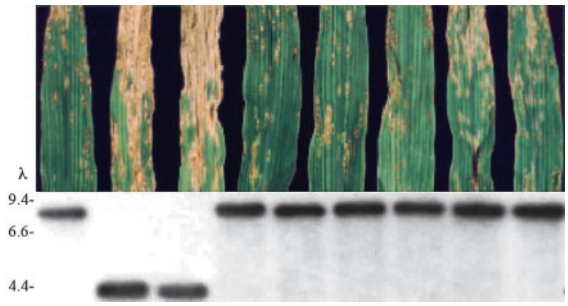

Fig. 3. Targeted disruption of *CPS1* in WT race T restored the original REMI mutant phenotype. (*Upper*) Corn plants (N cytoplasm) were inoculated with (*Left* to *Right*): REMI mutant R.C4.2696; WT race T (C4); WT race O (C5), and five race T disruptants obtained using linearized p214B7. (*Lower*) DNA gel blot showing that a single band (4.2 kb) is present in WT race T and race O (second and third lanes from *Left*), respectively, but is replaced by a 9.3-kb band (increased by the size, 5.1 kb, of pUCATPH) in all strains that showed the mutant phenotype in the plant assay shown (*Upper*). Genomic DNA was digested with *Bgl*II and probed successively with 5′ end and 3′ end flanking DNA fragments carried on p214B7 (strain order is the same as above).

types only: all hygromycin B-resistant progeny produced small lesions similar to those of the mutant parent; all hygromycin B-sensitive progeny produced WT lesions, indicating the mutant phenotype was caused by REMI vector insertion at a single site. This result was confirmed by targeted gene disruption by using plasmid p214B7 (Table 6); homologous integration resulted in the original REMI mutant phenotype (Fig. 3).

Analysis of DNA at the REMI Vector Insertion Site. DNA (11.3 kb) surrounding the REMI insertion site contained a 5.4-kb ORF designated *CPS1* (Fig. 4*A*). The ATG at nucleotide position 826 (GenBank accession no. AF332878) (CACCATGCT) initiates an ORF that starts 521 bp upstream of the *Kpn*I REMI vector insertion site (arrowhead) and ends at nucleotide position 6209. cDNA sequencing confirmed the presence of three introns.

Comparative Analysis of CPS1. BLASTP and TBLASTN searches using the deduced CPS1 protein sequence as query revealed high similarities to hypothetical proteins from fungi, i.e., the saprophytic pyrenomycete *Neurospora crassa*, the human pathogenic plectomycete *Aspergillus fumigatus*, the hemiascomycetes *Candida albican*s, and *Saccharomyces cerevisiae*, and two proteins from the archaeascomycete *Schizosaccharomyces pombe* and from animals, i.e., *Drosophila melanogaster, Caenorhabditis elegans*, *Mus musculus*, and *Homo sapiens* (Table 1; only CPS1 homologues in the GenBank database are shown).

BLAST searches found weaker similarities to the N-terminal acyl Co-A ligase-encoding modules of bacterial NRPSs, including bleomycin synthetase BlmVI from *Streptomyces verticillus* and saframycin Mx1 synthetase SafB from *Myxococcus xanthus* (Table 1). Similarly, N-terminal acyl Co-A ligase-encoding modules were found in bacterial PKSs and NRPS/PKS hybrid genes (data not shown).

The fungal and animal BLAST hits are all proteins of unknown function, although the fungal proteins SpYD22, SpYA84, and ScYor093C were annotated as probable membrane proteins, based on the presence of putative transmembrane domains. In our analysis, however, computational predictions (PredictProtein, Columbia University, http://cubic.bioc.columbia.edu/predictprotein/submit_def.html) and TMHMM (prediction of transmembrane helices in proteins, Technical University of Denmark, http://www.cbs.dtu.dk/services/TMHMM/) did not predict transmembrane helices for CPS1 or any of the three yeast proteins.

Protein domain searches against the Pfam database revealed that CPS1 and all significant BLAST hits contain AMP-binding domains

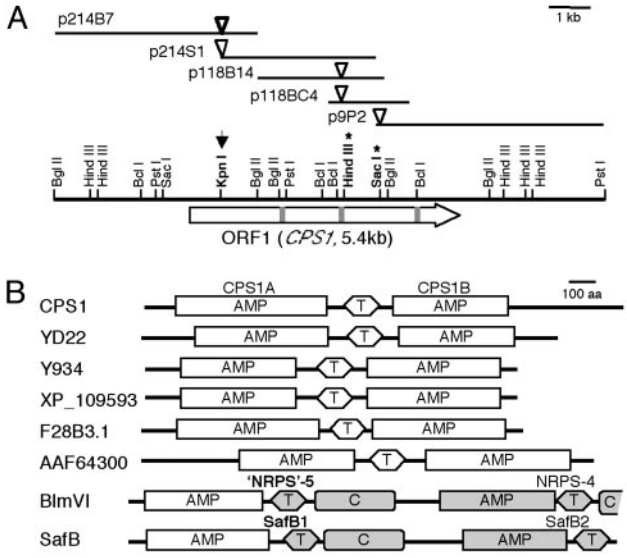

Fig. 4. Identification of *CPS1*. (*A*) Genomic DNA on both flanks of the REMI vector insertion site (solid arrowhead at *Kpn*I site) in the mutant R.C4.2696 genome was recovered by two successive cycles of plasmid rescue. A partial restriction enzyme map of the 11.3-kb region around the tagged site is shown. The asterisks indicate two sites (*Hind*III and *Sac*I) of vector integration, used for subsequent plasmid rescue. Locations of five rescued clones used for sequencing are shown by thin lines (triangle represents the position of pUCATPH in each clone). The ORF (ORF1 = *CPS1*) is indicated by the open arrow. Vertical bars in ORF1 indicate intron positions. (*B*) Domain organization of CPS1 and several of the top BLAST hits listed in Table 1. ORFs are indicated by thick black lines (drawn to scale with a maximum of 2,000 aa; the C terminus of BlmVI is not shown). AMP, AMP-binding; T, thiolation (phosphopantetheine attachment site); C, condensation. CPS1 and all hypothetical proteins have two AMP-binding domains (designated as CPS1A and CPS1B), both of which are similar to the AMP domain of the first modules of BlmVI and SafB.

(Pfam00501; Fig. 4*B* and Fig. 9, which is published as supporting information on the PNAS web site) characteristic of adenylate-forming enzymes that include NRPSs and various CoA ligase/synthetases such as acyl CoA ligase. Two AMP-binding domains (designated CPS1A and CPS1B) are recognizable in CPS1 and in fungal and animal homologs. This modular organization is similar to those of SafB, BlmVI, and other NRPSs (Fig. 4*B*).

Amino acid alignment of the putative AMP-binding domains of CPS1 and its fungal and animal orthologs confirmed the presence of conserved core sequences (cores 1–5) found in AMP domain-containing enzymes (20–22) (Fig. 9). AMP domain cores 3 and 4 are well conserved in CPS1 except that, in core 4 of domain A, there is a leucine instead of an aspartic acid, and in core 4 of domain B, there is a tyrosine instead of an aspartic acid that is present in the animal orthologs. Cores 1, 2, and 5 are weakly conserved compared with the adenylation domain consensus (Fig. 9).

Although database searches did not reveal other functional domains, manual alignment of amino acids between CPS1A and CPS1B suggested that CPS1 has a thiolation (T, also called peptidyl carrier protein) domain that is also found in all NRPSs 3′ of each AMP domain (21, 22) (Fig. 9). This domain is found only after CPS1A, not after CPS1B.

Phylogenetic analysis of AMP-binding domains indicated that CPS1 and the strong BLAST hits form a previously undescribed clade of eukaryotic proteins so far known only in ascomycete fungi and animals (Fig. 5*A* and Fig. 10, which is published as supporting information on the PNAS web site). There are two subclades, one with the CPS1A modules from all orthologous ORFs, and the other with the CPS1B modules, suggesting a duplication before the common ancestor of fungi and animals (Fig. 5*B*). The CPS clade has a common origin with the weaker BLAST hits (Fig. 10), i.e., the

Table 1. Sequence homologues of the predicted ORF1 protein

Protein (GenBank accession no.)	Size, amino acids	Organism	Proposed function*	Identity/similarity/ alignable amino acids	*E* value
YD22 (Q10250)	1,517	*S. pombe*	—	47/63/1395	0.0
YA84 (Q09773)	1,428	*S. pombe*	—	27/46/1403	e-154
Yor093 (NP014736)	1,648	*S. cerevisiae*	—	26/46/1559	e-153
Y934 (Q9Y2E4)	1,352	*H. sapiens*	—	24/43/728	5e-48
(XP_109593)	1,255	*M. musculus*	—	25/43/706	6e-47
(AAF64300)	1,640	*D. melanogaster*	—	22/42/793	3e-46
F28B3.1 (T34061)	1,394	*C. elegans*	—	23/42/731	1e-42
BlmVI (AAG02359)	2,675	*S. verticillus*	ACL	23/36/527	2e-20
SafB (T18551)	1,770	*M. xanthus*	ACL	22/37/642	3e-18

*ACL, acyl CoA ligase; first module of the protein similar to CPS1.

N-terminal modules of several proteins involved in biosynthesis of bacterial NRPSs (e.g., BlmVI and SafB), polyketides (e.g., *Streptomyces coelicolor*, GenBank accession no. CAA16183), and polyketide/peptide hybrids (e.g., *Bacillus subtilis* mycA, GenBank accession no. AAF08795). These N-terminal modules encode acyl CoA ligases. CPS1 in *C. heterostrophus* and other filamentous ascomycetes is extended at the 3′ end relative to orthologs in other organisms (Fig. 4*B*). There is no conservation in this extension except for a high percentage of polar amino acids. The NRPS/PKS-associated CoA ligases and closely related members of the CPS clade form a superclade that is distantly related to previously described members of the AMP-binding domain superfamily (Figs. 5*A* and 10). This latter group divides into two clades identified previously (23); one clade includes the NRPSs and closely related α-aminoadipic acid reductases, and the other includes several groups of CoA ligase/synthetases including acetyl-, phenylacetyl-, long chain fatty acid-, and 4-coumarate-CoA ligases. The sequence of *CPS1* has been deposited in the GenBank database (accession no. AF332878).

Phylogenetic analysis also suggests that *CPS* gene duplication and divergence occurred early in ascomycete evolution, forming *CPS1* and *CPS2* clades (Figs. 5*B* and 10). Only *CPS1* is retained in the filamentous ascomycetes, only *CPS2* is retained in the hemiascomycetes, and both have been retained in the archaeascomycetes, represented here by *S. pombe*. CPS1 is more highly conserved among ascomycetes than is CPS2. For example, SpCPS1 (YD22, deduced protein sequence of 1,517 amino acids) has 47% identity and 63% similarity to the ChCPS1 protein for 1,430 alignable amino acids. In contrast, SpCPS2 (YA84, deduced protein sequence of 1,428 amino acids) has only 27% identity and 46% similarity to the ChCPS1 protein for 1,403 alignable amino acids (Table 1). Overall, these data indicate that *CPS1* is highly conserved between both saprobic and pathogenic ascomycetous fungi and is likely involved in production of a similar molecule in all these fungi. On the basis of these data, we conclude that CPS1 represents a previously undescribed group of eukaryotic adenylate-forming enzymes.

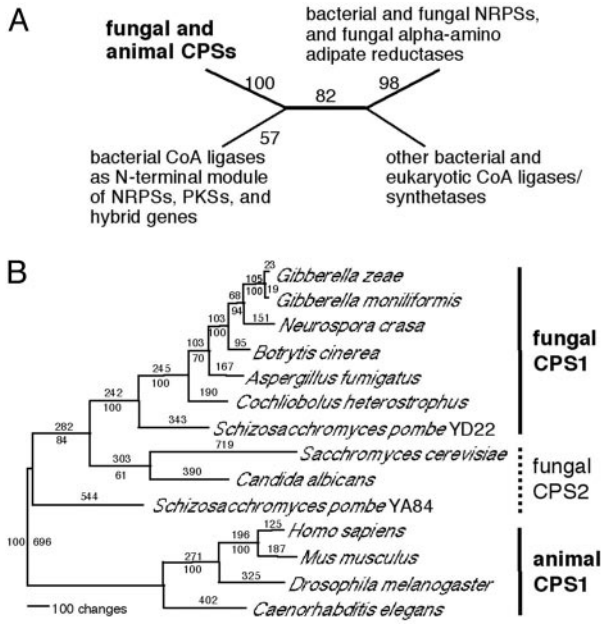

Fig. 5. Phylogenetic analysis of CPS1. (*A*) CPS belongs to the adenylate-forming enzyme superfamily, as shown in this simplified rendering, which highlights the relationships of the major enzyme families. Numbers are percentages of bootstrap support for each clade. See also Fig. 10. (*B*) CPS phylogeny showing an inferred duplication event resulting in CPS1 and CPS2 in ascomycetes. Shown is one most-parsimonious tree: 14 genes, 6,294 steps, 482 excluded characters, and 1,690 included characters, of which 1,244 are informative, 395 are uninformative, and 51 are constant. Rescaled consistency index = 0.653. Numbers above branches are amino acid changes; numbers below branches are bootstrap percentages estimating support for each branch.

Conservation of CPS1 Among Fungi. Genomic DNAs from 31 fungal species, representing 13 ascomycete genera in the loculoascomycetes and the pyrenomycetes, hybridized when probed with *CPS1* (Table 4 and Fig. 11, which is published as supporting information on the PNAS web site). All 16 *Cochliobolus* species, including apparent saprophytes (e.g., *Cochliobolus homomorphus* and *Cochliobolus peregianensis*), known plant pathogens (e.g., *C. carbonum*, *C. victoriae*, *C. miyabeanus*, and *C. sativus*), and representatives of five loculoascomycete genera closely related to *Cochliobolus* (i.e., *Pyrenophora*, *Setosphaeria*, *Bipolaris*, *Stemphylium*, and *Alternaria*) showed hybridization intensities comparable to that seen with *C. heterostrophus* itself (Fig. 11). DNAs of five distantly related pyrenomycete genera (e.g., *Magnaporthe grisea* and *G. zeae*) hybridized weakly to *CPS1* (Fig. 11*B*), whereas no signal was detected in DNA of the basidiomycete *Ustilago maydis* (data not shown). In addition, DNA gel blots (not shown) of the saprobic fungus *N. crassa* hybridized with a *CPS1* probe. BLAST searches confirmed the presence of *CPS1* orthologs in the genomes of *M. grisea*, *G. zeae*, *N. crassa*, *G. moniliformis*, *B. cinerea*, and *A. fumigatus* and absence in the genomes of the basidiomycetes *Phanaerochaete chrysosporium* and *Cryptococcus neoformans*, suggesting that Basidiomycota have lost *CPS1*.

Targeted Disruption of *C. victoriae CPS1*. Cv*CPS1* was disrupted by transformation of a victorin-producing isolate (HvW) with p118B14 (Table 6). Gel blot analysis showed that transformants Tx2, Tx5, Tx7, and Tx8 sustained disruptions of Cv*CPS1*. Ectopic integration of the vector occurred in transformants Tx4 and Tx9.

GENETICS

Fig. 6. Disruption of *C. victoriae CPS1* causes reduced virulence. Inoculations: (*Left* to *Right*) WT (HvW), Tx7, Tx2, Tx9, and water. Oat seedlings inoculated with WT or Tx9 (an ectopic transformant) were completely killed; those inoculated with Tx7 and Tx2 (disruptants) showed a mixture of dead (white arrow) and healthy plants.

Fig. 7. Disruption of *G. zeae CPS1* causes reduced virulence. *Left* to *Right*: WT, Tx4, Tx10, TxH1, Tx11, and water. Wheat heads inoculated with WT or TxH1 (ectopic transformant) were completely bleached, whereas those inoculated with Tx4, Tx10, and Tx11 (disruptants) showed minor symptoms.

Cv*cps1* disruptants (but not transformants with ectopic integrations) had reduced virulence on susceptible oats (Fig. 6 and Table 2), although they produced WT levels of victorin in culture as determined by HPLC analysis (Fig. 12, which is published as supporting information on the PNAS web site). This result is reminiscent of the effect of the *cps1* mutation on virulence of T-toxin-producing *C. heterostrophus* (Fig. 1 *B* and *C*), i.e., victorin production alone in *C. victoriae* is not sufficient for normal pathogenesis on oats; *CPS1* function is needed as well. The same fungal growth tests were done for Cv*cps1*$^-$ mutants as for Ch*cps1*$^-$ mutants (Fig. 2, Table 5), and no differences from WT growth rate were observed (data not shown).

Targeted Disruption of *G. zeae CPS1*. The GzCPS1 deduced protein has 68% identity and 80% similarity to ChCPS1 in a 1,472 alignable amino acid overlap, as well as the same motif sequences and domain organization as ChCPS1. The nucleotide sequence of *GzCPS1* has been deposited in the GenBank database (accession no. AY267334). Gel blot analysis confirmed that *CPS1* disruption occurred in transformants Tx4, Tx10, and Tx11 and ectopic integration occurred in transformants TxB1 and TxH1 (data not shown). Disruption of *G. zeae CPS1* resulted in reduced virulence of the Gz*cps1*$^-$ mutants (Fig. 7 and Table 3). Carbon and nitrogen utilization was the same for WT and mutant, except that the mutant grew more slowly than WT on nitrate-containing medium (not shown).

Discussion

The Role of CPS1 in Pathogenesis. Fungal plant pathogenesis is a complex biological process requiring diverse pathogenicity and virulence factors. These factors tend to fall into two categories: specific and general. Specific factors, e.g., host-specific toxins (1), detoxification enzymes (24), and avirulence proteins (25, 26), are usually restricted to certain species or even to a particular race. Genes controlling these factors are not broadly conserved. In contrast, genes for general factors, e.g., signal transduction pathway components (27), cell-wall-degrading enzymes (28–30), and melanin (31) are widely conserved among species and among large taxonomic groupings. The extensive conservation of *CPS1* and results of targeted gene disruption experiments conducted in both *Cochliobolus* spp. and *G. zeae* argue that the *CPS1*-controlled product may be a general virulence factor.

Unlike the general factors mentioned above, however, the role of *CPS1* in pathogenesis is not obvious. None of the *CPS1* homologs previously deposited in GenBank has a known function. Expression of *S. pombe* YD22 (SPAC56F8.02, an ortholog of *CPS1*) was up-regulated by heat shock and down-regulated by hydrogen peroxide treatment and during meiosis (www.sanger.ac.uk/perl/SPGE/geexview?group = 2&q = SPAC56F8.02). The *S. pombe* YA84 gene (SPAC22F3.04, a paralog of *CPS1*) was up-regulated during meiosis but not affected by stress conditions (www.sanger.ac.uk/perl/SPGE/geexview?q = SPAC22F3.04). *S. cerevisiae* YOR093C (a paralog of *CPS1*) expression was up-regulated by stress conditions such as heat shock or stationary growth (32), but functional analysis showed that the yor093c null mutant is viable and did not show an altered phenotype (33). The human Y934 gene is expressed at low to moderate levels and shows tissue specificity, e.g., moderate in heart and brain and low in lung and pancreas (34).

Table 2. Virulence of *C. victoriae* isolates to oats

	Germination, percent†		Survivors, percent	
Treatment*	Mean	SD	Mean	SD
Uninoculated	74.5	7.5	100.0	0.0
HvW (WT)	20.0	5.0	0.0	0.0
Tx2 (homologous)	44.4	5.9‡	59.0	6.7‡
Tx5 (homologous)	46.1	12.5‡	30.2	4.7‡
Tx7 (homologous)	62.9	5.6‡	62.5	7.3‡
Tx8 (homologous)	47.2	10.8‡	44.6	8.6‡
Tx4 (ectopic)	24.5	6.9	0.0	0.0
Tx9 (ectopic)	13.3	4.4	0.0	0.0

*All isolates were tested on susceptible seeds. Repeat experiments gave similar results.
†Percentage of survivors based on the number of plants visible at 8 days and the number of survivors visible at 24 days. Results are the average of three replicates (each including 60 inoculated seeds) from *t* test.
‡Significantly different from WT at $P < 0.05$.

Table 3. Virulence of *G. zeae* isolates to wheat

	Percentage of infection†			
	10^4		10^5	
Treatment*	Mean	SD	Mean	SD
Uninoculated	0.0	0.0	0.0	0.0
Z3639 (WT)	93.9	8.9	100.0	0.0
Tx4 (homologous)	7.0	8.5‡	41.5	7.9‡
Tx5 (homologous)	5.5	5.2‡	43.8	8.2‡
Tx10 (homologous)	7.8	10.9‡	47.5	14.4‡
Tx11 (homologous)	6.3	8.8‡	36.2	11.6‡
TxH2 (ectopic)	92.3	11.2	100.0	0.0
TxB1 (ectopic)	82.5	14.5	100.0	0.0

*Repeat experiments gave similar results.
†Infection referred to partially or completely bleached spikelets observed on day 7. Spikelets showing minor symptoms (tiny yellow or brown spots) not counted. Results are the average of eight inoculated wheat heads (12–18 spikelets per head) from *t* test.
‡Significantly different from WT at $P < 0.05$.

The *C. elegans* F28B3.1 gene is developmentally regulated and abundant only in adults (35). These data suggest that CPS1 and CPS1-like proteins may function in response to various environmental or developmental cues but actual function depends on the organism involved. For the fungi in our study, CPS1 function is required for normal virulence to plants.

Null mutations in ascomycete *CPS1* genes cause no discernable growth phenotype, indicating this gene is not required for primary metabolism. The yeast gene expression data mentioned above raised the possibility that this group of proteins might play a role in fungal survival under stress conditions. For fungal pathogens, coping with various stresses, e.g., oxidative burst, iron sequestration, and antimicrobial activity imposed by host defense mechanisms, is essential for pathogen ingress (36). It is notable that *cps1*$^-$ mutants of host-specific toxin-producing fungi are reduced in virulence despite their ability to continue producing WT levels of toxin. This observation suggests that these toxins are not sufficient to combat host-generated stresses that are inhibitory to fungal growth *in planta*, and that CPS1 function is required to overcome host defense. The conservation of *CPS1* in both pathogens and saprobes might reflect a general need of all aerobes to tolerate stresses present in natural environments. This would explain why saprobes like *S. pombe* and *N. crassa* maintain copies of *CPS1*.

CPS1 Belongs to a Superfamily of Adenylate-Forming Enzymes. The most prominent feature of the *CPS1*-encoded protein is the presence of AMP-binding domains, diagnostic of a large group of proteins collectively called adenylate-forming enzymes. These enzymes activate carboxylic substrates via an ATP-dependent covalent binding of AMP to form acyl adenylates required for various biochemical reactions (20, 21). Some adenylate-forming enzymes function alone, e.g., insect luciferase, acyl-CoA or aryl-CoA ligases, 4-coumarate-CoA ligase, long chain fatty acid-CoA ligase, acetyl-CoA synthetase and other closely related synthetases, and α-amino adipate reductase. These autonomous proteins are relatively small (400–700 amino acids) and have one module with one AMP-binding domain. In contrast, other adenylate-forming enzymes, e.g., NRPSs, are large multimodular complexes (20–22, 37, 38).

CPS1 has two AMP-binding domains, suggesting similarity to NRPSs. However, phylogenetic analyses distinguish CPS1 from NRPSs, or any other known adenylate-forming enzyme. CPS1 lacks the conserved thiolation and condensation domains characteristic of typical NRPSs (21, 22). Phylogenetic analyses indicate that *CPS1* is a member of a previously uncharacterized clade of eukaryotic genes whose closest relatives are a clade of bacterial acyl CoA ligases found in the N-terminal position of NRPSs, PKSs, and hybrid genes. Establishing the phylogenetic placement of *CPS1* gives us an advantage in understanding possible biochemical function, an ongoing project. Because the enzymatic activities and substrate specificities of the bacterial acyl-CoA ligases have not yet been well characterized, we expect discovery of the biochemical function of CPS1 will be a significant challenge.

Nevertheless, the value of the combined molecular genetic and comparative phylogenomic approach used here for gene annotation cannot be overemphasized. Without the latter, the discovery that CPS1 and CPS1-like proteins comprise a family of previously uncharacterized adenylate-forming enzymes would not have been possible. Although the enzymatic function of CPS1 remains to be identified, genetic tests indicate that CPS1 is important for virulence in at least two independently evolved plant pathogenic fungi (i.e., *Cochliobolus* spp. and *G. zeae*). However, the broad distribution of *CPS1* in all filamentous ascomycetes also suggests that this gene plays a role in some fundamental aspect (e.g., adaptation to stress conditions) of all filamentous ascomycetes, a role that is also necessary as a virulence factor.

We thank colleagues indicated in Table 4 for providing fungal isolates or DNA samples. This work was supported by grants to O.C.Y. and B.G.T. from the U.S. Department of Agriculture. S.K. was supported by Torrey Mesa Research Institute funds to Louise Glass and John Taylor (Plant and Microbial Biology Department, University of California, Berkeley).

1. Yoder, O. C., Macko, V., Wolpert, T. J. & Turgeon, B. G. (1997) in *The Mycota*, eds. Carroll, G. & Tudzynski, P. (Springer, Berlin), Vol. 5, pp. 145–166.
2. Berbee, M. L., Pirseyedi, M. & Hubbard, S. (1999) *Mycologia* **91,** 964–977.
3. Turgeon, B. G. & Lu, S.-W. (2000) in *Fungal Pathology*, ed. Kronstad, J. W. (Kluwer, Dordrecht, The Netherlands), pp. 93–126.
4. Booth, C. (1981) in *Fusarium: Disease, Biology, and Taxonomy*, eds. Nelson, P. E., Toussoun, A. T. & Cook, R. J. (Pennsylvania State Univ. Press, University Park, PA), pp. 446–452.
5. Sutton, J. C. (1982) *Can. J. Plant Pathol.* **4,** 195–209.
6. Leach, J., Lang, B. R. & Yoder, O. C. (1982) *J. Gen. Microbiol.* **128,** 1719–1729.
7. Christiansen, S. K., Wirsel, S., Yun, S.-H., Yoder, O. C. & Turgeon, B. G. (1998) *Mycol. Res.* **102,** 919–929.
8. Bowden, R. L. & Leslie, J. F. (1992) *Exp. Mycol.* **16,** 308–315.
9. Turgeon, B. G., Bohlmann, H., Ciuffetti, L. M., Christiansen, S. K., Yang, G., Schafer, W. & Yoder, O. C. (1993) *Mol. Gen. Genet.* **238,** 270–284.
10. Yang, G., Rose, M. S., Turgeon, B. G. & Yoder, O. C. (1996) *The Plant Cell* **8,** 2139–2150.
11. Yoder, O. C. (1988) in *Genetics of Plant Pathogenic Fungi*, ed. Sidhu, G. S. (Academic, San Diego), Vol. 6, pp. 93–112.
12. Ciuffetti, L. M., Yoder, O. C. & Turgeon, B. G. (1992) *Fungal Genet. Newslett.* **39,** 18–19.
13. Altschul, S. F., Madden, T. L., Schaeffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402.
14. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
15. Notredame, C., Higgins, D. G. & Heringa, J. (2000) *J. Mol. Biol.* **302,** 205–217.
16. Turgay, K. & Marahiel, M. A. (1994) *Pept. Res.* **7,** 238–241.
17. Swofford, D. L., Waddell, P. J., Huelsenbeck, J. P., Foster, P. G., Lewis, P. O. & Rogers, J. S. (2001) *Syst. Biol.* **50,** 525–539.
18. Thornton, J. W. & DeSalle, R. (2000) *Annu. Rev. Genomics Hum. Genet.* **1,** 41–73.
19. Klittich, C. R. J. & Bronson, C. R. (1986) *Phytopathology* **76,** 1294–1298.
20. Stachelhaus, T. & Marahiel, M. A. (1995) *J. Biol. Chem.* **270,** 6163–6169.
21. von Döhren, H., Keller, U., Vater, J. & Zocher, R. (1997) *Chem. Rev.* **97,** 2675–2705.
22. Marahiel, M. A. (1997) *Chem. Biol.* **4,** 561–567.
23. Turgay, K., Krause, M. & Marahiel, M. A. (1992) *Mol. Microbiol.* **6,** 529–546.
24. Bouarab, K., Melton, R., Peart, J., Baulcombe, D. & Osbourn, A. (2002) *Nature* **418,** 889–892.
25. Lauge, R. & DeWit, P. J. G. M. (1998) *Fungal Genet. Biol.* **24,** 285–297.
26. Jia, Y., McAdams, S. A., Bryan, G. T., Hershey, H. P. & Valent, B. (2000) *EMBO J.* **19,** 4004–4014.
27. Xu, J.-R. (2000) *Fungal Genet. Biol.* **31,** 137–152.
28. ten Have, A., Mulder, W., Visser, J. & Van, K. J. A. L. (1998) *Mol. Plant–Microbe Interact.* **11,** 1009–1016.
29. Reignault, P., Kunz, C., Delage, N., Moreau, E., Vedel, R., Hamada, W., Bompeix, G. & Boccara, M. (2000) *Mycol. Res.* **104,** 421–428.
30. Isshiki, A., Akimitsu, K., Yamamoto, M. & Yamamoto, H. (2001) *Mol.–Plant Microbe Interact.* **14,** 749–757.
31. Henson, J. M., Butler, M. J. & Day, A. W. (1999) *Annu. Rev. Phytopathol.* **37,** 447–471.
32. Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D. & Brown, P. O. (2000) *Mol. Biol. Cell* **11,** 4241–4257.
33. Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H., *et al.* (1999) *Science* **285,** 901–906.
34. Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Hirosawa, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. & Ohara, O. (1999) *DNA Res.* **6,** 63–70.
35. Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V. & Kim, S. K. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 218–223.
36. Wolpert, T. J., Dunkle, L. D. & Ciuffetti, L. M. (2002) *Annu. Rev. Phytopathol.* **40,** 251–285.
37. Cane, D. E., Walsh, C. T. & Khosla, C. (1998) *Science* **282,** 63–68.
38. Walsh, C. T., Chen, H., Keating, T. A., Hubbard, B. K., Losey, H. C., Luo, L., Marshall, C. G., Miller, D. A. & Patel, H. M. (2001) *Curr. Opin. Chem. Biol.* **5,** 525–534.