The Plant Journal (2001) **25**(3), 1–11

Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*

Yinong Han[1,†], Xiaoguang Liu[2,†], Ulla Benny[1], H. Corby Kistler[1,*,‡] and Hans D. VanEtten[2]
[1] *Graduate Program in Plant Molecular and Cellular Biology, Plant Pathology Department, University of Florida, Gainesville, FL 32611-0680, USA, and*
[2] *Department of Plant Pathology, University of Arizona, Tucson, AZ 85721, USA*

Received 7 August 2000; revised 13 November 2000; accepted 22 November 2000.
[*]For correspondence (fax +1 651 649 5054; e-mail hckist@puccini.crl.umn.edu).
[†]These authors contributed equally to this work.
[‡]Present address: USDA, ARS Cereal Disease Laboratory and Plant Pathology Department, University of Minnesota, 1551 Lindig St, St Paul, MN 55108, USA.

Summary

Three genes that contribute to the ability of the fungus *Nectria haematococca* to cause disease on pea plants have been identified. These pea pathogenicity (*PEP*) genes are within 25 kb of each other and are located on a supernumerary chromosome. Altogether, the *PEP* gene cluster contains six transcriptional units that are expressed during infection of pea tissue. The biochemical function of only one of the genes is known with certainty. This gene, *PDA1*, encodes a specific cytochrome P450 that confers resistance to pisatin, an antibiotic produced by pea plants. The three new *PEP* genes, in addition to *PDA1*, can independently increase the ability of the fungus to cause lesions on pea when added to an isolate lacking the supernumerary chromosome. Based on predicted amino acid sequences, functions for two of these three genes are hypothesized. The deduced amino acid sequence of another transcribed portion of the *PEP* cluster, as well as four other open reading frames in the cluster, have a high degree of similarity to known fungal transposases. Several of the features of the *PEP* cluster – a cluster of pathogenicity genes, the presence of transposable elements, and differences in codon usage and GC content from other portions of the genome – are shared by pathogenicity islands in pathogenic bacteria of plants and animals.

Keywords: *Pisum sativum*, *Fusarium solani*, pathogenicity island.

Introduction

The identity, genomic organization and transmission of genes controlling plant pathogenicity in fungi is poorly understood compared to our knowledge of these genes in prokaryotic pathogens. In the latter case, a framework of understanding is beginning to emerge in which pathogenicity determinants (e.g. toxins, adhesins, lytic enzymes) have been identified and shown to be exported by common secretion systems under the control of similar regulatory mechanisms (Collmer, 1998; Finlay and Falkow, 1997; Galán and Collmer, 1999; Salmond, 1994). Additionally, the genetic determinants for a number of these pathogenicity factors have been shown to be located on mobile elements such as plasmids or clustered on chromosomal segments known as 'pathogenicity islands' that are probably horizontally transmitted (Groisman and Ochman, 1996; Hacker *et al.*, 1997; Ochman *et al.*, 2000). Whether fungi will be found to have similar clustering and potential for mobility of pathogenicity determinants is unclear, although the arrangement of fungal toxin biosynthetic genes indicates that clustering is present in some instances (Brown *et al.*, 1996; Hohn *et al.*, 1993).

Nectria haematococca is an ascomycetous filamentous fungus pathogenic to a range of plant species including the common pea (*Pisum sativum*). A gene family (*PDA*) involved in pathogenicity to pea has been described for this fungus (VanEtten *et al.*, 1994). The *PDA* gene family encodes cytochrome P450 monooxygenases called pisatin demethylase that demethylate and thus detoxify pisatin, a

pathogen-induced antimicrobial compound ('phytoalexin') produced by pea. Site-directed disruption of the *PDA* gene in strains that contain only one member of this gene family (*PDA*1) increases the sensitivity of the fungus to pisatin and reduces the ability of the pathogen to cause necrotic lesions on pea, but does not render the gene-disrupted strains completely non-pathogenic (Wasmann and VanEtten, 1996). The observation that gene-disrupted strains were only reduced in pathogenicity rather than becoming non-pathogenic was unexpected, because previous conventional genetic studies (Kistler and VanEtten, 1984a; Kistler and VanEtten, 1984b) had indicated that *PDA1* was inherited as a single gene and was an absolute requirement for any pathogenicity on pea. However, the discovery that *PDA1* was on a 1.6 million base pair (Mb) supernumerary chromosome (Miao *et al.*, 1991; Wasmann and VanEtten, 1996) gave an explanation for the apparent discrepancy between gene disruption experiments and conventional genetic studies. The apparent Mendelian inheritance of pea pathogenicity and the absolute requirement of *PDA1* for pathogenicity could be attributed to the presence of other pea pathogenicity genes (*PEP* genes) on the *PDA1* supernumerary chromosome, and the possibility that this entire chromosome was transmitted as a single genetic entity in conventional genetic crosses (VanEtten *et al.*, 1994). Support for this hypothesis and the chromosomal location of these *PEP* genes was obtained by deletion analysis of the chromosome through telomere-mediated chromosome breakage (Kistler *et al.*, 1996). These experiments indicated that *PEP* genes were located somewhere in a 100 kb region between *PDA1* and the telomere. In this paper we show that *PEP* genes are clustered around *PDA1*, and provide evidence to suggest that this cluster may be analogous to bacterial pathogenicity islands.

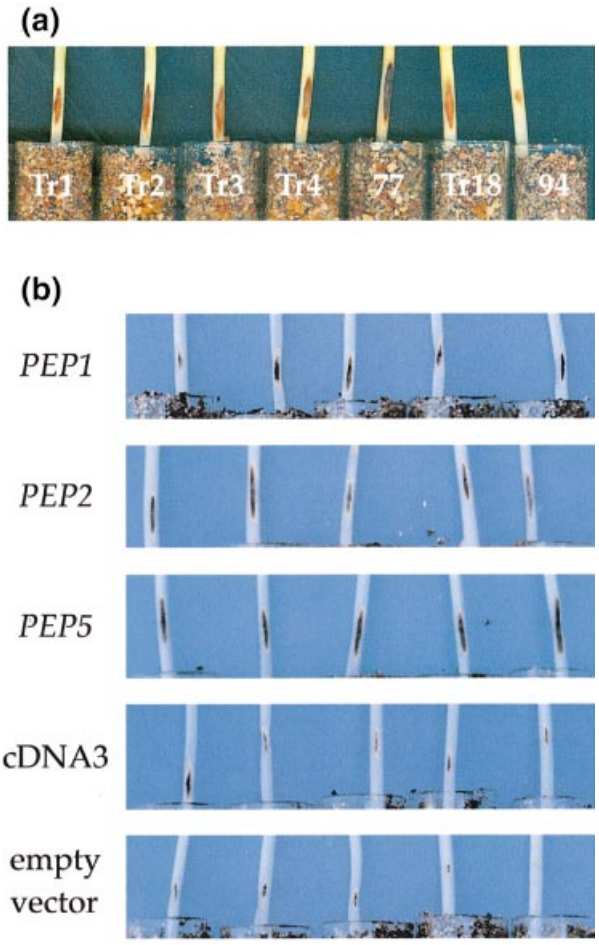

Figure 1.
(a) Effects on virulence of addition of cosmid 55-D-8 to recipient strain 94-6-1. Representative lesions caused by four independent transformants of 94-6-1 with 55-D-8 (labeled Tr1, Tr2, Tr3, Tr4) compared to lesions caused by 77-13-7 (77), Tr18.5 (Tr18) and 94-6-1 (94).
(b) Effect of addition of genomic sequences corresponding to *PEP1*, *PEP2*, *PEP5* and cDNA3 on lesion-causing ability of recipient strain 171-3. Representative lesions are shown 6 days after inoculation.

Results

Evidence for a pathogenicity gene cluster

To test if regions of the 1.6 Mb supernumerary chromosome immediately adjacent to *PDA*1 contained pathogenicity determinants, a cosmid library of *N. haematococca* strain Tr18.5 was constructed. Strain Tr18.5 was derived from the virulent strain 77-13-7, but has a disrupted *PDA*1 gene (Wasmann and VanEtten, 1996). A cosmid (55-D-8) that hybridized to a cloned fragment of *PDA1*, and that contained the disrupted *PDA1* locus along with flanking regions, was identified and used to transform 94-6-1, a non-pathogenic *N. haematococca* strain that lacks the 1.6 Mb chromosome. All four transformants recovered were capable of causing lesions on pea of similar size to those of isolate Tr18.5 (Figure 1a). The average lesion size for 94-6-1 was 1.7 ± 0.26 mm, whereas lesion sizes for transformants were significantly higher, ranging from 5.9 to 7.2 mm (Table 1). Lesion size caused by Tr18.5 was 6.9 ± 1.5 mm, while the wild-type strain 77-13-7 resulted in a lesion size of 11.1 ± 1.6 mm. Another strain (171-3) lacking the 1.6 Mb chromosome was also transformed with 55-D-8, with similar results (Table 1), and this strain was chosen for further studies because of its greater ease of transformation. Strain 171-3, however, produced slightly larger lesions than 94-6-1 and, based on our standard pea stem assay (VanEtten *et al.*, 1980), is classified as having 'low virulence' rather than as non-pathogenic. The lesion size caused by transformed 171-3 containing the vector pMOcosX was 3.9 ± 1.2 mm, whereas the four transformants of 171-3 with 55-D-8 caused lesions ranging in

Table 1. Tests of cosmid 55-D-8 and subclones for their ability to increase lesion size of *Nectria haematococca* strains 94-6-1 or 171-3

Strain/ transformant	Lesion length (number of tests)[a]
171-3	3.3 ± 1.12 (8)
94-6-1	1.7 ± 0.26 (6)
77-13-7	11.1 ± 1.64 (6)
Tr18.5	6.9 ± 1.50 (6)
94-6-1 + cosmid 55-D-8	
a	**6.6 ± 1.27 (6)**
b	**5.9 ± 1.80 (6)**
c	**7.2 ± 0.91 (6)**
d	**6.0 ± 1.61 (6)**
171-3 + pMOcosX	3.9 ± 1.21 (6)
171-3 + cosmid 55-D-8	
a	**6.4 ± 2.53 (6)**
b	6.2 ± 3.80 (6)
c	**9.6 ± 0.41 (6)**
d	5.6 ± 1.99 (6)
171-3 + vector (pCB1004) [2]	
a	4.0 ± 1.00 (7)
b	3.8 ± 0.97 (8)
c-f	4.5 ± 1.13 (6)
171-3 + cDNA1 genomic sequence	
a	3.4 ± 0.78 (3)
b	**6.2 ± 1.47 (13)** [3]
c	3.9 ± 1.20 (2)
d	**5.2 ± 0.69 (6)**
e	3.6 ± 1.22 (4)
171-3 + cDNA2 genomic sequence	
a	3.9 ± 1.34 (2)
b	3.8 ± 1.60 (3)
c	3.9 ± 1.25 (3)
d	4.3 ± 1.01 (8)
e	3.1 ± 1.27 (3)
f	**5.8 ± 1.22 (9)**
g	3.7 ± 0.49 (2)
h	5.5 ± 1.15 (3)
i	5.2 ± 0.50 (4)[4]
j	4.6 ± 2.40 (2)
k	4.5 ± 0.21 (2)
l	**6.7 ± 1.40 (5)**
m	4.1 ± 0.78 (2)
n	3.3 ± 1.34 (2)
o	**5.6 ± 2.21 (3)**
p	**6.3 ± 1.00 (12)**
q	3.9 ± 0.12 (2)
r	**6.3 ± 0.55 (4)**
s	4.1 ± 0.28 (2)
t	4.9 ± 1.13 (2)
u	4.2 ± 1.70 (2)
171-3 + cDNA3 genomic sequence	
a	3.3 ± 1.13 (2)
b	4.4 ± 1.06 (2)
c	4.8 ± 0.78 (2)
d	4.4 ± 0.35 (2)
e	4.3 ± 0.21 (2)
f	2.5 ± 0.14 (2)
g	3.9 ± 0.57 (2)
h	4.8 ± 1.63 (2)
171-3 + cDNA4 genomic sequence	
a	4.3 ± 0.87 (2)
b	3.3 ± 0.87 (3)
c	4.2 ± 0.21 (3)
d	2.9 ± 0.71 (2)
e	3.3 ± 0.14 (2)
f	5.8 ± 1.58 (10)[4]
g	3.0 ± 0.00 (2)
h	4.1 ± 0.00 (2)
i	4.9 ± 0.35 (2)
171-3 + cDNA5 genomic sequence	
a	5.4 ± 0.28 (2)
b	4.5 ± 0.85 (2)
c	**5.1 ± 1.34 (2)**
d	**5.7 ± 0.48 (3)**
e	**6.0 ± 0.73 (4)**
f	4.6 ± 1.12 (4)
g	3.8 ± 0.57 (2)
h	3.1 ± 1.13 (2)
i	5.2 ± 0.14 (2)
j	3.4 ± 0.64 (2)
k	4.7 ± 1.48 (3)
l	4.2 ± 0.64 (2)
m	4.2 ± 0.71 (2)
n	4.1 ± 1.34 (2)
o	5.6 ± 0.85 (2)
p	**5.3 ± 1.53 (4)**
q	5.0 ± 1.49 (2)
r	**6.9 ± 1.61 (4)**
s	3.9 ± 0.42 (2)
t	5.1 ± 0.42 (2)

[a]Mean and variance of lesion length in mm from independent pathogenicity tests. The number of independent pathogenicity tests is given in parentheses. Values for each pathogenicity test consisted of the mean lesion length from seven to 10 plants.
[b]Several independent transformants of 171-3 with pCB1004 were evaluated for virulence. The two transformants (a and b) used for pairwise comparisons with other strains were those showing the largest lesion length in the initial screen. Each strain represents an independently isolated transformant with the vector indicated.
[c]Values in bold are mean lesion lengths significantly larger ($P \leq 0.05$) than paired 171-3 + pMOcosX controls for transformants with 55-D-8. For transformants with individual genes, means in bold were significantly larger than 171-3 + pCB1004 controls in either all or a majority of pairwise comparisons.
[d]These transformants showed significant difference from the control in one or more pathogenicity tests, but the results could not be consistently repeated. We attribute these differences to type I statistical error.

size from 5.6 to 9.6 mm (Table 1). Perhaps because of the higher baseline virulence of 171-3, only two of these four transformants produced significantly larger lesions at

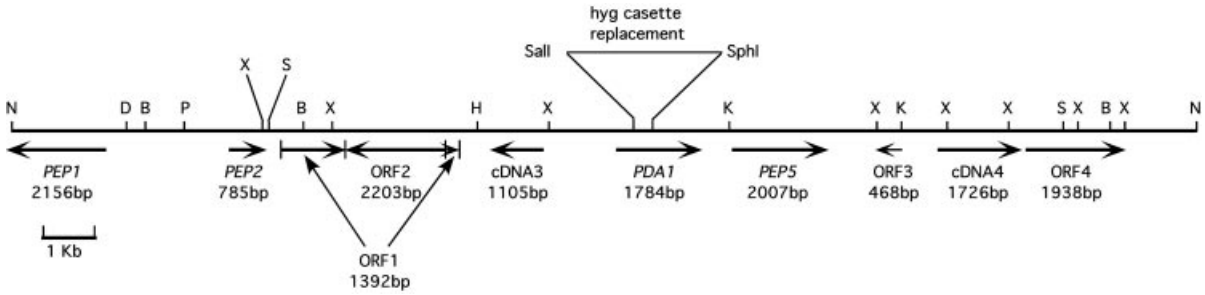

Figure 2. Physical map of cosmid 55-D-8 which contains the *PEP* cluster of *Nectria haematococca*.
Restriction sites: B, *Bgl*II; D, *Dra*I; H, *Hin*dIII; K, *Kpn*I; N, *Not*I (this restriction site is in the polylinker of the pMOcosX vector); P, *Pst*I; S, *Sac*I; X, *Xho*I. The only restriction sites shown for D, H, K, P and S are the sites used to clone genomic fragments corresponding to the cDNAs. ORF1 is interrupted by ORF2. Also shown is the ≈13 kb *Sal*I/*Sph*I fragment which contains DNA from the vector used to disrupt the *PDA*1 locus. Approximately 2.5 copies of this vector were integrated into this site. The insertion is not drawn in proportion to the rest of the figure.

$P \leqslant 0.05$, with the other two showing significance at the $P = 0.1$ level. Taken together, these results strongly imply that genes involved in pea pathogenicity are contained on cosmid 55-D-8, as the only other known gene conditioning pathogenicity, *PDA1*, is not functional in this clone.

To identify transcripts of other potential *PEP* gene(s) physically linked to *PDA1*, a cDNA library was constructed using mRNA from infected pea tissue collected 2 days after inoculation with 77-13-7. Restriction fragments representing all portions of the cosmid 55-D-8 insert were used as hybridization probes to screen the cDNA library, and clones corresponding to five independent transcripts (cDNA1 to cDNA5) were identified. The positions of these transcripts were mapped on the cosmid (Figure 2) and restriction sites flanking the transcriptional units were identified. Genomic fragments corresponding to the five cDNAs were subcloned in a plasmid (pCB1004) containing a hygromycin-resistance cassette (Sweigard *et al*., 1997).

The virulence of transformants containing a single genomic subclone was compared to virulence of controls transformed with the vector pCB1004. Some of the transformants containing genomic subclones corresponding to cDNA1, cDNA2 and cDNA5 were significantly increased in their ability to cause lesions on pea epicotyls, while none of the transformants containing genomic copies of cDNA3 or cDNA4 had such an effect (Table 1; Figure 1b). Because virulence increased only incrementally for single gene transfer, numerous repeated pathogenicity tests were conducted to assure the statistical significance of the results. Hereafter the sequences present in the genomic subclones corresponding to cDNA1, cDNA2 and cDNA5 will be referred to as genes and called *PEP1*, *PEP2* and *PEP5*, respectively. None of the genomic equivalents of the five cDNAs tested restored 171-3, a methionine auxotroph, to methionine prototrophy.

The DNA sequences of cDNAs and corresponding genomic sequences of *PEP1*, *PEP2*, cDNA3 and *PEP5* (GenBank accession numbers AF294788, AF315804, AF315805, AF315806, AF315808and AF315315) revealed the following. The *PEP1* cDNA is 2151 bp and potentially encodes two polypeptides, 330 and 320 amino acids in length and separated by 56 bp starting at position 1117 of the cDNA. These two ORFs are in different reading frames, suggesting the cDNA could encode two alternative functional products. The gene contains a 65 bp intron between the nucleotides corresponding to positions 1390 and 1391 of the cDNA sequence. The *PEP*2 cDNA is 810 bp and contains a single major ORF encoding a product of 233 amino acids. A 63 bp intron is located between positions corresponding to nucleotides 178 and 179 of the cDNA. The gene corresponding to the 1105 bp cDNA3 lacks introns and has no ORF longer than 195 bp. The 2008 bp sequence of the *PEP5* cDNA potentially encodes a 238 amino acid polypeptide. Three introns, 45, 57 and 57 bp, are located in the genomic region corresponding to cDNA5, respectively, at positions 138, 192 and 1161 of the cDNA.

Of the three *PEP* genes, only *PEP5* has significant deduced amino acid sequence similarity to accessions in the DNA sequence databases based on BLASTX and BLASTP searches conducted June 2000. *PEP5* appears to be related to members of the major facilitator superfamily (Pao *et al*., 1998) that includes proton-dependent multi-drug export systems (Figure 3). Of sequences with known function, *PEP5* is most closely related to the *Bacillus subtilis* gene *bmr3* which confers resistance to purimycin, tosufloxacin, ethidium bromide and other potentially inhibitory substances (Ohki and Murata, 1997). Also related to these membrane-associated transporters are toxin efflux systems which apparently allow filamentous fungi that produce toxic secondary metabolites to avoid self-inhibition by these compounds. Genes for these toxin transporters are associated with the gene clusters involved in toxin biosynthesis for cercosporin in *Cercospora*

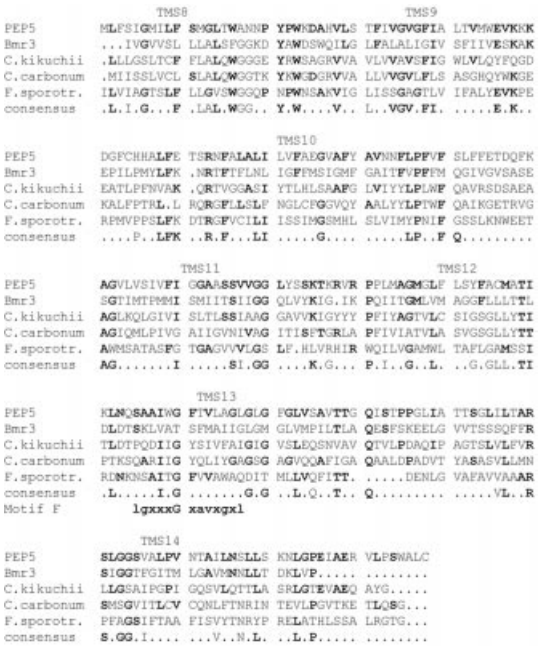

Figure 3. Multiple alignment and majority consensus of amino acid sequences.

Alignment and consensus of the amino acid sequences of *PEP*5, the multi-drug resistance gene *Bmr3* from *Bacillus subtilis* (accession number D50098) and toxin transporter from the fungi *Cercospora kikuchii* (accession number AF091042), *Cochliobolus carbonum* (accession number L48797) and *Fusarium sporotrichioides* (accession number AF011355). Amino acid identities between *PEP5* and other polypeptides are in bold type; the majority consensus amino acid sequence is shown below. Membrane-spanning domains predicted using TMpred (Hofmann and Stoffel, 1993) are overlined and numbered according to correspondence to the C terminus of 14 TMS family proteins. Motif F, found in the TMS-13 of the other 14 TMS family members (Paulsen *et al.*, 1996), is indicated below the consensus sequence.

kikuchii (Callahan *et al.*, 1999), trichothecenes in *Fusarium sporotrichioides* (Alexander *et al.*, 1999), and HC-toxin in *Cochliobolus carbonum* (Pitkin *et al.*, 1996). Hydrophobicity analysis of the *PEP5* encoded polypeptide using TMpred (Hofmann and Stoffel, 1993) predicts seven membrane-spanning domains.

The PEP cluster contains DNA sequences related to fungal mobile elements

In addition to the *PEP* genes and cDNA3, the chromosomal segment contained on cosmid 55-D-8 encompasses five regions with significant sequence similarity to different class II transposons from other filamentous fungi. Between *PEP2* and the gene corresponding to cDNA3 is a region (ORF1, Figure 2) containing similarity to the *Pot*3 transposase of *Magnaporthe grisea* (Farman *et al.*, 1996). Amino acid sequences derived from ORF1 contain 48% identity (244/513 amino acids) and 65% similarity (331/513) to the *Pot3* transposase, but with three frameshift mutations. The *Pot3* similarity is also interrupted by a 2204 bp region (ORF2, Figure 2) which contains a transposon of the *Nht*1 family from *N. haematococca*. *Nht*1 transposable elements are located predominantly on a 1.6 Mb supernumerary chromosome in some strains of *N. haematococca* (Enkerli *et al.*, 1997). The derived amino acid sequence of the *Nht*1-like transposase here is 65% identical (357/549 amino acids) and 77% similar (427/549 amino acids) to *Nht*1, and includes two presumed introns at the same positions as the introns in *Nht1*. The copy on cosmid 55-D-8 is defective because the predicted transposase reading frame contains two mutations that would lead to premature termination of translation. The *Nht*1-like element is framed by imperfect terminal inverted repeats consisting of a 39 or 40 bp unique region plus copies of an imperfect 17 bp direct repeat [TACCCCTCAC(C/T)CGGTCA]. Two copies of the direct repeat are external to the segment containing the 39 bp unique region on one side of the *Nht1* homologue, and three copies of the direct repeat are external on the side containing the 40 bp region.

Between *PEP5* and the gene corresponding to cDNA4 is a 1119 bp pseudogene (ORF3, Figure 2) with a derived amino acid sequence 32% identical (118/362 amino acids) and 44% similar (163/362 amino acids) to the *Tc1-mariner*-like transposase of the *Fusarium oxysporum* transposon *impala* (Langin *et al.*, 1995). To the right of the gene corresponding to cDNA4 is a 1938 bp open reading frame (ORF4, Figure 2) with a derived amino acid sequence (discounting a stop codon at position 1423) 32% identical (219/674 amino acids) and 50% similar (345/674 amino acids) to the transposase of the *restless* transposon of *Tolypocladium inflatum* (Kempken and Kück, 1996). Finally, the gene corresponding to cDNA4 itself has 29% amino acid identity (90/341 amino acids) and 45% sequence similarity (153/341 amino acids) to the transposase of element *Fcc1* found within the *TOX2* locus of *Cochliobolus carbonum* (Panaccione *et al.*, 1996). Clearly, the *PEP* cluster of *N. haematococca* is associated with a considerable number of putative mobile elements and/or remnants of such elements and associated repeats.

Similarity between PEP *gene cluster and bacterial pathogenicity islands*

Pathogenicity islands are clusters of genes involved in pathogenicity, and are further defined by their genetic instability, discontinuous distribution within a bacterial species, proximity to mobile genetic elements and associated repeated sequences, and dissimilarity to other regions of the chromosome based on $G + C$ content (Hacker *et al.*, 1997). The *PEP* gene cluster of *N. haematococca* has many characteristics of a bacterial pathogenicity island. The instability of the supernumerary chromosome containing the *PDA1* locus has been previously established

during genetic crosses and by loss during manipulation of cultures for transformation experiments (Wasmann and VanEtten, 1996). As determined by Southern hybridization, all genes of the *PEP* cluster are found in all field isolates examined of *N. haematococca* virulent on pea, while portions of the cluster are missing in all field isolates of *N. haematococca* non-pathogenic to pea (E. Temporini and H. VanEtten, unpublished results).

Like bacterial pathogenicity islands, the *PEP* gene cluster of *N. haematococca* appears distinct from other regions of the genome with respect to percentage G + C and codon usage bias. To compare G + C content, the nucleotide sequence of major open reading frames of cDNA1 to cDNA5 and the *PDA1* gene were compared to coding sequences for eight *N. haematococca* genes not associated with the *PEP* gene cluster. Mean G + C content was significantly different ($P < 0.001$) between the two groups of genes, with virtually no overlap. The mean G + C content of the *PEP* cluster genes was 50.4% (range 47.1–53.4%), whereas the mean G + C content of the other genes (cutinase, dynein, *ERG*1, *ERG*3, *PEL*A, *PEL*B, *PEL*C, *PEL*D) was 57.5% (range 53.3–60.7%). In addition, using the CORRESPOND program from the sequence analysis software package GCG (Genetic Computer Group, Inc., Madison, WI, USA), codon usage frequencies of the two sets of genes were compared and the pattern of codon usage was found to be different ($D^2 = 0.78$). It should be emphasized, however, that while these differences were found, the absolute number of genes available for comparison was small so that sampling error cannot be ruled out as a major contributing factor to the differences.

Discussion

For several years the pisatin demethylase genes of *N. haematococca* MPVI have been recognized as pea-specific pathogenicity determinants (Ciuffetti and VanEtten, 1996; Schäfer *et al.*, 1989; Wasmann and VanEtten, 1996). We report here that several physically linked genes can also influence the outcome of this disease interaction. Like the pisatin demethylase gene (Ciuffetti and VanEtten, 1996), addition of *PEP1*, *PEP2* or *PEP5* can result in a small but significant increase in the size of necrotic lesions on pea epicotyls caused by some transformants. However, many transformants carrying *PEP1, PEP2* or *PEP5* were not increased in virulence. Because only certain genes from the *PEP* cluster could confer increased virulence, and because lesion sizes were compared to those produced by transformants with the vector pCB1004 alone, the change in virulence is probably attributable to the *PEP* gene rather than a non-specific effect associated with transformation. Similar results were obtained previously on addition of a *PDA* gene to non-pathogenic strains, as a significant increase in virulence occurred in only ≈35% of the transformants (Ciuffetti and VanEtten, 1996), even though disruption of this gene always decreases the virulence of a pathogenic isolate (Wasmann and VanEtten, 1996).

It is puzzling why only some of the *PDA, PEP1, PEP2 and PEP5* transformants show an increase in virulence. Results with other fungal plant pathogens (e.g. Romeis *et al.*, 1997) or plants themselves have also shown that not all transformants show the expected change in phenotype with introduction of a new gene. The lack of uniform change is often attributed to events which result in unreliable expression of the introduced gene. However, some *PDA* transformants which did not show an increase in virulence had enzymatic activity equivalent to wild-type, virulent strains, at least when measured *in vitro* (Ciuffetti and VanEtten, 1996). Perhaps the lack of a consistent increase in virulence when *PDA* or *PEP* genes are introduced into non-pathogenic strains is because these genes are normally clustered and may require other members of the cluster to consistently function in virulence. By specifically disrupting each gene of the cluster, we will be able to test whether those genes that show no effect in gene-addition experiments (e.g. cDNA3) may also be shown to contribute to virulence.

Of the three newly described pathogenicity genes presented here, only the *PEP5* gene has sequence similarity to any gene with known function. *PEP5* shows similarity to transport systems in family two (Pao *et al.*, 1998) of the major facilitator superfamily, called family one by Paulsen *et al.* (1996). Major facilitator transporters are single-polypeptide carriers capable of transporting small solutes driven by chemiosmotic ion gradients. Members of family two are known to function in drug efflux and normally have 14 transmembrane-spanning (TMS) domains (Pao *et al.*, 1998). However, the predicted *PEP5* protein is only about half the size of authentic members of this family, containing only seven predicted TMS domains that correspond to the carboxyl terminus of these proteins. As such, it lacks the 13 amino acid signature motif A found between TMS-2 and TMS-3 in most members of the superfamily (Pao *et al.*, 1998), and also lacks other less conserved motifs found in the amino half of 14 TMS proton antiporters (Paulsen *et al.*, 1996). However, it has retained the weakly conserved motif F, found in the predicted membrane-spanning region corresponding to TMS-13 (Figure 3) as it is found in other members of the 14-TMS family. While the functional capabilities of the *PEP*5 protein have yet to be tested, it is interesting to note that a gene with similarity to a drug resistance system is immediately adjacent to a gene (*PDA1*) known to be involved in overcoming toxicity of the host phytoalexin, pisatin. It is also intriguing that an apparent ATP-dependent pump recently has been reported to be a pathogenicity factor in the rice blast fungus *M. grisea* (Urban *et al.*, 1999).

While no sequence similarities were detected by BLASTP using the derived polypeptide sequence of *PEP2*, a more refined PSI-BLAST search (Altschul *et al.*, 1997) of DNA databases, identified an RNP-1 motif (Burd and Dreyfuss, 1994) corresponding to the conserved RNA-binding domain of polyadenylate-binding proteins and other RNA-binding proteins (Y. Han and H.C. Kistler, unpublished results). The predicted 27 kDa *PEP2* polypeptide has now been expressed *in vitro* and is being tested for RNA-binding and biological activity.

The concept of pathogenicity islands has changed views on how pathogens evolve and acquire disease determinants (Ochman *et al.*, 2000). Rather than slowly accumulating new genetic capabilities by mutation or genomic rearrangement and selection, disease-causing capability may develop rapidly by lateral transfer of well adapted pathogenicity genes arranged on DNA cassettes. This 'evolution by quantum leaps' explains the rapid emergence of troublesome strains of normally innocuous bacteria such as enteropathogenic and uropathogenic *Escherichia coli* (Groisman and Ochman, 1996; Mecsas and Strauss, 1996). Lateral transfer of pathogenicity islands has also been invoked to explain the curious discontinuous distribution of pathogenicity genes among naturally occurring *E. coli* strains, and the observation that strains causing the same symptoms do not form a monophyletic grouping (Boyd and Hartl, 1998; Hacker *et al.*, 1997; Whittam *et al.*, 1993). Lateral transfer might also explain why fungal strains that cause disease on the same host are not monophyletic (O'Donnell *et al.*, 1998).

Horizontal transfer of DNA between organisms has even been proposed as the driving force in the formation of gene clusters (Rosewich and Kistler, 2000). According to the 'selfish operon' hypothesis (Lawrence and Roth, 1996), horizontal transfer accelerates gene clustering because genetic rearrangements that bring genes with co-operating products closer together increase the chance that the genes will be co-mobilized. If the co-operating genes are conditionally non-essential but have adaptive value for colonizing certain ecological niches, then the nascent cluster can be maintained by positive selection in those environments (Lawrence, 1997). The model also proposes that following horizontal transfer, introgressed DNA containing the loosely clustered genes will be foreign to the host, and will not be essential for growth of the recipient cell. Therefore the DNA between the co-operating genes with adaptive significance can be subject to spontaneous deletion, ultimately bringing the loosely clustered genes into closer proximity. We speculate that the processes of horizontal transfer, deletion and conditional positive selection for pea pathogenicity may have allowed for the formation of the *PEP* gene cluster.

The observation that pathogenicity genes in *N. haematococca* may be organized as a pathogenicity island may also explain major questions concerning the acquisition of pathogenicity by fungal pathogens. For example, all strains of *N. haematococca* pathogenic to pea contain *PDA* genes, while strains non-pathogenic to pea often lack *PDA* homologues. Additionally, pea pathogenic strains of a different broad host-range phytopathogenic fungal species, *F. oxysporum*, also have pisatin demethylase genes with remarkable DNA sequence similarity to those from *N. haematococca* (McCluskey and VanEtten, 1995; C. Wasmann, unpublished results); *F. oxysporum* strains from other hosts usually lack *PDA*1 homologues (data not shown). The fact that the genes correlate to host range rather than to phylogenetic relationships for these fungi may be explained by lateral gene transfer, as has been observed for bacterial pathogenicity islands.

While lateral transfer might explain the clustering, acquisition and distribution of pea pathogenicity genes in *N. haematococca* and perhaps *F. oxysporum*, the potential mechanism for transfer is unclear. The *N. haematococca PEP* cluster is associated with sequences related to transposable elements, but these features may explain mobilization only within the genome of the resident strain. In this regard, it is of interest to note that the *PEP* cluster for the strain examined here is located on a supernumerary chromosome. Supernumerary chromosomes have been shown to have the capacity for transfer between otherwise genetically isolated strains of a plant pathogenic fungus (He *et al.*, 1998).

The 1.6 Mb supernumerary chromosome containing the *PEP* gene cluster may also encode other potentially adaptive characteristics for pathogenicity to pea. It has been observed that pea root exudates contain the somewhat unusual amino acid homoserine, and the ability of *N. haematococca* to utilize homoserine as a sole carbon and nitrogen source appears to be encoded by the same supernumerary chromosome that contains the *PEP* gene cluster (Rodriguez and H.D. VanEtten, unpublished results). Thus transfer of this chromosome to a strain of *N. haematococca* non-pathogenic to pea could potentially impart several distinct adaptive features enhancing the recipient strain's ability to cause disease on its host. If host specificities can be transferred horizontally between lineages of this fungal pathogen, strains pathogenic to a given host may emerge rapidly in a genetic background pre-adapted for fitness in any habitat currently occupied by *N. haematococca*.

Q1

Experimental procedures

Fungal strains

Isolates of *Nectria haematococca* Berk. & Br., mating population MVI (anamorph *Fusarium solani*) were grown as described previously (VanEtten *et al.*, 1980). The strain 77-13-7 contains a

single gene for pisatin demethylase (*PDA1*) and is virulent toward pea (Kistler and VanEtten, 1984a; Kistler and VanEtten, 1984b). Tr18.5, which is derived from 77 to 13-7, has the *PDA1* locus disrupted by several copies of a transformation vector containing a hygromycin B resistance gene (*hph*) (Wasmann and VanEtten, 1996). Tr18.5 is virulent toward pea, and construction of a genomic library from Tr18.5 DNA allows for potential pathogenicity genes to be directly determined by a virulence assay because the only other known pathogenicity gene, *PDA1*, is not functional in this isolate.

The fungal strain 94-6-1 (Kistler and VanEtten, 1984a), used as a recipient for fungal transformation, is not pathogenic on pea (Ciuffetti and VanEtten, 1996). Recipient strain 171-3 (American Type Culture Collection 90516) was derived through a series of crosses (Tegtmeier, 1984; P. Matthews and H.D. VanEtten, unpublished results) with the purpose of developing sexually fertile, genetically marked strains deficient in the production of naphtharazin toxins. This strain lacks the 1.6 Mb chromosome containing *PDA1*, and is a methionine auxotroph.

Construction and screening of the cosmid library

For DNA extraction, fungal cultures were grown for 3–5 days in ≈75 ml potato dextrose broth (24 g l^{-1}, PDB, Difco Laboratories, Detroit, MI, USA) at room temperature without shaking. DNA was extracted as described by Koenig *et al.* (1997) and a cosmid library of Tr18.5 DNA was constructed in the vector pMOcosX as described by Orbach (1994). Filter replicates of the ordered cosmid library were screened by hybridization using an internal 1.2 kb *Sac*I fragment ('SacB') from the *N. haematococca PDA* T9 gene (Maloney and VanEtten, 1994). Probe labeling and hybridization were performed using a random primer fluorescein labeling kit, following the procedures provided by the manufacturer (Dupont NEN 'Renaissance', E.I. du Pont de Nemours & Co., Boston, MA, USA). DNA from colonies which hybridized strongly were analyzed further by Southern blotting.

Construction of cDNA library

Diseased pea tissue was produced by the 'test tube' assay method (VanEtten *et al.*, 1980), and a cDNA library was constructed in the vector Uni-ZAP XR from mRNA derived from 77-13-7-infected pea tissue collected 2 days after inoculation. The mRNA was isolated using streptavidin paramagnetic particles and biotinylated oligo(dT). Synthesis of cDNA, and construction of the library in vector Uni-ZapXR (Stratagene, La Jolla, CA, USA) was according to the manufacturer's instructions. This library was then screened using different portions of the cosmid 55-D-8.

DNA sequence analysis

All cDNA and genomic clones obtained in this research were sequenced by the DNA Sequencing Core Laboratories at either the University of Florida or the University of Arizona. DNA sequences for each clone were analyzed by the sequence similarity search tool BLAST 2.0 (Gapped BLAST) (Altschul *et al.*, 1997) using the BLASTX or BLASTP algorithms with default parameters, provided by the National Center for Biotechnology Information. Codon usage comparisons were conducted using the CORRESPOND program from the sequence analysis software package GCG (Genetic Computer Group, Inc, Madison, WI, USA).

Subcloning of candidate PEP *genes*

Genomic equivalents containing the entire open reading frames as well as the native 5′ and 3′ sequences of the five cDNAs were subcloned into pCB1004 (Sweigard *et al.*, 1997), which carries a hygromycin-resistance cassette so that it can be transformed into filamentous fungi. Genomic sequences corresponding to cDNAs 2–5 were obtained from cosmid 55-D-8, while the cDNA1 region was obtained from cosmid 21-C-9 which contained the entire *PEP1* open reading frame. The size of the subcloned fragments and the restriction enzyme(s) used to create them are as follows: cDNA1 (6.0 kb, *Dra*I); cDNA2 (1848 bp, *Pst*I/*Sac*I or 1767 bp, *Pst*I/*Xho*I); cDNA3 (1522 bp, *Xho*I/*Hin*dIII); cDNA4 (3587 bp, *Kpn*I/*Sac*I); cDNA5 (3545 bp, *Kpn*I/*Xho*I).

N. haematococca *transformation and virulence assays*

Fungal transformations with cosmid 55-D-8 and candidate *PEP* genes were performed as described by Powell and Kistler (1990) without linearization of the vectors. Pathogenicity assays (Kistler and VanEtten, 1984b) were conducted by wound-inoculating pea epicotyls using the *Pisum sativum* cultivar Alaska (Gurney's Seed Company, Yankton, SD, USA). Virulence was determined by measuring the lengths of the epicotyl lesion 6 days after inoculation. The values for lesion length in Table 1 are the mean and standard deviation of independent tests, with seven or more plants assayed in each test. Statistical analysis of the assays was based on a one-tailed Student's *t*-test (two sample assuming unequal variances).

Acknowledgements

This work was supported in part by grants from the National Science Foundation, Division of Molecular and Cellular Biosciences (9905422) and the US Department of Agriculture, National Research Initiative (96-35303-3487).

References

Alexander, N.J., McCormick, S.P. and Hohn, T.M. (1999) TRI12, a trichothecene efflux pump from *Fusarium sporotrichioides*: gene isolation and expression in yeast. *Mol. Gen. Genet.* **261**, 977–984.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389–3402.

Boyd, E.F. and Hartl, D.L. (1998) Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. *J. Bacteriol.* **180**, 1159–1165.

Brown, D.W., Yu, J.-H., Kelkar, H.S., Fernandes, M., Nesbitt, T.C., Keller, N.P., Adams, T.H. and Leonard, T.J. (1996) Twenty-five co-regulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. *Proc. Natl Acad. Sci. USA*, **93**, 1418–1422.

Burd, C.G. and Dreyfuss, G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.

Callahan, T.M., Rose, M.S., Meade, M.J., Ehrenshaft, M. and Upchurch, R.G. (1999) *CFP*, the putative cercosporin transporter of *Cercospora kikuchii*, is required for wild type cercosporin production, resistance, and virulence on soybean. *Mol. Plant–Microbe Interact.* **12**, 901–910.

Ciuffetti, L.M. and VanEtten, H.D. (1996) Virulence of a pisatin demethylase-deficient *Nectria haematococca* MPVI isolate is increased by transformation with a pisatin demethylase gene. *Mol. Plant–Microbe Interact.* **9**, 787–792.

Collmer, A. (1998) Determinants of pathogenicity and avirulence in plant pathogenic bacteria. *Curr. Opin. Plant Biol.* **1**, 329–335.

Enkerli, J., Bhatt, G. and Covert, S.F. (1997) *Nht1*, a transposable element cloned from a dispensable chromosome in *Nectria haematococca*. *Mol. Plant–Microbe Interact.* **10**, 742–749.

Farman, M.L., Taura, S. and Leong, S.A. (1996) The *Magnaporthe grisea* DNA fingerprinting probe MGR586 contains the 3′ end of an inverted repeat transposon. *Mol. Gen. Genet.* **251**, 675–681.

Finlay, B.B. and Falkow, S. (1997) Common themes in microbial pathogenicity revisited. *Microbiol. Rev.* **61**, 136–169.

Galán, J.E. and Collmer, A. (1999) Type-III secretion machines: bacterial devices for protein delivery into host cells. *Science*, **284**, 1322–1328.

Groisman, E.A. and Ochman, H. (1996) Pathogenicity islands: bacterial evolution in quantum leaps. *Cell*, **87**, 791–794.

Hacker, J., Blum-Oehler, G., Mühldorfer, I. and Tschäpe, H. (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol. Microbiol.* **23**, 1089–1097.

He, C.Z., Rusu, A.G., Poplawski, A.M., Irwin, J.A.G. and Manners, J.M. (1998) Transfer of a supernumerary chromosome between vegetatively incompatible biotypes of the fungus *Colletotrichum gloeosporioides*. *Genetics*, **150**, 1459–1466.

Hofmann, K. and Stoffel, W. (1993) TMbase – a database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler*, **347**, 166.

Hohn, T.M., McCormick, S.P. and Desjardins, A.E. (1993) Evidence for a gene cluster involving trichothecene-pathway biosynthetic genes in *Fusarium sporotrichioides*. *Curr. Genet.* **24**, 291–295.

Kempken, F. and Kück, U. (1996) *restless*, an active *Ac*-like transposon from the fungus *Tolypocladium inflatum*: structure, expression, alternative RNA splicing. *Mol. Cell. Biol.* **16**, 6563–6572.

Kistler, H.C. and VanEtten, H.D. (1984a) Three non-allelic genes for pisatin demethylation in the fungus *Nectria haematococca*. *J. Gen. Microbiol.* **130**, 2595–2603.

Kistler, H.C. and VanEtten, H.D. (1984b) Regulation of pisatin demethylation in *Nectria haematococca* and its influence on pisatin tolerance and virulence. *J. Gen. Microbiol.* **130**, 2605–2613.

Kistler, H.C., Meinhardt, L.W. and Benny, U. (1996) Mutants of *Nectria haematococca* created by a site-directed chromosome breakage are greatly reduced in virulence toward pea. *Mol. Plant–Microbe Interact.* **9**, 804–809.

Koenig, R.L., Ploetz, R.C. and Kistler, H.C. (1997) *Fusarium oxysporum* f. sp. *cubense* consists of a small number of divergent and globally distributed clonal lineages. *Phytopathology*, **87**, 915–923.

Langin, T., Capy, P. and Daboussi, M.J. (1995) The transposable element *impala*, a fungal member of the *Tc1-mariner* superfamily. *Mol. Gen. Genet.* **246**, 19–28.

Lawrence, J.G. (1997) Selfish operons and speciation by gene transfer. *Trends Microbiol.* **5**, 355–359.

Lawrence, J.G. and Roth, J.R. (1996) Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics*, **143**, 1843–1860.

Maloney, A.P. and VanEtten, H.D. (1994) A gene from the fungal plant pathogen *Nectria haematococca* that encodes the phytoalexin-detoxifying enzyme pisatin demethylase defines a new cytochrome P450 family. *Mol. Gen. Genet.* **243**, 506–514.

McCluskey, K. and VanEtten, H. (1995) Phylogenetic relationships among PDA genes. *Fung. Genet. Newsl.* **42A**, 44(abstract).

Mecsas, J. and Strauss, E.J. (1996) Molecular mechanisms of bacterial virulence: type III secretion and pathogenicity islands. *Emerg. Infect. Dis.* **2**, 271–288.

Miao, V.P.W., Matthews, D.E. and VanEtten, H.D. (1991) Identification and chromosomal locations of a family of cytochrome P-450 genes for pisatin detoxification in the fungus *Nectria haematococca*. *Mol. Gen. Genet.* **226**, 214–223.

Ochman, H., Lawrence, J.G. and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299–304.

O'Donnell, K., Kistler, H.C., Cigelnik, E. and Ploetz, R.C. (1998) Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. *Proc. Natl Acad. Sci. USA*, **95**, 2044–2049.

Ohki, R. and Murata, M. (1997) *bmr3*, a third multidrug transporter gene of *Bacillus subtilis*. *J. Bacteriol.* **179**, 1423–1427.

Orbach, M.J. (1994) A cosmid with a Hy^R marker for fungal library construction and screening. *Gene*, **150**, 159–162.

Panaccione, D.G., Pitkin, J.W., Walton, J.D. and Annis, S.L. (1996) Transposon-like sequences at the *TOX2* locus of the plant-pathogenic fungus *Cochliobolus carbonum*. *Gene*, **176**, 103–109.

Pao, S.S., Paulsen, I.T. and Saier, M.H., Jr (1998) Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **62**, 1–34.

Paulsen, I.T., Brown, M.H. and Skurray, R.A. (1996) Proton-dependent multidrug efflux systems. *Microbiol. Rev.* **60**, 575–608.

Pitkin, J.W., Panaccione, D.G. and Walton, J.D. (1996) A putative cyclic peptide efflux pump encoded by the *TOXA* gene of the plant-pathogenic fungus *Cochliobolus carbonum*. *Microbiology*, **142**, 1557–1565.

Powell, W.A. and Kistler, H.C. (1990) *In vivo* rearrangement of foreign DNA by *Fusarium oxysporum* produces linear self-replicating plasmids. *J. Bacteriol.* **172**, 3163–3171.

Romeis, T., Kämper, J. and Kahmann, R. (1997) Single-chain fusions of two unrelated homeodomain proteins trigger pathogenicity in *Ustilago maydis*. *Proc. Natl Acad. Sci. USA*, **94**, 1230–1234.

Rosewich, U.L. and Kistler, H.C. (2000) Role of horizontal gene transfer in the evolution of fungi. *Annu. Rev. Phytopathol.* **38**, 325–363.

Salmond, G.P.C. (1994) Secretion of extracellular virulence factors by plant pathogenic bacteria. *Annu. Rev. Phytopathol.* **32**, 181–200.

Schäfer, W., Straney, D., Ciuffetti, L., VanEtten, H.D. and Yoder, O.C. (1989) One enzyme makes a fungal pathogen, but not a saprophyte, virulent on a new host plant. *Science*, **246**, 247–249.

Sweigard, J.A., Chumley, F., Carroll, A., Farrall, L. and Valent, B. (1997) A series of vectors for fungal transformation. *Fung. Genet. Newsl.* **44**, 52.

Tegtmeier, K.J. (1984) The role of naphtharazin production in *Nectria haematococca* MP VI (anamorph *Fusarium solani*) in pathogenesis on pea (*Pisum sativum*). *J. Cell Biol. Suppl.* **8B**, 254.

Urban, M., Bhargava, T. and Hamer, J.E. (1999) An ATP-driven efflux pump is a novel pathogenicity factor in rice blast disease. *EMBO J.* **18**, 512–521.

VanEtten, H.D., Matthews, P.S., Tegtmeier, K.J., Dietert, M.F. and Stein, J.I. (1980) The association of pisatin tolerance and

demethylation with virulence on pea in *Nectria haematococca*. *Physiol. Plant Pathol.* **16**, 257–268.

VanEtten, H., Funnell-Baerg, D., Wasmann, C. and McCluskey, K. (1994) Location of pathogenicity genes on dispensable chromosomes in *Nectria haematococca* MPVI. *Antonie Van Leeuwenhoek*, **65**, 263–267.

Wasmann, C.C. and VanEtten, H.D. (1996) Transformation-mediated chromosome loss and disruption of a gene for pisatin demethylase decrease the virulence of *Nectria haematococca* on pea. *Mol. Plant–Microbe Interact.* **9**, 793–803.

Whittam, T.S., Wolfe, M.L., Wachsmuth, I.K., Orskov, F., Orskov, I. and Wilson, R.A. (1993) Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. *Infect. Immun.* **61**, 1619–1629.

DNA sequence database accession numbers: AF294788, AF315804, AF315805, AF315806, AF315807, AF315808, AF315315.

Author Query Form

Journal: Plant Journal

Article: tpj969

Dear Author,

During the preparation of your manuscript for publication, the questions listed below have arisen. The numbers pertain to the numbers in the margin of the proof. Please attend to these matters and return the form with this proof.

Many thanks for your assistance.

Query Refs.	Query	Remarks
Q1	Au: please give initial/s for Rodriguez	
Q2	Au: please provide a summary sentence for the whole figure.	