Disparate Independent Genetic Events Disrupt the Secondary Metabolism Gene perA in Certain Symbiotic Epichloë Species

Daniel Berry
Massey University, New Zealand

Johanna E. Takach
The Samuel Roberts Noble Foundation

Christopher L. Schardl
University of Kentucky, chris.schardl@uky.edu

Nikki D. Charlton
The Samuel Roberts Noble Foundation

Barry Scott
Massey University, New Zealand

See next page for additional authors

Follow this and additional works at: https://uknowledge.uky.edu/plantpath_facpub

Part of the Plant Pathology Commons

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Repository Citation
Berry, Daniel; Takach, Johanna E.; Schardl, Christopher L.; Charlton, Nikki D.; Scott, Barry; and Young, Carolyn A., "Disparate Independent Genetic Events Disrupt the Secondary Metabolism Gene perA in Certain Symbiotic Epichloë Species" (2015). Plant Pathology Faculty Publications. 34. https://uknowledge.uky.edu/plantpath_facpub/34

This Article is brought to you for free and open access by the Plant Pathology at UKnowledge. It has been accepted for inclusion in Plant Pathology Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
Disparate Independent Genetic Events Disrupt the Secondary Metabolism Gene perA in Certain Symbiotic Epichloë Species

Digital Object Identifier (DOI)
http://dx.doi.org/10.1128/AEM.03721-14

Notes/Citation Information
Published in Applied and Environmental Microbiology, v. 81, no. 8, p. 2797-2807.

Copyright © 2014, American Society for Microbiology. All Rights Reserved.
The copyright holder has granted permission for posting the article here.

Authors
Daniel Berry, Johanna E. Takach, Christopher L. Schardl, Nikki D. Charlton, Barry Scott, and Carolyn A. Young

This article is available at UKnowledge: https://uknowledge.uky.edu/plantpath_facpub/34
Disparate Independent Genetic Events Disrupt the Secondary Metabolism Gene perA in Certain Symbiotic *Epichloë* Species

Daniel Berry, Johanna E. Takach, Christopher L. Schardl, Nikki D. Charlton, Barry Scott, Carolyn A. Young

Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand; The Samuel Roberts Noble Foundation, Ardmore, Oklahoma, USA; Department of Plant Pathology, University of Kentucky, Lexington, Kentucky, USA

Peramine is an insect-feeding deterrent produced by *Epichloë* species in symbiotic association with C₃ grasses. The perA gene responsible for peramine synthesis encodes a two-module nonribosomal peptide synthetase. Alleles of perA are found in most *Epichloë* species; however, peramine is not produced by many perA-containing *Epichloë* isolates. The genetic basis of these peramine-negative chemotypes is often unknown. Using PCR and DNA sequencing, we analyzed the *Epichloë* species; however, peramine is not produced by many of the method was developed to predict peramine chemotypes by combining PCR product size polymorphism analysis with sequencing first and second adenylation domains (A₁ and A₂, respectively) were common sites for such mutations. Using this information, a mutations responsible for abolishing peramine production in full-length perA alleles were also identified. The regions encoding the first and second adenylation domains (A₁ and A₂, respectively) were common sites for such mutations. Using this information, a method was developed to predict peramine chemotypes by combining PCR product size polymorphism analysis with sequencing of the perA adenylation domains.

Fungal secondary metabolites are a diverse group of important but often nonessential organic compounds with a wide range of properties that are likely to be advantageous for the producing organism or in some cases essential for pathogenicity or developmental stages (1–3). These low-molecular-weight compounds tend to only be produced under certain environmental or growth conditions. The biosynthetic pathways for production of any particular class of secondary metabolites are common to many fungi, but production of a specific secondary metabolite is often unique to a small phylogenetic group of species (4). *Epichloë* species are fungal endophytes of C₃ grasses that are known to produce several bioactive alkaloids that provide bioprotective properties to the host plant (5). These secondary metabolites include the indole-diterpenes, ergot alkaloids, lolines, and peramine (Fig. 1) (6, 7). The indole-diterpene lolitrem B and ergot alkaloid ergovaline have significant detrimental effects on the health and production of stock animals that graze infected pastures (7, 8). The lolines are insecticidal (9), and peramine is a potent deterrent of feeding by insects, including the agriculturally important invertebrate pest *Listronotus bonairesis* (Argentine stem weevil) (10–12).

Peramine synthesis is catalyzed by the two-module nonribosomal peptide synthetase (NRPS), peramine synthetase (PerA), encoded by the 8.3-kb gene perA (12). The first module of PerA contains an adenylation (A₁) domain responsible for selection and activation of the proposed substrate amino acid 1-pyrroline-5-carboxylate and a thiolation (T₁) domain that bonds this sub-strate as a thioester via a 4′-phosphopantetheine (4′PPT) linker. The second module contains adenylation (A₂) and thiolation (T₂) domains for selection, activation, and thiolation of the substrate proposed to be arginine. The second module also contains a methylation (M) domain proposed to N-methylate the alpha-amine of the arginine moiety, a condensation (C) domain that catalyzes peptide bond formation, and a variant reductase domain (R⁺) at the C terminus, proposed to be responsible for intramolecular cyclization and release of the dipeptide product.

The genus *Epichloë* (including former *Neotyphodium* spp.) consists of sexual nonhybrid species and asexual, nonpathogenic endophytes that are derived either directly from the sexual species or by hybridization of two or more *Epichloë* progenitors (14, 15). Hybrid *Epichloë* species contain duplicate or even triplicate copies of most genes due to inheritance of an allele from each progenitor. Alleles of perA are found in nearly all *Epichloë* species, with the notable exceptions of *Epichloë* *glyceriae* and *Epichloë* *gansuensis* (16), but perA null alleles are common. One such allele, first identified in the genome sequence of *Epichloë festucae* isolate E2368 (16), has a deletion of the region encoding the C-terminal R⁺ domain of PerA. This deletion is associated with the insertion of the miniature inverted-repeat transposable element (MITE) designated 3m (17). However, there are many other cases of peramine-negative (per⁻) isolates for which the genetic basis is unknown (18, 19).

Peramine production is an important trait when considering...
endophyte strains for deployment in forage grasses and likely provides a selective advantage to endophyte-infected wild grasses. As such, diagnostic methods are useful to identify suitable endophyte isolates and associations for use in agriculture around the world. The objective of this study was to identify and characterize the mutations causing perA null alleles in a collection of hybrid and nonhybrid *Epichloë* species and strains. Using this information, we developed a PCR method to predict the peramine chemotype of endophytes from pure culture and in endophyte-infected plant material.

**MATERIALS AND METHODS**

**Endophyte strains and growth conditions.** Isolates of *Epichloë* species (see Table S1 in the supplemental material) were grown and maintained as previously described (20, 21). Endophyte-infected plant samples (see Table S2 in the supplemental material) were obtained from plant lines maintained under greenhouse conditions at 23°C during the day and 20°C at night, with 16 h of light starting at 0600; light intensity varied throughout the year, depending on the season.

**Peramine analysis.** Peramine was analyzed by AgResearch Grasslands (Palmerston North, New Zealand) from plant material using a modification of the method described by Rasmussen et al. (22). A 50-mg freeze-dried sample taken from endophyte-infected *Lolium perenne* whole tillers was extracted for 1 h with 1 ml of extraction solvent (50% [vol/vol] methanol) with 2.064 mg/ml homopereamine nitrate (AgResearch Grasslands) as an internal standard. The sample was then centrifuged for 5 min at 8,000 × g, and a 500-μl aliquot of the supernatant was transferred to an amber 12- by 32-mm high-performance liquid chromatography (HPLC) vial via a 0.22-μm-pore polyvinylidene difluoride (PVDF) syringe filter. Separation was achieved on a Synergi Polar-RP 100- by 2.00-mm (2.5-μm) column (Phenomenex, Torrance, CA) using a linear gradient profile (where eluent A is aqueous 0.1% formic acid and eluent B is acetonitrile), with time 0 (T₀) at 5% B, T₅₀ at 40% B, T₁₀₀ at 90% B, and T₁₅₀ at 90% B, followed by equilibration to initial conditions over the following 8 min. Peramine was quantified by mass spectroscopy (using homopereamine as an internal standard) according to the parameters described by Rasmussen et al. (22). Peramine is expected to have a retention time of 8.6 min with an MS ion of 248.1 m/z, and homopereamine is expected to have a retention time of 9.9 min with an MS ion of 262.1 m/z. A 5-μl injection volume gave a limit of detection for this technique of 0.1 μg/g for herbage.

**Genomic DNA isolation.** Genomic DNA was isolated from freeze-dried mycelium of *Epichloë* species using the ZR Fungal/Bacterial DNA MiniPrep kit (Zymo Research, Irvine, CA) as per the manufacturer’s instructions. Total plant DNA (including endophyte) was extracted using the MagAttract 96 DNA plant core kit (Qiagen, Inc., Valencia, CA) as per the manufacturer’s instructions.

**Primer design.** Primers for PCR amplification and sequencing (see Table S3 in the supplemental material) were designed using a multiple-sequence alignment of all available perA and flanking gene sequences from the *Epichloë* genome database (www.endophyte.uky.edu) (16), which included 10 strains from seven species. Primers were designed to maximize conservation of the target binding sequence between species. Primers for sequencing specific alleles from hybrid species were designed to contain at least two single nucleotide polymorphisms (SNPs) specific to each allele, with one of these SNPs located at the 3’ terminus wherever possible.

**PCR amplification and product purification**. Genomic DNA templates were amplified using GoTaq DNA polymerase (Promega, Madison, WI) under the conditions described by Takach et al. (18). PCR products for sequencing were purified using the QiAquick PCR purification kit (Qiagen). Where insufficient PCR product was produced for direct sequencing, a second PCR using a 10^- or 10^-fold dilution of the initial reaction was used as a template to increase PCR product concentration; this was often necessary when amplifying the perA gene directly from endophyte-infected plant material.

**Sequencing of perA.** Three overlapping DNA fragments, perA-1, perA-2, and perA-3, or perA-3R, covering the whole perA or perA-ΔR gene, was amplified using primer sets defined in Table S3 in the supplemental material and sequenced with BigDye chemistry v3.1 (Applied Biosystems, Foster City, CA) using an Applied Biosystems 3730 DNA analyzer. Sequences from nonhybrid isolates were assembled using MacVector 12.6 with Assembler (MacVector, Inc., Cary, NC), with further sequencing completed using isolate-specific primers where required. Sequences from hybrid isolates were similarly assembled, but this assembly was then used as a reference to design allele-specific primers based on polymorphic regions to sequence each perA allele.

**Phylogenetic reconstruction.** A2 domain DNA sequences from 39 perA alleles (1,785 bp in length from positions 3575 to 5358 for *perA* from *E. festucae* Fl1) were aligned using ClustalW (23) and manually edited where necessary with MacVector 12.6. DNA sequences spanning from the middle of the T2 domain until the perA-ΔR truncation location were similarly aligned from 38 perA alleles. These alignments were analyzed using Mega 5.1 (24) via the maximum likelihood method using the gamma-distributed (5 categories) Tamura three-parameter nucleotide substitution model (25) and the subtree-pruning-regrafting (level 3) heuristic method on all sites and codons. The bootstrap method with 1,000 repetitions was used to test the phylogeny.

**Nucleotide sequence accession numbers.** The GenBank accession numbers for perA sequences generated by this study are listed as follows: KP347845 to KP347877 and KP719965 to KP719973. Details about these and additional perA accession numbers from other studies (5, 22) are presented in Table S4 in the supplemental material.

**RESULTS**

Peramine chemotypes of *E. festucae* isolates. The distribution of peramine production within *E. festucae* was evaluated from herbage samples of *Lolium perenne* plants symbiotic with *E. festucae* isolate E189, Fl1, Fl3, Fr5, FrC7, Fr1, or Frr1. Of these associations, only the plants infected with Fl1, FrC7, or Frr1 contained peramine (Per<sup>−</sup>) (Table 1). These data demonstrated that pe-
ramine chemotypes can be highly variable and discontinuous, even between isolates of a single *Epichloë* species.

**Analysis of perA across multiple Epichloë species.** Genomic DNA extracted from mycelium of 34 different isolates spanning nine nonhybrid *Epichloë* species, including the *E. festucae* isolates mentioned above, was used in a PCR-based size polymorphism analysis to evaluate the presence and integrity of the *PER* locus. PCR primers were designed to amplify each *perA* domain in overlapping DNA fragments, as well as the conserved flanking genes *mfsA* and *qcrA*. The genes *mfsA* and *qcrA*, but not *perA*, were detected in *E. glycerae* E2772 and *Epichloë elymi* E184 (Fig. 2). The majority of isolates (30/34) gave either a full complement of *perA* fragments or all fragments except the R*-domain fragment (no amplification with the primer pair perA3_3/perA3_R), indicating that these alleles likely lacked R* (which we designate perA-ΔR*). Although the regions encoding the A2 and C domains did not amplify from *Epichloë baconii* As6 and *Epichloë bromicola* E799, respectively, we know from sequencing and other PCR that these fragments are present (data not shown).

Alleles of *perA-ΔR* were first observed in the genome sequence of *E. festucae* E2368, *E. festucae* var. lolii Lp14, and *Epichloë typhina* subsp. *poae* E5819 (16, 17). The region encoding the *perA* R* domain was also missing from both *Epichloë sylvatica* isolates tested and was discontinuously distributed within *E. baconii*, *E. bromicola*, *E. festucae*, and *E. typhina* (Fig. 2). An additional deletion was observed within the region encoding the T1 domain of *perA-ΔR* from *E. sylvatica* isolates E354 and E503 (Fig. 2). Also detected was a deletion in the A2 domain fragment from the otherwise full-length *E. festucae* Fl1 *perA* allele (Fig. 2). The identification of *perA-ΔR* in *E. festuca* E189, Frc5, and Fr1 and an A2 domain deletion in *E. festucae* Fl1 explains the observed per*”* chemotypes (18, 19, 29). Each of these isolates appeared to contain at least one full-length *perA* allele (Fig. 4). The presence of full-length copies of *perA* suggested that small mutations within these gene copies likely generated *perA* null alleles.

The *perA-ΔR* allele was identified in *E. typhina* OR10, *Epichloë siegelii* e915, and an undescribed endophyte, *Epichloë* sp. isolate e4768, from *Festuca versuta* (Fig. 4). The absence of the R*-domain product in *E. siegelii* e915, a two-parent hybrid, suggests that this strain contains two *perA-ΔR* alleles. For *Epichloë* sp. strain e4768, the successful amplification of both the R*-domain and ΔR* deletion-specific PCR product indicates this isolate is a hybrid containing both the *perA* and *pera*-ΔR* alleles.

A draft genome sequence of *E. siegelii* e915 was used to explore the region flanking the two *pera*-ΔR* alleles (Fig. 3C). Annotation of *pera*-ΔR*, *mfsA*, *qcrA*, and repeat sequences that flank these genes revealed that the *pera*-ΔR* allele 1 was nearly identical to the arrangement found in *E. festucae* E2368 (Fig. 3B). Interestingly, the e915 *pera*-ΔR* allele 2, originating from the *E. bromicola* progenitor, was oriented toward *mfsA* rather than *qcrA*, indicating a gene inversion event has occurred. Although the common 17-bp region was still associated with this allele, there were no longer any downstream repetitive sequences. The

### Table 1: Peramine concentrations of whole tillers from *Lolium perenne* infected by different *Epichloë festuca* isolates

<table>
<thead>
<tr>
<th><em>E. festuca</em> strain</th>
<th>Peramine concn (ppm)*</th>
<th>Gene feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>E189</td>
<td>ND*</td>
<td><em>perA</em>-ΔR*</td>
</tr>
<tr>
<td>Fl1</td>
<td>ND</td>
<td>Deletion in A2 domain</td>
</tr>
<tr>
<td>F1</td>
<td>15–90</td>
<td><em>pera</em> functional</td>
</tr>
<tr>
<td>Frc5</td>
<td>ND</td>
<td><em>perA</em>-ΔR*</td>
</tr>
<tr>
<td>Frc7</td>
<td>19.5</td>
<td><em>pera</em> functional</td>
</tr>
<tr>
<td>Fr1</td>
<td>ND</td>
<td><em>perA</em>-ΔR*</td>
</tr>
<tr>
<td>Fr1r</td>
<td>137.2</td>
<td><em>pera</em> functional</td>
</tr>
</tbody>
</table>

* Determined by combined liquid chromatography-mass spectroscopy. The limit of detection was 0.1 ppm, and the limit of quantification was 0.5 ppm.

<table>
<thead>
<tr>
<th>Gene feature</th>
<th>Gene feature</th>
<th>Gene feature</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>perA</em></td>
<td><em>mfsA</em></td>
<td><em>qcrA</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><em>mfsA</em></th>
<th><em>qcrA</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><em>mfsA</em></th>
<th><em>qcrA</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>
perA-ΔR* allele of *E. bromicola* isolate E799 contained a similar orientation (data not shown). Linkage between the contigs from e915 containing perA-ΔR* allele 2 and qcrA allele 2 cannot be determined from this sequence and were not able to be connected by PCR, likely due to the AT-rich repeat sequence that flanks qcrA (Fig. 3C).

Sequencing and characterization of perA variants. To determine why some isolates appeared to contain an intact perA gene...
FIG 3 Analysis of perA-ΔR* downstream repeat sequences. Schematic representation of *Epichloë* isolates that lack the perA-R* domain. (A) Overview of the functional perA gene required for peramine production and the associated flanking gene qcrA from *E. festucae* F11. The domains of PerA are detailed in Fig. 2. (B) Schematic comparisons of regions from the perA-T2 domain to qcrA. The regions from *E. festucae* isolates F11 and E2368 and *E. typhina* subsp. *poae* E5819 were drawn from perA GenBank accession no. AB205145, JN640287, and JN640289, respectively, and *E. bromicola* AL0426_2 was generated from a genome sequence. The remaining examples were amplified with primer set perA-mid2-F2/Ef104-R using genomic DNA. Maps are arranged to illustrate synteny and do not necessarily suggest an evolutionary history. Syntenic regions (which may include small indels) between sequences are indicated by light gray polygons. Black vertical boxes indicate MITE 3m or 25m. Repeat sequence 16 (a putative retrotransposon) (16) is indicated by a dark gray horizontal box. The primer region for perA-17bp_R is shown in all perA-ΔR* sequences. (C) Schematic representation of the PER loci from a draft genome sequence of the hybrid species *E. siegelii* e915 demonstrating the inverted orientation of perA-ΔR* allele 2 relative to the flanking gene mfsA. Linkage between perA-ΔR* allele 2 and qcrA allele 2 is likely but cannot be proven due to the position of these genes on the ends of their respective contigs. The progenitor species from which each *E. siegelii* allele is derived is indicated in parentheses as “Efe” for *E. festucae* and “Ebm” for *E. bromicola*. Repeat sequences Ono and 28 (putative retrotransposons) (16) are indicated by dark gray horizontal boxes.
but did not produce peramine, the perA alleles from 27 nonhybrid and seven hybrid isolates were sequenced or evaluated from genome sequences (Fig. 5; see Table S4 in the supplemental material). Of these 34 isolates, 11 were known to be Per/H11001, and 17 were known to be per/H11002; the peramine chemotypes of the remaining 6 isolates were unknown.

The identification of perA-ΔR* explained the per/H11002 chemotype for 13 per/H11002 isolates (Fig. 5C), and the presence of a deletion in the A2 domain in E. festucae Fg1 explained the per/H11002 chemotype of this isolate (Fig. 5A). The remaining three per/H11002 isolates, E. cabralii BlaTG-2 isolate NFe661, E. uncinata isolate e167, and Epichloë sp. FaTG-2 G3 isolate NFe45115, were hybrids (18, 19, 29). A 1-bp insertion causing a frameshift mutation was identified in E. cabralii BlaTG-2 NFe661 allele 1 (Fig. 5B), and an SNP that resulted in a nonsense mutation was identified in allele 2 (Fig. 5B). Analyses of the allele sequences from the hybrid isolate E. uncinata e167 identified independent frameshift mutations in alleles 1 and 2, generating perA null alleles (Fig. 5A). An SNP identified in the first adenylation domain of Epichloë sp. FaTG-2 G3 isolate NFe45115 allele 2 resulted in a nonsense mutation, while deletions are present in both A domains of allele 1 (Fig. 5A). These mutations explain the per/H11002 chemotype of all three isolates. The peramine chemotype was unknown for 6 of the 34 isolates from which perA was sequenced, so predictions were made by

FIG 4 Analysis of perA integrity from infected plant material. PCR analysis of perA from plants infected with Epichloë species. The PCR products were produced using primers designed to conserved sequence in or near each of the major perA domains and the two flanking genes mfsA and qcrA (see Table S3 in the supplemental material). Gray shading indicates the regions amplified. The dashed lines indicate the region that will amplify with primers perA-mid2-F2/perA_17bp_R only from isolates containing perA-ΔR* (Fig. 3). The PerA domains shown above the perA gene map are listed in Fig. 2. Isolates for which perA was subsequently sequenced are shown in boldface. Endophytes with unknown hybrid status are labeled “unknown.” LAE, Lolium-associated endophyte; Eam, E. amarillans; Ebm, E. bromicola; Eel, E. elymi; Efe, E. festucae; Ety, E. typhina; Etp, E. typhina subsp. poae; nt, not tested. Known peramine chemotypes are indicated as Per/H11001 (peramine producer) or per/H11002 (peramine nonproducer) (5, 6, 18, 19, 29, 43, 44). Peramine production was predicted for isolates with unknown peramine chemotypes based on the presence of expected PCR products amplified from all perA domains and are indicated by + or −.
analyzing these sequence data. Of the six isolates, *E. baconii* E1031 and *E. bromicola* AL0434 were predicted to be Per⁺ because they both contained full-length alleles with no nonsense, frameshift deletion, or insertion mutations. Of the four remaining isolates, *E. baconii* As6 and *E. bromicola* AL0426_2 were nonhybrid isolates containing perA-ΔR* alleles and were therefore predicted to be per⁻. The perA allele from *Epichloë mollis* AL9924 contained multiple small (<10-bp) insertions causing frameshift mutations (Fig. 5A) and was therefore predicted to be per⁻. Based on phylogenetic analysis, we judged the partial sequence from hybrid *Epichloë* sp. e4768 allele 2 to be derived from an *E. typhina* progenitor (Fig. 6B), and it is considered a perA null allele due to a...
1-bp deletion that results in a frameshift mutation (Fig. 5B). Isolate e4768 allele 1 was \textit{perA-D}\textsuperscript{R*} and was derived from an \textit{E. festucae} E2368-like progenitor (Fig. 6B). Given that this isolate is a hybrid and has two null alleles, it is predicted to have a per\textsuperscript{-}chemotype.

Analysis of the 18 sequenced \textit{perA-D}\textsuperscript{R*} alleles identified eight isolates that contained frameshift mutations, large deletions, and/or nonsense mutations in addition to the known 3M MITE-associated R\textsuperscript{*}-domain deletion (Fig. 5C). A large deletion spanning the junction between the regions encoding the A1 and T1 domains and two small insertions were present in the alleles from \textit{E. sylvatica} E354 and E503 (Fig. 5C), and nonsense mutations that should significantly truncate the translated protein were present in the \textit{E. typhina} E505 and \textit{E. typhina} subsp. \textit{clarkii} Holcus 3 alleles. The \textit{E. bromicola} AL0426\texttt{2} \textit{perA-D}\textsuperscript{R*} allele contained an SNP that disrupted the start codon, but a potential alternate ATG codon was located 189 bp downstream that does not truncate any conserved A-domain motifs. The nonsense mutations identified in the \textit{perA-D}\textsuperscript{R*} alleles from \textit{E. bromicola} E799 and \textit{E. typhina} ORE04 and ORE06 were located close to the existing \textit{perA-D}\textsuperscript{R*} truncation, and no other \textit{perA} domains were affected (Fig. 5C). Both alleles from the hybrid isolate \textit{E. siegelii} e915 were confirmed to be \textit{perA-D}\textsuperscript{R*}. A nonsense mutation that would result in truncation of the translated protein was identified in e915 allele 2 (Fig. 5C).

**Phylogeny of \textit{perA} A2- and T2-domain DNA sequences.** Unrooted maximum likelihood phylogenetic trees were generated from the A2-domain DNA sequence of 39 \textit{perA} and \textit{perA-D}\textsuperscript{R*} alleles, and the DNA sequence spanning from the middle of the T2 domain to the location of the \textit{perA-D}\textsuperscript{R*} truncation (T2-\textit{ΔR*}) from 38 \textit{perA} and \textit{perA-D}\textsuperscript{R*} alleles (Fig. 6). The A2-domain phylogeny revealed that all \textit{E. typhina} complex-derived \textit{perA-D}\textsuperscript{R*} alleles grouped in a clade distinct from \textit{E. typhina} complex \textit{perA}.
alleles (69% bootstrap support). The E. bromicola-derived perA-ΔR* alleles also grouped separately from the E. bromicola perA alleles (80% bootstrap support). The perA-ΔR* alleles derived from E. festucae and E. baconii isolates grouped separately from perA alleles from Epichloë amarillians, E. bacconi, and E. festucae isolates (97% bootstrap support). In contrast, the phylogeny of the T2-ΔR* region showed all perA-ΔR* alleles grouped together separated from the perA alleles despite originating from multiple different species (Fig. 6). The A2 domain of FaTG-2 G3 NFe45115 perA allele 2 (derived from E. festucae) grouped with the E. bacconi- and E. festucae-derived perA-ΔR* alleles, despite this allele still retaining the R* domain, as was previously observed by Takach et al. (18) (Fig. 6). In contrast the T2-ΔR* region from NFe45115 allele 2 grouped with related perA alleles (Fig. 6).

**DISCUSSION**

Peramine has been reported as the most commonly produced alkaloid by Epichloë species (66%), yet a discontinuous distribution is found within and between species (5, 6). In this study, we identify nonfunctional perA alleles from hybrid and nonhybrid Epichloë species. Although E. glyceriae and E. gansuensis have been previously shown to lack perA (16), we show the discontinuous distribution of peramine producers across Epichloë species is most frequently associated with mutations within perA that abolish peramine production. Analyses of apparently nonfunctional perA alleles show that each inactivating mutation is either isolate specific or is shared between closely related isolates of nonhybrids or closely related genomes in hybrids. Also, the DNA sequences encoding the first and second adenylation (A1 and A2) domains are common sites for such inactivating mutations (Fig. 5). These data indicate that independent mutation events have inactivated perA many times. The only exception to this rule is the identical R*-domain deletion found in all perA-ΔR* alleles. Identification of these inactivating perA mutations provides information to aid in prediction of peramine producers for isolates with unknown peramine chemotypes and genetic diagnosis for isolates known to be per−.

In this study, the PER locus was evaluated using PCR to assess the integrity of each perA domain from 67 isolates representing at least 20 Epichloë species. We were able to distinguish the full-length perA alleles found in peramine producers and the perA-ΔR* alleles found in strains that are unable to produce peramine. Amplification of genomic DNA using a primer pair specific to perA-ΔR* (perA-T2_F and perA-17bp_R) identified 18 (27%) isolates missing the R* domain, of which two isolates, E. siegelli e915 and Epichloë sp. e4768, were hybrid species and the remaining were nonhybrids. However, the PCR analyses used to detect the presence of each domain did not reveal all perA mutants. Sequencing of perA from isolates unable to make peramine revealed frameshift and nonsense mutations predominantly within the regions encoding the A1 and A2 domains that would render perA nonfunctional. From these sequence data, we were able to explain the mutations responsible for the peramine-negative chemotype previously identified in E. uncinata e167, Epichloë sp. FaTG-2 G3 isolate NFe45115, and E. cabralii BlaTG-2 isolate NFe661 (18, 19, 29).

The isolate- or lineage-specific nature of mutations that have resulted in perA null alleles contrasts sharply with the taxonomic distribution of perA-ΔR*. The perA-ΔR* alleles are distributed widely within the Epichloë genus, occurring in a subset of isolates from each of the E. baconii, E. bromicola, E. festucae, and E. typhina complex (ETC) clades (Fig. 6). Given that the R* domain was deleted at identical sites within perA and there is high sequence conservation immediately downstream of the perA-ΔR* alleles, it is unlikely this deletion occurred more than once. In support of the possibility that a single event was responsible is the consistent association of perA-ΔR* with a downstream MITE 3m sequence and a unique 17-bp sequence containing an in-frame stop codon. Both of these features were absent from all perA alleles, so the MITE 3m insertion seems likely to have been involved in the deletion of the region encoding the R* domain.

The evolution of the perA-ΔR* alleles appears particularly complex, considering the disparate phylogenies of the regions encoding A2 and T2-ΔR* (Fig. 6). The T2-ΔR* phylogeny placed perA and perA-ΔR* into separate clades, each independently reflecting relationships of the Epichloë species. The T2-ΔR* phylogeny suggested transspecies polymorphism (TSP), whereby the corresponding sequences in perA and perA-ΔR* diverged early during, or even before, evolution of the genus Epichloë. This pattern is similar to evidence of TSP in other systems, such as vertebrate major histocompatibility loci (30) and fungal vegetative incompatibility loci (31).

In contrast to the T2 phylogeny, the phylogeny of the A2-encoding sequences consistently grouped perA-ΔR* alleles with perA of the same or closely related species, although in most species, the separation of the perA-ΔR* and perA subclades seemed deeply rooted in the species. The disparity between the A2 and T2-ΔR* phylogenies suggests multiple recombination events. What appears to be the most recent example affected perA allele 2 in FaTG-2 G3 isolate NFe45115. The region encoding the A2 domain of this perA allele groups with the E. festucae and E. baconii perA-ΔR* alleles, whereas the T2-ΔR* sequence groups with E. festucae perA alleles (Fig. 6). The fact that multiple species clades exhibit the disparate A2 and T2-ΔR* phylogenies suggests that a recombination event could exist between these two portions of perA.

The only two hybrid isolates containing perA-ΔR* alleles were E. siegelli isolate e915 and Epichloë sp. e4768. This is perhaps surprising given the wide distribution seen in the sexual isolates (44% of isolates tested in this study) (Fig. 2) and the number of hybrid species we tested (21 isolates representing 10 species) that contain E. festucae (7 isolates) and E. typhina (14 isolates) ancestral progenitors.

Previous studies of Epichloë alkaloid biosynthesis loci, such as the ergot alkaloid (EAS), indole-diterpene (IDT/LTM), and loline (LOL) gene clusters, have shown the presence or absence of pathway-specific genes to be the primary factor determining chemotype diversity of these alkaloids (16, 32, 33). The IDT/LTM and EAS loci are localized to dynamic subterminal regions of chromosomes (16), and both these and the LOL gene cluster are closely associated with a variety of transposable elements (16, 17, 34). These factors provide mechanisms through which genes from these clusters, or even an entire gene cluster, may be lost via recombination when selective pressure for a cluster is reduced. In contrast, with the exception of perA-ΔR* alleles, full-length perA alleles have not been found in association with transposable elements (16). In the absence of selective pressure, perA is likely to be retained longer than genes from the other secondary metabolite gene clusters, and this could explain the observed increase of perA inactivation by nonsense, frameshift, or deletion mutations rela-
tive to gene loss events common to the EAS, IDT/LTM and LOL gene clusters.

Diagnostic PCR utilizing markers developed from sequences of housekeeping and secondary metabolite biosynthetic genes is an effective approach to identify and quantify potential contamination from mycotoxin-producing fungi within foodstuffs for human and animal consumption (35, 36). For example, multiplex PCRs have been successfully used to simultaneously detect multiple fungal genera found in cereals that are likely to produce ochratoxins and trichothecenes (37). A quantitative PCR (qPCR) assay that detects polymorphisms within TRI12 can identify different trichothecene genotypes within Fusarium species from field samples (38). Chemotype prediction using PCR to detect the presence of biosynthesis genes has also been very successful when evaluating Epichloë species for the ability to produce ergot alkaloids, indole-diterpenes, and lolines and provides insight into the bioactive potential of any given endophyte isolate (18, 19, 32, 39–41). In all of these approaches, the ability to directly analyze infected plant material by PCR provides rapid detection methods for a wide range of organisms and their biosynthetic potential. To determine whether an endophyte isolate is likely to produce peramine, we have refined the PCR approach described previously (18, 19, 39, 41) in order to identify the presence and integrity of all domains encoded by perA. In addition, sequence analysis of the regions encoding the A1 and A2 domains can be used to identify the most commonly found mutations. Using this pipeline, specific isolates with known and unknown peramine chemotypes were screened to identify perA-R^* alleles and other observable deletions, and sequence analysis was used to identify frameshift and nonsense mutations that would render perA nonfunctional. Although this method will not eliminate the need to evaluate peramine production, especially for determination of the levels of peramine produced by a given isolate, it does provide insight into the likelihood of peramine production. Evaluation of endophyte-infected plant germplasm for potential peramine producers as well as production of other bioactive alkaloids will help us understand the bioprotective potential of Epichloë species and facilitate investigation into the effects of different geographic and selective pressures on the evolution of this locus.

ACKNOWLEDGMENTS

We acknowledge Ginger A. Swoboda (The Samuel Roberts Noble Foundation) for technical assistance, Pierre-Yves Dupont (Massey University) for assistance with phylogeny reconstruction methodology, and Wade Mace (AgResearch) for measurement of peramine concentrations. We thank Adrian Leuchtmann (ETH Zurich) for access to Epichloë bromicola (AL0434 and AL0426_2) genome sequences.

This research was supported by a Massey University Ph.D. scholarship and a grant from the Royal Society of New Zealand Marsden Fund (contract MAU1002). Genome sequencing was supported by United States Department of Agriculture grants 2012-67013-19384 and 2010-34457-31010.

REFERENCES


