

# Genomes of Plant-Associated Clavicipitaceae

Christopher L. Schardl<sup>\*,1</sup>, Carolyn A. Young<sup>†</sup>, Neil Moore<sup>‡</sup>,  
Nicholas Krom<sup>†</sup>, Pierre-Yves Dupont<sup>§</sup>, Juan Pan<sup>\*</sup>, Simona Florea<sup>\*</sup>,  
Jennifer S. Webb<sup>\*</sup>, Jolanta Jaromczyk<sup>\*</sup>, Jerzy W. Jaromczyk<sup>‡</sup>,  
Murray P. Cox<sup>§</sup>, Mark L. Farman<sup>\*</sup>

<sup>\*</sup>Department of Plant Pathology, University of Kentucky, Lexington, Kentucky, USA

<sup>†</sup>The Samuel Roberts Noble Foundation, Ardmore, Oklahoma, USA

<sup>‡</sup>Department of Computer Science, University of Kentucky, Lexington, Kentucky, USA

<sup>§</sup>Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand

<sup>1</sup>Corresponding author: e-mail address: schardl@uky.edu

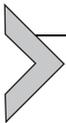
## Contents

1. Introduction	292
1.1 Biology of plant-associated Clavicipitaceae	292
1.2 Phylogenetic relationships	296
2. Sequenced Genomes of the Clavicipitaceae	299
2.1 Repeat content	303
2.2 Telomeric regions	304
2.3 Mating-type loci	305
2.4 Mitochondrial genomes	305
2.5 Gene ontology categories	306
2.6 Variation of SM clusters	308
3. Alkaloid Gene Loci	309
3.1 Relationships between gene contents and alkaloid structures	311
3.2 Variation in gene order	319
3.3 Variations in gene expression	320
4. Future Perspectives	321
Acknowledgments	322
References	322

## Abstract

Fungi of family Clavicipitaceae serve as models for evolution on the symbiotic continuum from pathogenic to mutualistic. Clavicipitaceous fungi associate with plants, invertebrates, and other fungi. Most plant-associated Clavicipitaceae systemically colonize shoots, but the fungal fruiting structures are localized to inflorescences, florets, buds, leaves, or nodes. Many Clavicipitaceae decrease or eliminate host seed production, but some have evolved such intimate symbioses with plant hosts that they disseminate

clonally in seeds (vertical transmission) without damage or any reduction in plant fertility. In such cases, the fungi dramatically enhance host fitness by producing defensive alkaloids and through other mechanisms. To date, sequences have been assembled for 26 Clavicipitaceae representing 21 species in seven genera. These include three *Claviceps* species that fruit on and replace host ovaries, two *Metarhizium* species that parasitize insects and associate with plant roots, and 21 strains of systemic plant parasites or symbionts. Of the latter, 14 are capable of vertical transmission, and of those, 7 are strictly seed-borne mutualists in genera *Epichloë* and *Periglandula*. Alkaloid biosynthetic genes are widely distributed among these fungi. Gene clusters for ergot alkaloids and indole-diterpenes, both of which are neurotoxins in vertebrates and invertebrates, are present in members of all seven genera. The genes for anti-insect loline alkaloids and peramine have a more restricted distribution, but are present in many of the vertically transmissible *Epichloë* species. The availability of these genome sequences will facilitate studies of the evolution and mechanisms underlying the diversity of metabolism, host interactions, and niche adaptation of plant-associated Clavicipitaceae.



## 1. INTRODUCTION

### 1.1. Biology of plant-associated Clavicipitaceae

The family Clavicipitaceae, order Hypocreales, is composed of fungi that interact with a broad range of invertebrate animals and plants and occasionally with other fungi. The associations with plants span a symbiotic continuum including parasitism, mutualism, and pleiotropic symbioses where the relative benefits to host and symbiont depend on the developmental stages and modes of transmission and may be mediated by environmental variables. Diversity among the plant-associated Clavicipitaceae and their potential to protect host plants are also starkly evident in the variety of specialized (secondary) metabolites (SMs) that they produce, including antiherbivore alkaloids belonging to four different chemical classes and exhibiting considerable structural variants within those classes. Recent sequencing of genomes from a wide diversity of plant-associated Clavicipitaceae (Schardl, Young, Hesse, et al., 2013; Schardl, Young, Pan, et al., 2013) should facilitate comparative genomic analyses that can help address the mechanistic basis for mutualism and parasitism; the basis for variation in expression, location, and structure of fruiting bodies; variation in the ability and efficiency of vertical transmission in host seeds; and metabolic diversity. For example, such a comparative analysis has already revealed that the alkaloid gene loci in *Epichloë* species are highly dynamic and tend to have much more abundant repeat sequence than in representatives of genera *Aciculosporium*, *Claviceps*, *Metarhizium*, and

*Periglandula* (Schardl, Young, Hesse, et al., 2013; Schardl, Young, Pan, et al., 2013). The dynamics of SM genes and loci, as revealed through comparative genomics, is the main focus of this chapter.

### 1.1.1 Symbiosis and transmission strategies of plant-associated Clavicipitaceae

The interactions of Clavicipitaceae with plant hosts range from pathogenic to highly mutualistic (Table 10.1). Some (e.g. *Claviceps* species and *Neodaviceps monostipa*) replace host seeds with their fruiting structures (Pažoutová, Kolarik, & Kolinska, 2004; Tudzynski & Scheffer, 2004), while others (*Metarhizium* species) are plant root associates and insect pathogens (Gao et al., 2011). Many form systemic (endophytic) associations throughout host shoots but fruit in a highly localized manner on leaves, nodes, buds, or inflorescences. Such systemic associations characterize *Aciculosporium* (Tanaka & Tanaka, 2008), *Atkinsonella* (Leuchtman & Clay, 1989), *Balansia* (Diehl, 1950), *Ephelis* (Tanaka & Tanaka, 2008), *Epichloë* (Leuchtman, Bacon, Schardl, White, & Tadych, 2014), *Heteroepichloë* (Tanaka & Tanaka, 2008), *Myriogenospora* (Glenn, Rykard, Bacon, & Hanlin, 1998), and *Parepichloë* (White & Reddy, 1998) species and possibly also *Cephalosporium phalaridis* (Walker, 2004), *Corallocytophthora* species (Pažoutová et al., 2004), *Neodaviceps monostipa* (White & Reddy, 1998), and *Nigrocornus scleroticus* (Ryley, 2003). Some species grow in intercellular spaces (endobiotic growth), whereas others are restricted to surfaces of, and spaces between, plant tissue layers (epibiotic growth). We consider “endophytes” to encompass all species that grow in asymptomatic plant parts, whether endobiotically or epibiotically.

Remarkably, some plants have established symbioses with Clavicipitaceae that are heritable, being vertically transmitted via host seeds. In particular, many cool-season grasses (Poaceae subfamily Poöideae) possess seed-borne *Epichloë* species (including *Neotyphodium* species) (Schardl, 2010), and many morning glories (Convolvulaceae tribe Ipomoeae) have seed-borne *Periglandula* species (Steiner, Leibner, Schardl, Leuchtman, & Leistner, 2011). These same plant groups can possess other heritable symbionts (An et al., 1993; Cook et al., 2013), but the associations with Clavicipitaceae are particularly common and are the best-documented defensive mutualisms. Symbiotic and parasitic Clavicipitaceae often produce a wide array of alkaloids that antagonize invertebrate and sometimes vertebrate herbivores (Schardl, Young, Faulkner, Florea, & Pan, 2012; Schardl, Young, Hesse, et al., 2013).

**Table 10.1** Characteristics of genera in plant-associated Clavicipitaceae<sup>a</sup>

<b>Genus</b>	<b>Hosts</b>	<b>Seed transmission</b>	<b>Systemic growth</b>	<b>Fruiting type</b>	<b>Fruiting location</b>
<i>Aciculosporium</i>	Poaceae, Bambusoideae	No	Endobiotic	Stroma	Bud
<i>Atkinsonella</i>	Poaceae	Yes	Epibiotic	Stroma	Inflorescence
<i>Balansia</i>	Poaceae or Cyperaceae	No	Endobiotic or epibiotic	Stroma	Bud, leaf, node, or inflorescence
<i>Cepsiclava</i>	Poaceae	No	Endobiotic	Sclerotium	Floret
<i>Claviceps</i>	Poaceae	No	None	Sclerotium	Floret
<i>Corallocytostroma</i>	Poaceae	No	ND	Sclerotium	Inflorescence
<i>Epichloë</i>	Poaceae, Poöideae	Yes	Endobiotic	Stroma	Inflorescence
<i>Heteroepichloë</i>	Poaceae, Bambusoideae	No	Epibiotic	Stroma	Leaf
<i>Metarhizium</i>	Plants and insects	No	In insect	Stroma	Insect
<i>Myriogenospora</i>	Poaceae, Panicoideae	No	Epibiotic	Stroma	Leaf
<i>Neoclaviceps</i>	Poaceae	No	ND	Hypothallus	Floret
<i>Nigrocornus</i>	Poaceae	No	ND	Sclerotium	Bud
<i>Parepichloë</i>	Poaceae, Panicoideae	No	Epibiotic	Stroma	Inflorescence
<i>Periglandula</i>	Convolvulaceae, Ipomoeae	Yes	Epibiotic	None	None
<i>Villosiclava</i>	Poaceae, Oryzoideae	No	ND	Sclerotium	Floret

<sup>a</sup>Only genera described from sexual structures (teleomorphs) are listed. ND, not determined.

### 1.1.2 Vertically transmitted symbionts, including asexual *Epichloë* species

The process of vertical transmission has best been described for *Epichloë* species in grasses (Freeman, 1904; Philipson & Christey, 1986; Sampson, 1937). These fungi can colonize most aerial portions of the host plant, including meristematic zones. In most cases, even the ovary and eventually the embryo are infected but undamaged, leading to vertical transmission. Among the Clavicipitaceae, this heritability is known only for *Epichloë* and *Periglandula* species and for *Atkinsonella hypoxylon* (Clay, 1994). It seems somewhat ironic, then, that sexual *Epichloë* species fruit on immature host inflorescences and arrest their development, thereby preventing seed production from the symptomatic tillers. This phenomenon, known as “choke disease”, also occurs on most or all reproductive tillers of grasses infected with *Atkinsonella* or *Corallocytostroma* species (Pažoutová et al., 2004) and on grasses and sedges (Cyperaceae) infected with certain *Balansia* and *Ephelis* species (Diehl, 1950; Tanaka & Tanaka, 2008). What is remarkable about the sexual *Epichloë* species is that, in most infected hosts, the majority of reproductive tillers are completely asymptomatic and give rise to normal seeds bearing the fungal symbiont. In such symbioses, both the grass plant and its symbiotic *Epichloë* species have the benefit of a complete range of reproductive capabilities.

Asexual *Epichloë* (*Neotyphodium*) species primarily or exclusively use vertical transmission in host seeds as their means of dissemination (Schardl, Leuchtman, & Spiering, 2004). The vast majority of the asexual *Epichloë* species have not been reported to produce any external hyphae or spores under natural conditions, but a few have been observed to produce hyphal nets or stromata bearing mitotic spores (conidia) (Tadych, Ambrose, Bergen, Belanger, & White, 2012). Though claims have been made that conidia can mediate the horizontal transmission of *Epichloë typhina* ssp. *poae* (Tadych et al., 2012), direct infection by germinating conidia (rather than from growing cultures) has not been demonstrated. However, hyphal proliferation enhanced by arthropod activities and host wounds, as observed for *E. typhina* (Alderman, 2013), may well provide the means for horizontal transmission. An overwhelming body of evidence nevertheless supports vertical transmission in seeds as the primary and often exclusive means of transmission for the asexual *Epichloë* species.

Most asexual *Epichloë* species produce several antiherbivore alkaloids, often in high abundance (Schardl et al., 2012; Schardl, Young, Pan, et al., 2013), as would be expected when such systems are subject to selection

due to reliance on vertical transmission (Clay & Schardl, 2002; Selosse & Schardl, 2007). However, asexual reproduction can also exact a significant genetic cost in the long term, and it may be for that reason that most asexual *Epichloe* species are recently derived interspecific hybrids with two or even three genomes traceable to sexual ancestors (Schardl, 2010; Selosse & Schardl, 2007).

### 1.1.3 Biology and life history of ergot fungi

Although alkaloid production and vertical transmissibility can be a basis for mutualistic symbiosis, alkaloid production is not limited to the mutualistic symbionts, but is widespread in the Clavicipitaceae. Particularly important sources of alkaloids are the *Claviceps* species, the infamous “ergot” fungi that can contaminate grain supplies. These fungi infect individual host florets via the stigma, ultimately engulfing the host ovary and developing into dense sclerotia (resting structures) that are also called ergots (from old French, “argot”, “cock’s spur”, in reference to their shape) (Pažoutová et al., 2004; Tudzynski & Scheffer, 2004). Ergots resemble seeds in density and sometimes size and shape, are not easily removed with the chaff, and were a dangerous source of mycotoxins until the advent of modern mechanical techniques to remove them (Schardl, Panaccione, & Tudzynski, 2006). One reason that wheat (*Triticum aestivum*) is preferable among the grains is that it self-pollinates before florets open, largely eliminating opportunities for *Claviceps* spores to access susceptible young stigmata. Open-pollinated grains such as rye (*Secale cereale*) and seed heads of forage grasses such as tall fescue (*Lolium arundinaceum*) are much more susceptible to infection by *Claviceps* species. Compounds of the ergot alkaloid chemical class have been purified from ergot fungi and sometimes chemically modified into important pharmaceuticals or into the illicit drug, lysergic acid diethylamide (LSD) (Hofmann, 1978; Schardl et al., 2006). Ergots can also contain tremorgenic indole-diterpenes, which have received less consideration in human health but are a recognized cause of livestock poisoning (Schardl, Young, Hesse, et al., 2013; Uhlig, Botha, Vrålstad, Rolén, & Miles, 2009).

## 1.2. Phylogenetic relationships

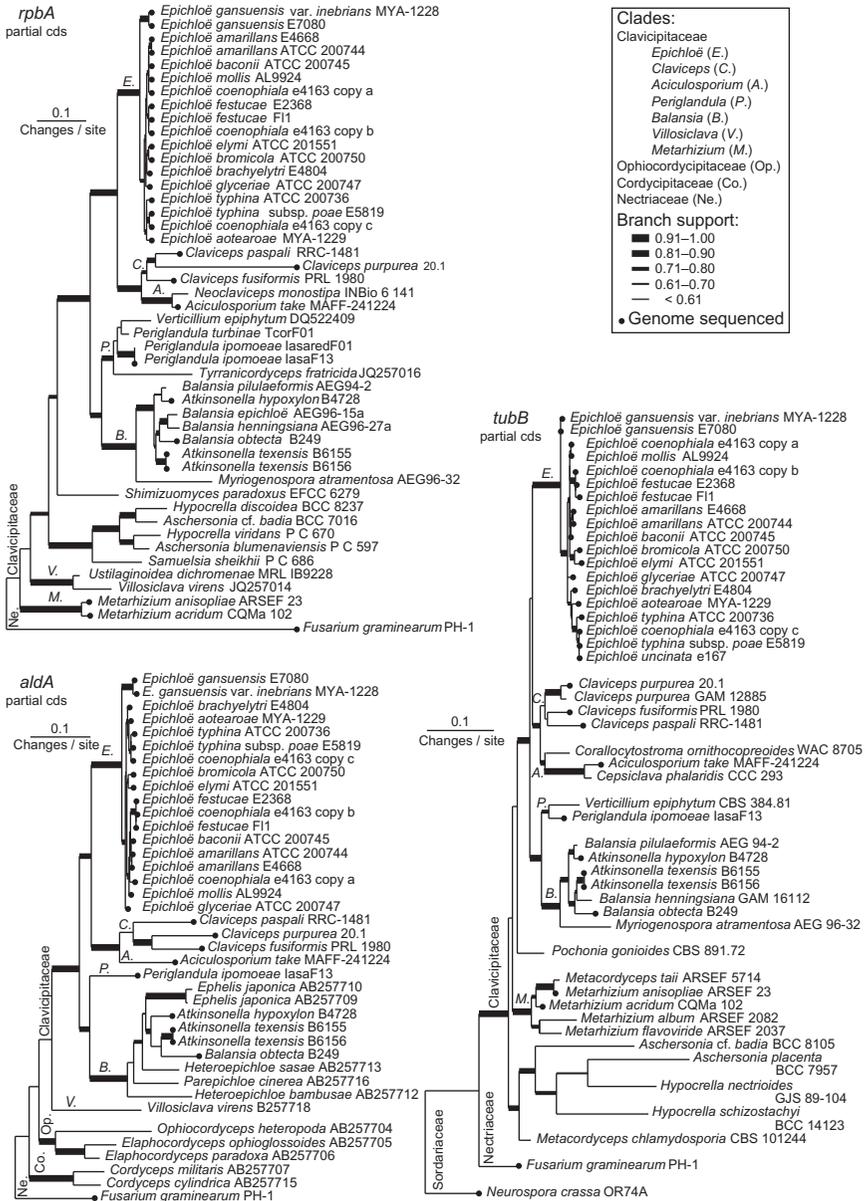
### 1.2.1 Gene trees

Phylogenetic studies of the order Hypocreales have recently resulted in the description of two new families, Cordycipitaceae and Ophiocordycipitaceae, comprising species previously aligned with Clavicipitaceae (Kepler et al.,

2012; Sung, Sung, Hywel Jones, & Spatafora, 2007). Whereas all three families are dominated by pathogens of insects, the Clavicipitaceae *sensu stricto* include a very large number of plant parasites and plant symbionts. We investigated relationships among the plant-associated Clavicipitaceae by maximum likelihood analysis of aligned sequences from public databases and sequenced genomes. Phylograms were generated based on partial coding sequences of genes for the largest subunit of RNA polymerase II (*rpbA*),  $\beta$ -tubulin (*tubB*), and aldehyde dehydrogenase I (*aldA*) (Fig. 10.1). Results were consistent for all three phylograms. Plant-associated Clavicipitaceae are grouped into clades associated with genera *Epichloë*, *Claviceps*, *Aciculosporium*, *Balansia*, *Periglandula*, and *Villosiclava*. Included in some clades were anamorphs linked with respective teleomorphic genera: *Ephelis* with *Balansia* and *Ustilaginoidea* with *Villosiclava*. Additionally, the anamorph genus *Neotyphodium* has recently been aligned with *Epichloë* (Leuchtman et al., 2014), and the new names are used here (e.g. *E. coenophiala* = *Neotyphodium coenophialum*). Some clades had multiple teleomorphic genera as well. The *Balansia* clade (B) included *Atkinsonella* species, *Heteroepichloë* species, *Myriogenospora atramentosa*, and *Parepichloë cinerea*, and the *Aciculosporium* clade included *Cepsiclava phalaridis*, *Corallocytostroma ornithocopreoides*, and *Neodaviceps monostipa*. Particularly interesting was that the *rpbA* gene indicated a close relationship of *Periglandula* species—which are seed-transmitted symbionts of morning glories—with *Verticillium epiphytum* and *Tyranicordyceps fraticida*, both of which are fungal parasites of other fungi (Kepler et al., 2012).

### 1.2.2 Phylogenetic relationships of pathogenic and symbiotic life histories

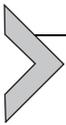
The ergot fungi (*Claviceps* species; clade C; Fig. 10.1) cause highly localized infections of their hosts. They are restricted to the host florets, where they engulf the ovaries, access nutrients, and develop sclerotia, which then drop to the ground to germinate later and complete their development and infection cycle (Pažoutová, Olšovská, Linka, Kolínská, & Flieger, 2000). Thus, the ergot infections are more transient and localized compared to the long-term (i.e. constitutive) systemic infections of plant hosts that typify members of clades A, B, E, and P. Taken together, these clades form a group paraphyletic to clade C, implying that the ability to form constitutive localized infections was ancestral and was lost in the evolution of *Claviceps* species, which may have instead developed a much more efficient process of contagious spread.



**Figure 10.1** Phylogenies inferred from genes for RNA polymerase II largest subunit (*rpba*), aldehyde dehydrogenase I (*aldA*), and  $\beta$ -tubulin (*tubB*). Partial coding sequences were obtained from sequenced genomes and public databases, aligned with MUSCLE (Edgar, 2004), and trees were inferred by maximum likelihood with PhyML implemented by Phylogeny.fr (Dereeper et al., 2008). Species names are followed by isolate identifiers or, if those were unavailable, the GenBank accession numbers.

Coding sequences of *aldA*, *rpbA*, and *tubB* were very similar among *Epichloë* species (clade *E*) (Fig. 10.1). Nevertheless, there was sufficient phylogenetic signal in all three sequence alignments to indicate a root to this clade. The two taxa symbiotic with *Achnatherum inebrians*, *Epichloë gansuensis*, and *E. gansuensis* var. *inebrians* formed a subclade separated by the root from the rest of the sequenced *Epichloë* species.

The characteristics and life cycles of *Epichloë* species are not particularly unique to that genus. The formation of stromata on inflorescences is common in clades *E* and *B* (Fig. 10.1), and highly efficient vertical transmission, which is so important for their mutualistic symbiotic associations, is shared with *Periglandula* species of clade *P*. However, the capability for a delicately balanced symbiosis that mixes stroma production on some tillers and seed transmission on others is common among sexual *Epichloë* species, but so far not characteristic of any other Clavicipitaceae or, for that matter, any other fungi described to date. Therefore, it appears that the features that characterize the *Epichloë* life history and host interactions, as well as most of their alkaloid biosynthetic capability, evolved well before the emergence of that fungal genus, whereas the particular combination characteristics that are crucial to their widespread and often mutualistic symbioses in grasses (Schardl et al., 2008) were selected in the origin of the genus *Epichloë*.



## 2. SEQUENCED GENOMES OF THE CLAVICIPITACEAE

Recent publications (Gao et al., 2011; Schardl, Young, Hesse, et al., 2013; Schardl, Young, Pan, et al., 2013) report genome sequences for a total of 17 genomes of Clavicipitaceae. Here, we report additional sequences, bringing the total to 26 (Table 10.2), including representatives of seven genera (considering *Epichloë* and *Neotyphodium* as a single genus), and 21 species. These, together with a large number of genome sequences from the other families of Hypocreales (Bushley et al., 2013; Cuomo et al., 2007; Güldener et al., 2006; Martinez et al., 2008; Wiemann et al., 2013; Zheng et al., 2011), make this order one of the most intensively studied groups of eukaryotes at the genomic level.

All of the published genome sequences in the Clavicipitaceae are haploids, although the genus *Epichloë* contains a very large number of polyploid hybrids (Leuchtman et al., 2014; Moon, Craven, Leuchtman, Clement, & Schardl, 2004). Comprehensive genome sequencing of the polyploid hybrids has only just begun (Schardl, Young, Pan, et al., 2013) because it requires much more sequence coverage. Nevertheless, a reasonable assembly

**Table 10.2** Statistics for sequenced genomes of plant-symbiotic and plant-parasitic Clavicipitaceae

Organism	Strain	MT	Genome assembly length (bp) <sup>a</sup>	Genes	Total genic (Mb)	Total CDS (Mb)	% CDS	% Non-Rpt-IG	% Rpt	GC proportions			
										Genome	CDS	Non-Rpt-IG	Rpt
<i>Aciculosporium take</i>	MAFF-241224	B	58,707,902	8863	14.0	9.4	16.0	16.9	60.9	0.40	0.59	0.52	0.31
<i>Atkinsonella hypoxylon</i>	B4728	A	35,553,953	9763	15.6	12.7	35.7	32.8	23.4	0.44	0.52	0.48	0.31
<i>Atkinsonella texensis</i>	B6155	A	28,241,655	8878	13.8	12.5	44.2	45.4	8.2	0.50	0.53	0.49	0.36
<i>At. texensis</i>	B6156	B	28,146,212	8890	13.8	12.5	44.4	45.6	8.2	0.50	0.53	0.49	0.36
<i>Balansia obtecta</i>	B249	B	30,388,664	9779	14.9	12.9	42.4	48.0	3.5	0.48	0.53	0.47	0.27
<i>Claviceps fusiformis</i>	PRL 1980	B <sup>b</sup>	52,335,178	9784	19.2	11.9	22.7	30.4	42.3	0.37	0.55	0.41	0.23
<i>Claviceps paspali</i>	RRC-1481	B	28,922,829	8631	14.4	10.0	34.7	37.6	16.9	0.48	0.58	0.48	0.23
<i>Claviceps purpurea</i>	20.1	A	32,108,429	9452	17.5	12.2	39.4	41.6	7.7	0.52	0.55	0.50	0.49
<i>Epichloë amarillans</i>	ATCC 200744	B	37,962,913	10198	15.9	10.9	28.7	23.3	36.8	0.44	0.55	0.49	0.33
<i>E. amarillans</i>	E4668	B	40,693,583	11154	16.9	13.4	32.9	20.6	38.6	0.44	0.53	0.50	0.35

<i>Epichloë aotearoae</i>	MYA-1229	A	34,338,902	11045	17.9	13.1	38.1	28.1	20.4	0.44	0.54	0.49	0.27
<i>Epichloë baconii</i>	ATCC 200745	A	38,004,484	12669	20.5	13.5	35.5	26.9	19.5	0.42	0.53	0.47	0.26
<i>Epichloë brachyelytri</i>	E4804	B	44,051,370	12772	16.5	10.9	24.7	21.0	33.4	0.40	0.54	0.48	0.29
<i>Epichloë coenophiala</i> <sup>c</sup>	e4163	AAA	97,712,391	30268	51.3	36.1	36.9	12.6	35.5	0.43	0.53	0.48	0.36
<i>Epichloë elymi</i>	ATCC 201551	A	31,756,217	8426	15.4	10.8	33.9	29.9	30.4	0.47	0.55	0.50	0.34
<i>Epichloë festucae</i>	E2368	A	34,661,749	8306	15.9	11.1	32.5	24.4	30.6	0.44	0.55	0.48	0.28
<i>E. festucae</i>	F11	B	34,904,508	8649	15.6	10.9	31.9	28.7	26.1	0.44	0.55	0.48	0.28
<i>Epichloë gansuensis</i>	E7080	B	39,525,795	9030	16.6	11.7	29.7	25.7	38.6	0.44	0.54	0.49	0.33
<i>E. gansuensis</i> var. <i>inebrians</i>	MYA-1228	A	29,794,493	9823	16.1	11.2	38.5	35.9	15.0	0.47	0.54	0.49	0.27
<i>Epichloë glyceriae</i>	ATCC 200747	A	49,319,608	11483	20.4	14.9	30.3	24.4	41.0	0.45	0.54	0.49	0.36
<i>Epichloë mollis</i>	AL9924	B	36,111,612	10406	17.0	13.1	36.2	26.1	27.2	0.44	0.54	0.49	0.30

Continued

**Table 10.2** Statistics for sequenced genomes of plant-symbiotic and plant-parasitic Clavicipitaceae—cont'd

Organism	Strain	MT	Genome assembly length (bp)	Genes	Total genic (Mb)	Total CDS (Mb)	% CDS	% Non-Rpt-IG	% Rpt	GC proportions			
										Genome	CDS	Non-Rpt-IG	Rpt
<i>Epichloë typhina</i>	ATCC 200736	A	41,288,070	8584	15.5	10.3	30.9	22.5	44.5	0.42	0.55	0.49	0.28
<i>E. typhina</i> ssp. <i>poae</i>	E5819	A	34,036,313	8770	15.2	10.5	25.1	27.9	32.1	0.43	0.55	0.48	0.24
<i>Periglandula ipomoeae</i> <sup>d</sup>	IasaF13	A <sup>b</sup>	35,301,553	12102	22.5	15.9	45.0	46.4	0.2	0.51	0.53	0.49	0.44

<sup>a</sup>Based on total of scaffolds (supercontigs) or contigs  $\geq 500$  bp.

<sup>b</sup>*C. fusiformis* PRL 1980 mating-type genes include *mtBA* and *mtAC*. *P. ipomoeae* IasaF13 mating-type genes *mtAA* and *mtAC* appear to have premature stop codons.

<sup>c</sup>Repeat statistics for the hybrid *E. coenophiala* were determined by masking the genetic regions prior to determining the repeat content.

<sup>d</sup>Statistics for *P. ipomoeae* are tentative because the assembly was filtered by selecting only contigs containing tBLASTx matches to genome sequences from other Clavicipitaceae.

Abbreviations: CDS, coding sequence; MT, mating type; non-Rpt-IG, nonrepetitive intergenic DNA; Rpt, repetitive DNA; GC, proportion of sequence that is G or C.

has been obtained for *Epichloë coenophiala*, an economically important hybrid endophyte with three ancestral genomes. This particular fungus has considerable economic and ecological importance as the common endophyte of tall fescue, which is widely distributed on several continents, and the most widely planted forage grass in the United States (Rudgers, Holah, Orr, & Clay, 2007; Schardl, Scott, Florea, & Zhang, 2009). Haploid genome sizes (Table 10.2) vary roughly twofold from just under 30 Mb for *Atkinsonella texensis*, *Claviceps paspali*, and *Epichloë gansuensis* var. *inebrians* to nearly 60 Mb for *Aciculosporium take* and *Claviceps fusiformis*. However, genetic content does not correlate with genome size in this group of fungi, which all have approximately the same number of genes. Genome size differences are almost entirely due to repeat DNA content. There is no obvious taxonomic or phylogenetic pattern associated with variation in genome size or repeat content. However, in the *Epichloë* species, alkaloid loci tend to have large blocks of repeat sequences interspersed among the genes, whereas the members of other genera have far less repeat sequence within those loci (Schardl, Young, Hesse, et al., 2013).

## 2.1. Repeat content

The estimated proportion of repetitive sequences for each genome was found to vary, ranging from the fewest in the *P. ipomoeae* (0.2%) genome, which could be a feature of the assembly, to the most in *A. take* (60.9%) (Table 10.2). Only five genomes, *P. ipomoeae*, *B. obtecta*, *At. texensis* B6155, *At. texensis* B6156, and *C. purpurea*, contained repeat content less than 10% of the total genome. The GC proportion of the repetitive sequences was low, ranging from 0.24 to 0.36 (average of 0.3 GC), apart from *C. purpurea* (0.49) and *P. ipomoeae* (0.44). The comparison of repeat sequences within a genome has indicated those with low GC content contain many C to T and G to A transitions that are likely due to repeat-induced point mutations (RIPs) (Fleetwood, Scott, Lane, Tanaka, & Johnson, 2007; Schardl, Young, Hesse, et al., 2013; Young et al., 2005). The evaluation of repeat sequences within the *Epichloë* species indicates there are more class I retrotransposon elements than class II DNA transposons. However, it is unlikely that any of the transposable elements are still functional since they have been rendered highly degenerate due to their extensive mutations.

The fragmented nature of the assemblies (largely due to AT-rich repeats) makes it difficult to determine the repeat distribution and integration bias patterns. However, it is very apparent that repeat sequences are

overrepresented in the alkaloid biosynthesis clusters of *Epichloë* species compared to most other SM clusters and to alkaloid clusters in most members of the other genera (Schardl, Young, Hesse, et al., 2013). The extensive repeat blocks are likely to have impacted alkaloid cluster stability (Schardl, Young, Hesse, et al., 2013) (discussed later). Also prevalent in the genomes of *Epichloë* species, including their alkaloid biosynthesis clusters, are miniature inverted-repeat transposable elements (MITEs), particularly in gene promoter regions (Fleetwood et al., 2011; Schardl, Young, Hesse, et al., 2013). It appears that gene clusters have rearranged due to repetitive elements and genes being rendered nonfunctional because of integrating elements. Interestingly, although *At. texensis* has markedly fewer repeats than the *Epichloë* species, the *IDT* clusters found in the *At. texensis* isolates appear to also have been fragmented by repeat sequences.

## 2.2. Telomeric regions

The telomere repeat sequence, (TTAGGG)<sub>n</sub> when read towards the chromosome end, was identified as tandem repeats at the start or end of contigs in many of the sequenced genomes. Contigs that contained at least two copies of a telomere repeat were counted to determine the likely number of chromosome ends. The average chromosome number of nonhybrid and hybrid *Epichloë* species was 4.5 and 8.75, respectively. The predicted chromosome numbers for the two *E. festucae* isolates and *E. coenophiala* are consistent with data generated from chromosome separations using pulsed-field gel electrophoresis and from Southern-blot hybridizations to telomere repeat probes (Kuldau, Tsai, & Schardl, 1999; Schardl, C. L., Young, C. A., & Andreeva, K., unpublished data). Unfortunately, telomere repeat sequences were underrepresented in the genomes of *C. purpurea* 20.1, *Fusarium graminearum* PH-1, the two *Metarhizium* species, and some *Epichloë* species (namely, *E. typhina* E8, *E. elymi* E56, and *E. bromicola* E502).

Many of the predicted telomeres were contained on contigs with large subtelomeric AT-rich repeat regions. The telomere-linked *recQ* helicase (*TLH*) genes, commonly associated within the telomeres of *Magnaporthe oryzae* (Rehmeier et al., 2006), did not appear to be associated with subtelomeric regions in the Clavicipitaceae. In fact, the total number of *recQ* genes identified in clavicipitaceous species was limited to one to four copies. Subterminal genetic regions did not appear to be shared across the Clavicipitaceae, but among the *Epichloë* species, the *EAS* and *IDT* clusters were often identified as subterminal (Schardl, Young, Hesse, et al., 2013).

### 2.3. Mating-type loci

By far, the most common genetic system governing mating types in Pezizomycotina (filamentous ascomycetes) involves one locus with two idiomorphs designated *MAT1-1* and *MAT1-2* (Turgeon & Yoder, 2000). Finding that nomenclature a bit cumbersome, we refer to the locus as *MT*, the idiomorphs as *MTA* and *MTB*, and the genes as *mtAA*, *mtAB*, and *mtAC* for the three genes of the *MTA* idiomorph and *mtBA* for the sole gene of *MTB*. With some exceptions (Vaillancourt, Du, Wang, Rollins, & Hanau, 2000), obligately outcrossing (heterothallic) species have a haploid genome with either *MTA* or *MTB* at the *MT* locus, whereas self-compatible (homothallic) species have linked *MTA* and *MTB* (Turgeon & Yoder, 2000).

Inspection of the sequenced genomes suggests that most Clavicipitaceae are heterothallic (obligately outcrossing), because their *MT* loci have either one idiomorph (*MTA*) or the other (*MTB*) (Table 10.2). Two strains seem exceptional: *Claviceps fusiformis* PRL 1980 has an *mtBA* gene and an *mtAC* gene. In contrast, *Periglandula ipomoeae* IasaF13 has premature stops in *mtAA* and *mtAC*. It is possible that neither of these strains is competent for sexual crosses. The *C. fusiformis* strain has been propagated clonally in culture for many decades, as a model for studying ergot alkaloid biosynthesis (Gröger & Floss, 1998), and no teleomorph (sexual state) is known for *P. ipomoeae*. *Balansia obtecta* B249 also appeared exceptional in that the *apnB* (DNA lyase) gene adjacent to *MTB* is a pseudogene (Schardl, C. L., unpublished data). Apparently, functional *apnB* genes are almost always located adjacent to *MT* loci in Pezizomycotina.

Hybrid *Epichloë* species are generally asexual (Moon et al., 2004), with the reported exception of *E. liyangensis* (Kang et al., 2011), yet they generally possess *MT* idiomorphs. There is no general pattern of the *MT* idiomorphs in hybrids. *Epichloë coenophiala* has an *MTA* idiomorph from each of its three ancestors (Takach & Young, 2014), the *E. festucae* x *E. typhina* hybrid strain Lp1 has *MTB* idiomorphs from both of its ancestors (unpubl. data of the authors), and *MTA* and *MTB* genotypes are found in *E. canadensis* (Charlton, Craven, Mittal, Hopkins, & Young, 2012), *E. uncinata*, and *Epichloë* sp. FaTG-4 from decaploid *Lolium arundinaceum* (Takach & Young, 2014).

### 2.4. Mitochondrial genomes

Complete mitochondrial genomes (mtDNA) were identified in the genome assemblies of *E. festucae* E2368 and *E. bromicola* E502 (Young, C. A. &

Schardl, C. L., unpublished data), and the mtDNA sequence was nearly complete in the *C. purpurea* 20.1 assembly (accession number FO082257). The A+T content in mtDNA of the two *Epichloë* strains and *C. purpurea* was 73% and 65%, respectively, consistent with other published sequences (Pantou, Kouvelis, & Typas, 2008). In the mtDNA of both *E. festucae* E2368 and *E. bromicola* E502, 15 protein-coding genes were identified, as were genes for the small- and large-subunit rRNAs of the mitochondrial ribosomes and for 27 tRNAs (predicted by tRNAscan-SE; Lowe & Eddy, 1997). The genes were ordered consistently with the mtDNA of Hypocreales when compared to the other members of Sordariomycetes (Pantou et al., 2008). The assembled *C. purpurea* mtDNA contained the same genes as the *Epichloë* isolates, except the *atp9* gene was absent, and although *nad4* was present, it was no longer adjacent to *nad1*. It is yet to be determined if these are real differences or an artefact of incomplete assembly of the *C. purpurea* mtDNA.

Mitochondrial genome sizes are variable in fungi, and this is also reflected in the Clavicipitaceae where *M. anisopliae* mtDNA (accession number NC\_008068; Ghikas, Kouvelis, & Typas, 2006) is 24,673 bp, *C. purpurea* mtDNA is estimated at 55,537 bp, *E. bromicola* E502 mtDNA is 69,466 bp, and *E. festucae* mtDNA is the largest at 72,701 bp. The size differences between these genomes are largely due to variations of intron number and size and the presence of a large variable region found in E2368 and E502 associated with the tRNA cluster containing 12 tRNA genes flanked by *rnl/rps3* and *nad2*. The variable region in E2368 and E502 extends over 10 kb and appears to contain sequence similarity to *dpoA*, a DNA-directed RNA polymerase gene encoded on the opposite strand to the other mitochondrial genes. Multiple stop codons present throughout the coding region of *dpoA* suggest the gene is nonfunctional. In many fungi, *dpoA* genes not only are often found on a linear plasmid but also can be stably integrated into the mitochondrial genome in single or multiples sites (Formighieri et al., 2008). *Fusarium* species also contain a large variable coding region between *rnl/rps3* and *nad2*, but its encoded function is yet to be elucidated and it does not appear to be related to *dpoA* (Al-Reedy, Malireddy, Dillman, & Kennell, 2012).

## 2.5. Gene ontology categories

We compared the representation of genes across higher-order gene ontology (GO) terms for four fungi with annotated genome sequences, the symbiont *E. festucae* E2368, the grass pathogens *Fusarium graminearum* PH-1 and

*Magnaporthe oryzae* 70-15, and the saprophyte *Neurospora crassa* OR74A. All GO annotations were downloaded from Ensembl using BioMart. Then, the GO hierarchy was determined for each term. Level 1 nodes in the hierarchy are the three roots, biological process, molecular function, and cellular component. Level 2 nodes are the first-generation children of the level 1 nodes. Table 10.3 lists the proportion and the number of genes having a GO annotation linked to each level 2 GO term for the molecular function and biological process roots. Only the 10 most frequent categories are displayed.

Despite their different ecological niches and host interactions and differences in total gene numbers, the four fungi showed little difference in

**Table 10.3** Gene ontology (GO) categories of inferred genes from genomes of *Epichloë festucae* E2368 (*Ef*), *Fusarium graminearum* PH-1 (*Fg*), *Magnaporthe oryzae* 70-15 (*Mo*), and *Neurospora crassa* OR74A (*Nc*)<sup>a</sup>

GO id and term	<i>Ef</i>		<i>Fg</i>		<i>Mo</i>		<i>Nc</i>	
	No.	%	No.	%	No.	%	No.	%
<b>Biological process</b>								
GO:0008152 metabolic process	4768	58	4437	59	4686	61	2929	58
GO:0051234 establishment of localization	1220	15	1341	18	1314	17	852	17
GO:0065007 biological regulation	813	10	706	9	586	8	445	9
GO:0044699 single-organism process	596	7	566	8	500	6	483	10
GO:0009987 cellular process	543	7	222	3	407	5	144	3
GO:0071840 cellular component organization or biogenesis	150	2	138	2	136	2	117	2
GO:0050896 response to stimulus	75	1	73	1	64	1	75	1
GO:0051704 multiorganism process	14	<1	9	<1	7	<1	2	<1
GO:0048511 rhythmic process	2	<1	0	0	0	0	1	<1
GO:0040007 growth	1	<1	1	<1	1	<1	1	<1
Total	8182	100	7493	100	7701	100	5049	100
<b>GO id and term</b>	<b><i>Ef</i></b>		<b><i>Fg</i></b>		<b><i>Mo</i></b>		<b><i>Nc</i></b>	
<b>Molecular function</b>	<b>No.</b>	<b>%</b>	<b>No.</b>	<b>%</b>	<b>No.</b>	<b>%</b>	<b>No.</b>	<b>%</b>
GO:0003824 catalytic activity	7302	54	5709	46	6876	48	2771	40

Continued

**Table 10.3** Gene ontology (GO) categories of inferred genes from genomes of *Epichloë festucae* E2368 (*Ef*), *Fusarium graminearum* PH-1 (*Fg*), *Magnaporthe oryzae* 70-15 (*Mo*), and *Neurospora crassa* OR74A (*Nc*)—cont'd

GO id and term	<i>Ef</i>		<i>Fg</i>		<i>Mo</i>		<i>Nc</i>	
	No.	%	No.	%	No.	%	No.	%
GO:0005488 binding	5273	39	5315	43	6309	44	3453	50
GO:0005215 transporter activity	344	3	458	4	420	3	242	3
GO:0001071 nucleic acid binding transcription factor activity	253	2	369	3	235	2	157	2
GO:0005198 structural molecule activity	149	1	141	1	143	1	143	2
GO:0009055 electron carrier activity	102	1	176	1	185	1	92	1
GO:0030234 enzyme regulator activity	62	<1	54	<1	53	<1	44	1
GO:0060089 molecular transducer activity	59	<1	59	<1	47	<1	40	1
GO:0004872 receptor activity <sup>b</sup>	52	<1	0	0	21	<1	8	<1
GO:0000988 protein binding transcription factor activity	26	<1	18	<1	20	<1	6	<1
Total	13622	100	12299	100	14309	100	6956	100

<sup>a</sup>Note that some genes may be assigned to multiple GO categories, and unassigned genes are not listed.

<sup>b</sup>The large differences in this category appear to be an artefact of differences in gene annotations by curators of the respective genomes.

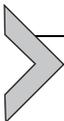
proportions of genes represented by the 20 GO terms that we compiled. The apparent difference in receptor activity genes was likely an artefact of differences in the original annotations by curators. None of the other terms showed obvious differences in representation. Nevertheless, it will be worthwhile in the future to conduct analyses at other levels of GO terms and among a broader range of fungi.

## 2.6. Variation of SM clusters

The repertoire of genes and gene clusters associated with potential SM production were determined using the SMURF and antiSMASH pipelines

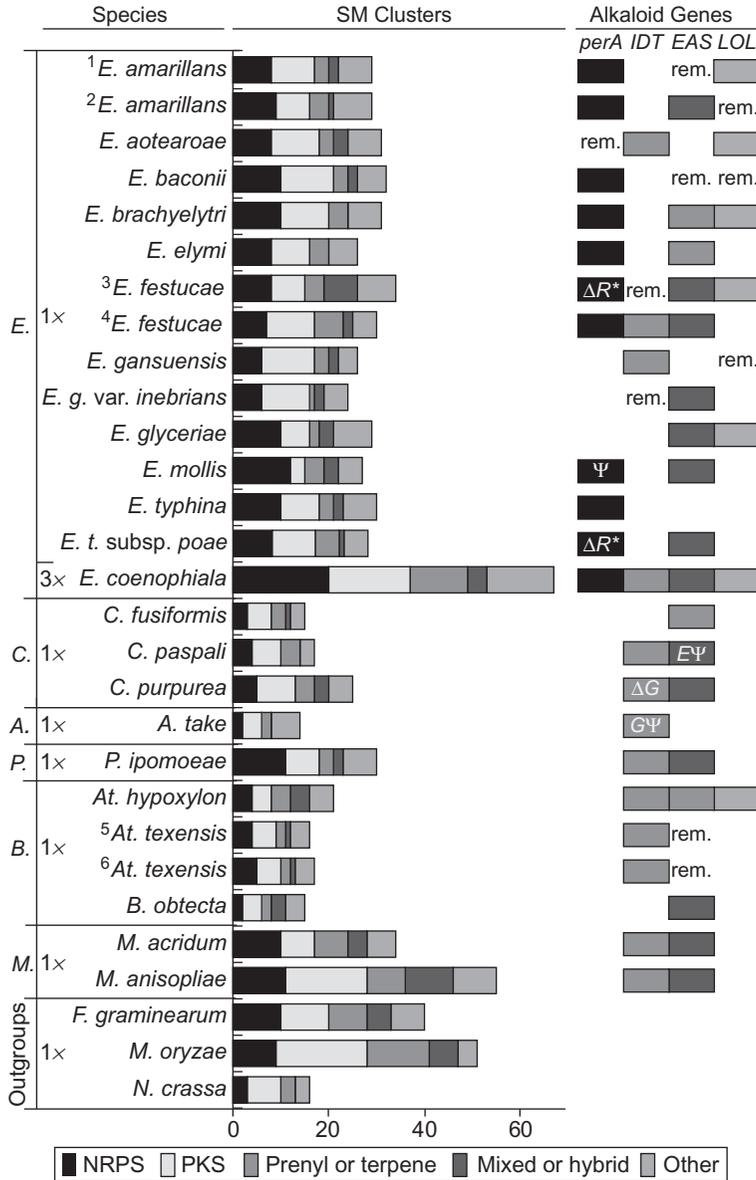
(Blin et al., 2013; Khaldi et al., 2010). Similar trends were seen with either analysis, and Fig. 10.2 presents antiSMASH results supplemented with manual identification of those alkaloid clusters missed by antiSMASH and SMURF. These programs failed to detect the *LOL* clusters due to a lack of key SM signature genes at this locus. They also failed to detect *idtG* from the *IDT* clusters, although antiSMASH often associated the *idtG* paralog, *ggsA* (required for primary metabolism production of geranylgeranyl diphosphate), with SM cluster prediction. The classification of the clusters is based on predicted functions that are considered signatures of SM (Bok et al., 2006), although sometimes orthologous clusters can be classified differently. In particular, *EAS* clusters for the production of the simple ergot alkaloids such as chanoclavine are placed in the prenyl class, while those clusters required for more complex ergot alkaloids are considered in the mixed class because both prenyl (DmaW) and NRPS (LpsA and LpsB) functionality are required. In addition, the fragmented nature of the *EAS* and *IDT* clusters also affected correct SM identification and class placement of these clusters.

Numbers of SM clusters identified within the *Epichloë* nonhybrid (1x ploidy) species appeared consistent, averaging 29 (range = 24–35), although more variation could be seen among individual classes. As expected, this number dramatically increased in the hybrid, *E. coenophiala*, whereby 69 clusters were identified, consistent with its higher (3x) ploidy level. Generally, the *Epichloë* species show an increase in SM clusters compared to members of related clades, *C*, *B*, and *A*, but not *P*. Based on fungal lifestyle, the fungal clades could be ranked fewest to most SM clusters, whereby saprophytes (e.g. *N. crassa*) contained fewest, followed by replacement parasites (both systemic and ergot; clades *C*, *B*, and *A*), then seed-transmitted endophytes (clades *E* and *P*), and finally “classic” pathogens of plants and animals (e.g. *Metarhizium* spp., *Fusarium graminearum*, and *Magnaporthe oryzae*). We speculate that in plant-associated Clavicipitaceae, an increase in SM clusters found in seed-transmitted endophytes represents selection for increased host protection in both *E* and *P*.



### 3. ALKALOID GENE LOCI

On a broad scale, there is considerable diversity of alkaloid profiles at all taxonomic levels from species on up. For example, different strains of *Epichloë festucae* can produce any of the four classes of alkaloids—ergot alkaloids, indole-diterpenes, lolines, and peramine—in various combinations, although only some strains of *E. coenophiala* (including e4163) are



**Figure 10.2** Numbers of predicted specialized metabolite (SM) clusters by class and fungal strain. SM classes are based on signature enzyme families as follows: NRPS, non-ribosomal peptide synthase; PKS, polyketide synthase; prenyl or terpene, prenyl transferase or terpene synthase/cyclase; mixed or hybrid, clusters with a mixture of classes or hybrid NRPS/PKS; and others, known SM clusters lacking the aforementioned families. Cluster numbers were adjusted to account for each of the known alkaloid gene clusters. Specifically, the *LOL* clusters that were not identified by antiSMASH were added into the

known to produce all four (Schardl et al., 2012; Schardl, C. L., Young, C. A., Dinkins, R. D., Nagabhyru, P., & Panaccione, D. G., unpublished data). Comparing the syntenic loci among strains, it is apparent that those that lack any particular alkaloid class most often lack some or all of the genes at the corresponding locus. At least three of the four alkaloid classes (ergot alkaloids, indole-diterpenes, and lolines) exhibit considerable variation in structures, and such variation has implications for spectra and modes of activity (Pan et al., 2014; Schardl et al., 2006, 2012; Schardl, Young, Hesse, et al., 2013; Young et al., 2009).

### 3.1. Relationships between gene contents and alkaloid structures

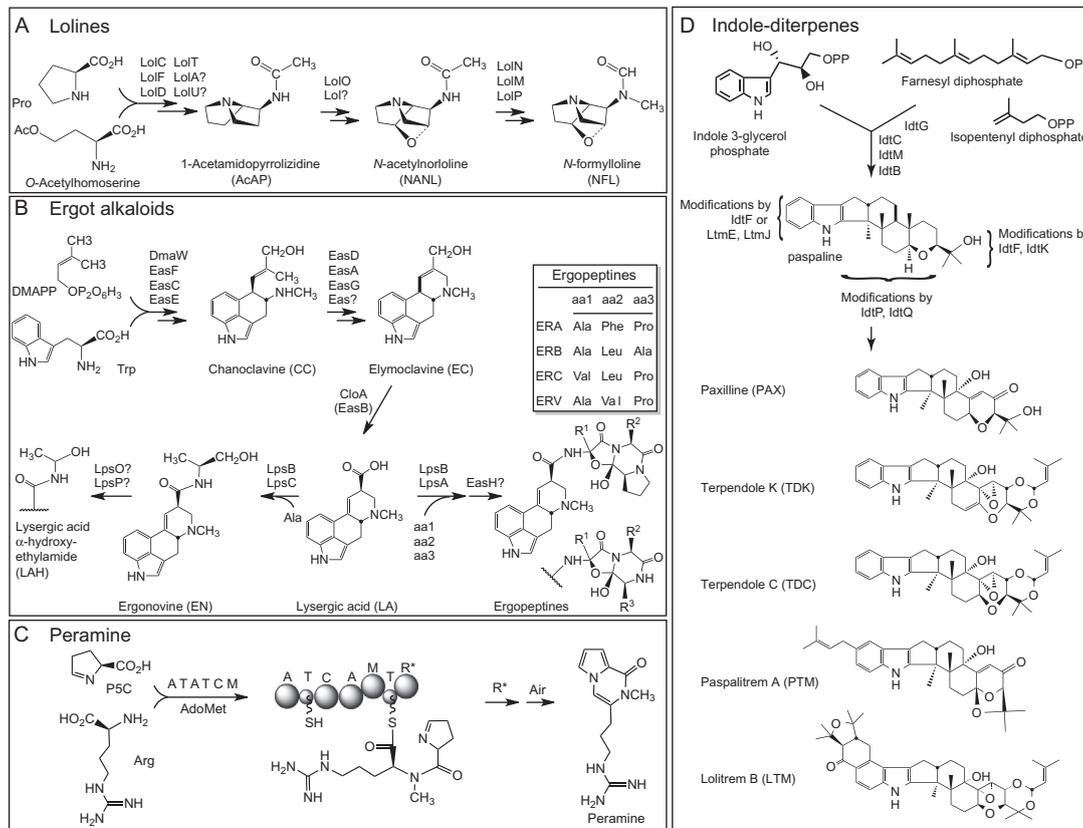
#### 3.1.1 Ergot alkaloids

Ergot alkaloids are known from a wide range of filamentous ascomycetes (Schardl et al., 2006). Simpler ergot alkaloids called clavines are known from several orders, but the Clavicipitaceae can produce an extraordinarily diverse range of ergot alkaloids from simple clavines such as chanoclavine I to much larger and more complex amides of lysergic acid (Fig. 10.3). The most complete ergot alkaloid biosynthesis (*EAS*) gene cluster has been identified in the morning glory symbiont, *Periglandula ipomoeae*, comprising 14 genes (Fig. 10.4). This endophyte produces the ergopeptine, ergobalansine, and the simpler lysergic acid amides, ergonovine and D-lysergic acid  $\alpha$ -hydroxyethylamide (LAH). The latter two metabolites are also associated with *Claviceps paspali*, which shares with *P. ipomoeae* the genes designated *lpsC*, *easO*, and *easP*. (However, the sequenced *C. paspali* strain has an inactive *easE* gene, so it should produce only one or two early intermediates of the ergot alkaloid pathway.)

The presence of *lpsC* in genomes of *C. purpurea* and *Epichloë gansuensis* var. *inebrians* is in keeping with the finding that the LPS3 subunit encoded by this gene is required, together with the *lpsB* gene product (LPS2), for ergonovine biosynthesis (Ortel & Keller, 2009). The LPS2 subunit also is required together with the *lpsA* gene product, LPS1, in the biosynthesis

---

others class, while the *EAS* and *IDT* clusters split by repeat sequences were condensed into their correct SM classes. The known alkaloid biosynthesis genes are indicated on the right with shaded boxes indicating SM class. Pseudogenes ( $\Psi$ ), gene or domain deletions ( $\Delta$ ) that preclude production of the alkaloids, and cluster remnants (rem.) are also indicated. Genome ploidies are given as 1x or 3x. Where multiple strains of the same species were sequenced, they are distinguished by superscript numbers as follows: 1, ATCC 200744; 2, E4668; 3, E2468; 4, F11; 5, B6155; and 6, B6156.



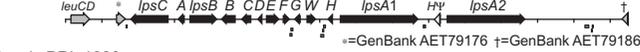
**Figure 10.3** Alkaloid brief pathways. Summaries of biosynthetic pathways for four classes of alkaloids produced by Clavicipitaceae. (A) Loline alkaloids. (B) Ergot alkaloids. (C) Peramine. (D) Indole-diterpenes. Double arrows indicate multiple steps, and arrows are labelled according to the gene products (enzymes) that catalyse the steps or, in the case of peramine biosynthesis (panel C), the individual domains of the multifunctional peramine synthetase (*perA* gene product).

## Species and strains

*Periglandula ipomoeae* lasaF13

Alkaloids

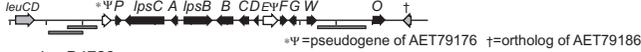
LAH EN ERB

*Claviceps purpurea* 20.1

ENERA ERC

*Claviceps fusiformis* PRL 1980

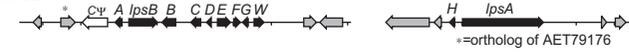
EC

*Claviceps paspali* RRC-1481

MeDMAT

*Atkinsonella hypoxylon* B4728

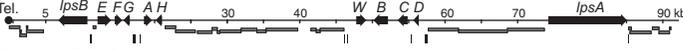
(CC)

*Balansia obtecta* B249

ERB

*Epichloë gansuensis* var. *inebrians* e818

LAH EN

*Epichloë festucae* F11

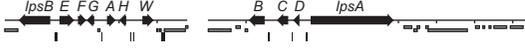
ERV

*Epichloë typhina* E5819

ERV

*Epichloë amarillans* E4668

(ERV)

*Epichloë mollis* AL9924

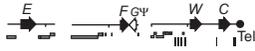
ERV

*Epichloë glyceriae* E277

(ERV)

*Epichloë brachyelytri* E4804

CC

*Epichloë elymi* E56

CC

*Epichloë coenophiala* e4163

EAS1 (left)

EAS1 (right)

ERV



**Figure 10.4** Structures of representative *EAS* loci and variation in flanking sequences. Associated ergot alkaloids are indicated at the right, abbreviated as indicated in Fig. 10.3 or as MeDMAT for *N*-methyl-dimethylallyltryptophan. The clusters have been ordered to show similarity of cluster gene order and orientation and common flanking genes. *EAS* genes are represented with black arrows and those labelled by single letters are as follows: A, *easA*; B, *cloA*; C, *easC*; D, *easD*; E, *easE*; F, *easF*; G, *easG*; H, *easH*; O, *easO*; P, *easP*; and W, *dmaW*. Other *EAS* genes, encoding nonribosomal peptide synthetase subunits, are *IpsA*, *IpsB*, and *IpsC*. Pseudogenes, white arrows; flanking genes, grey arrows; and telomeres (tel.), filled black circles. Depicted beneath each map are horizontal bars indicating transposon-derived repeats and vertical bars indicating miniature inverted-repeat transposable elements (MITEs).

of ergopeptine alkaloids. Variation in LPS1 determines which three amino acids are incorporated into the ergopeptide lactam (Riederer, Han, & Keller, 1996), which is then converted by EasH to the corresponding ergopeptine (Havemann, Vogel, Loll, & Keller 2014). Thus, *Periglandula* and *Balansia* species have been found to produce ergobalansine, and *Epichloë* species have been found to produce ergovaline (Schardl, Young, Hesse, et al., 2013) (Fig. 10.3). The two forms of *lpsA* in *C. purpurea* 20.1 are responsible for two different ergopeptines, ergotamine and ergocryptine (Haarmann, Lorenz, & Tudzynski, 2008). In all, 20 different amino acid combinations are evident in natural ergopeptines (some of which also are known in dihydroergopeptines) (Cheng, Coyle, Panaccione, & O'Connor, 2010; Schardl et al., 2006).

While *P. ipomoeae* is the only species so far found to have all 14 of the *EAS* genes known in Clavicipitaceae (although the fumigaclavine biosynthesis gene cluster of *Aspergillus fumigatus* has additional genes; Coyle & Panaccione, 2005), other ergot alkaloid producers lack functional copies of between 2 and 10 of these genes. So, in the extreme cases, only *dmaW*, *easF*, *easC*, and *easE* remain, and the fungus produces chanoclavine I as its pathway end product (Schardl, Young, Hesse, et al., 2013). Remnants of some of *EAS* genes in such reduced clusters indicate that gene loss is a major evolutionary process in ergot alkaloid diversification. Other processes that diversify these profiles include variation in *lpsA*, as discussed earlier in the text, and variation in *easA* associated with the production of either the classical 8–9 unsaturated ergot alkaloids or the dihydroergot alkaloids with a saturated D ring (Coyle, Cheng, O'Connor, & Panaccione, 2010).

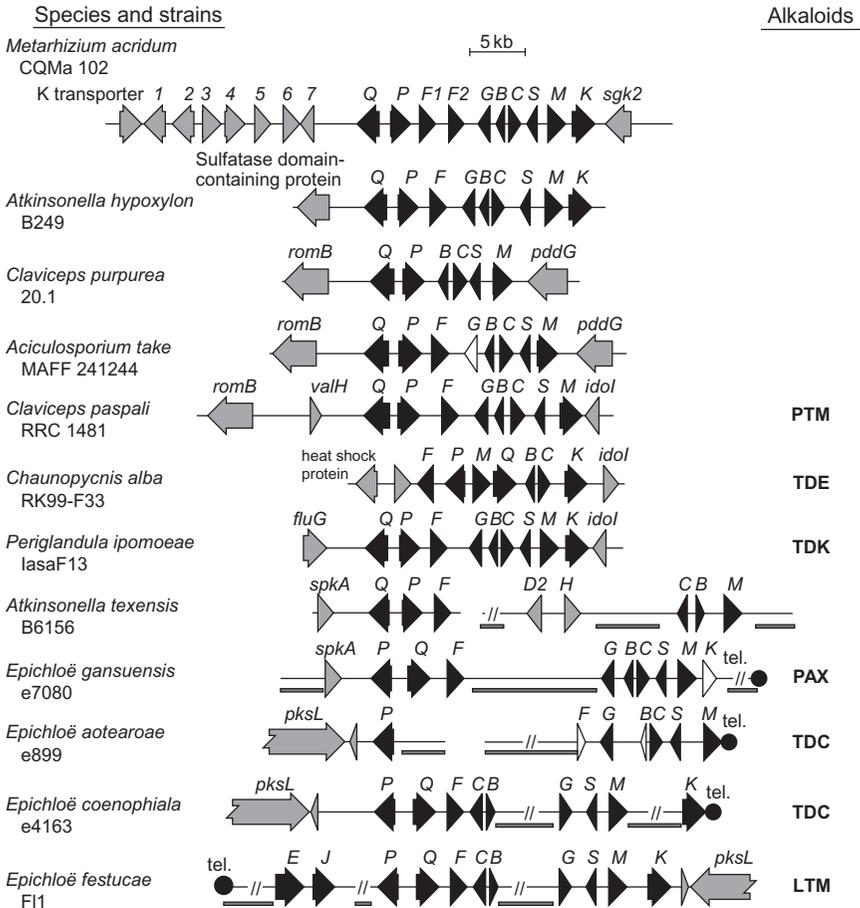
### 3.1.2 Indole-diterpenes

Also produced by a range of filamentous ascomycetes, the indole-diterpenes represent a diverse array of compounds synthesized from indole-3-glycerol phosphate and geranylgeranyl diphosphate and followed by various modifications such as prenylations, epoxidations, and hydroxylations to provide structural diversity (Fig. 10.3) (Saikia et al., 2012). The reconstitution of the *Penicillium paxilli* paxilline biosynthetic pathway in both homologous and heterologous systems has clearly established the early pathway steps to the simple indole-diterpene, paspaline (Saikia, Parker, Koulman, & Scott, 2006; Tagami et al., 2013). Four gene products are required to produce paspaline: IdtG to form geranylgeranyl diphosphate, IdtC to form 3-geranylgeranyl indole, and followed by IdtM and IdtB for the epoxidation and cyclization steps, respectively. Beyond paspaline production,

the indole–diterpene pathway can be more complex and pathway diversity is determined by both the presence and absence of *IDT* genes and variations in substrate and product specificity of pathway-specific gene products.

A key feature of *IDT* clusters (Fig. 10.5) is usually the presence of an *IDT* pathway-specific geranylgeranyl diphosphate synthase encoded by *idtG*, whereas a paralog, *ggsA*, is located elsewhere in the genome and specifies geranylgeranyl diphosphate synthase for primary metabolism. Conceivably, *ggsA* could serve both roles in some fungi, but genetic tests in *P. paxilli* demonstrate a requirement of *paxG* (ortholog of *idtG/ltmG*) for indole-diterpene biosynthesis (Saikia et al., 2006; Young, McMillan, Telfer, & Scott, 2001). Interestingly, an *IDT* gene cluster was recently sequenced from *Chaunopycnis alba* (which shares an ordinal relationship with the Clavicipitaceae), which is capable of producing terpendole E despite the absence of an identifiable *idtG* gene in the cluster (Motoyama, Hayashi, Hirota, Ueki, & Osada, 2012). However, without a complete sequence for *Ch. alba*, it remains unknown if an *idtG* ortholog may be present elsewhere in its genome. It is also conceivable that some fungi can utilize the housekeeping *ggsA* gene to direct the first step in the pathway. Sequenced genomes of some Clavicipitaceae have revealed *IDT* clusters that lack *idtG* (in *C. purpurea* 20.1 and *A. texensis* B6155 and B6156) or are predicted to contain a nonfunctional *idtG* (*A. take* MAFF-241224), but whether any of these clusters is functional is currently unknown.

Lolitrems B production is one of the most well-studied indole–diterpene pathways, elucidated based on the identification of intermediates present in perennial ryegrass (*Lolium perenne*) infected with *E. festucae* var. *festucae* and var. *lolii* (Munday-Finch, Wilkins, & Miles, 1998) and from gene deletion and heterologous expression experiments (Saikia et al., 2006, 2012; Young et al., 2006, 2009). In total, 11 genes are found at the *IDT/LTM* locus for *Epichloë* species capable of producing lolitrems B (Schardl, Young, Hesse, et al., 2013; Young et al., 2006). However, the production of lolitrems B occurs via a very complex metabolic grid apparently lacking any specific order of oxygenations (catalysed by IdtP and IdtQ) and prenylations (catalysed by IdtF, IdtK, IdtE, and IdtJ) of the geranylgeranyl and indole moieties. Such a metabolic grid gives a much larger number of intermediate products than would a direct linear pathway (Saikia et al., 2012). The *ltmE* and *ltmJ* genes required for the indole–ring prenylation steps seem to have arisen by duplications and neofunctionalizations of other *IDT* genes (Schardl, Young, Hesse, et al., 2013).



**Figure 10.5** Structures of representative *IDT* loci and variation in flanking sequences. Associated indole-diterpenes are indicated at the right, abbreviated as indicated in Fig. 10.3 or as TDE for terpendole E (Motoyama et al., 2012). The clusters have been ordered to show similarity of cluster gene order and orientation and common flanking genes. *IDT* genes are represented with black arrows and genes are labelled by single letters, whereby Q = *idtQ*, *ltmQ* (*E. festucae*), or *terQ* (*Ch. Alba*; accession AB725916), etc. The *IDT* genes in the *M. acridum* cluster (accession GL698534) (representing a *Metarhizium IDT* example) have been renamed to represent homologs present in other Clavicipitaceae. The *M. acridum* genes labelled 1–7 have predicted functions possibly associated with SM. The *IDT* genes from B6156 were chosen as the *At. texensis* example as B6155 was identical but appeared to lack *idtF*. Additional *IDT* genes, likely required for TDC production, are present in e899 but each gene is embedded in AT-rich regions (Schardl, Young, Pan, et al., 2013). Pseudogenes, white arrows; flanking genes, grey arrows; telomeres (tel.), filled black circles; and //, regions of the cluster containing repetitive elements.

Other variations within the *IDT* clusters have revealed that *idtF* is sometimes nonfunctional and *idtK* can be absent or nonfunctional, leading to the accumulation of terpendoles (Schardl, Young, Hesse, et al., 2013; Schardl, Young, Pan, et al., 2013; Young et al., 2009). Additional indole-diterpene pathway variation can also stem from specificity differences of the P450 monooxygenases *IdtP* and *IdtQ* (Saikia et al., 2012). Recently, *IDT* gene clusters have been identified in genome sequences of a number of other Clavicipitaceae, namely, *M. acridum*, *M. anisopliae* (Schardl, Young, Hesse, et al., 2013), *At. hypoxylon*, and *At. texensis* (Young, C.A., and Schardl, C.L. unpublished data), but indole-diterpene products are yet to be identified from these species.

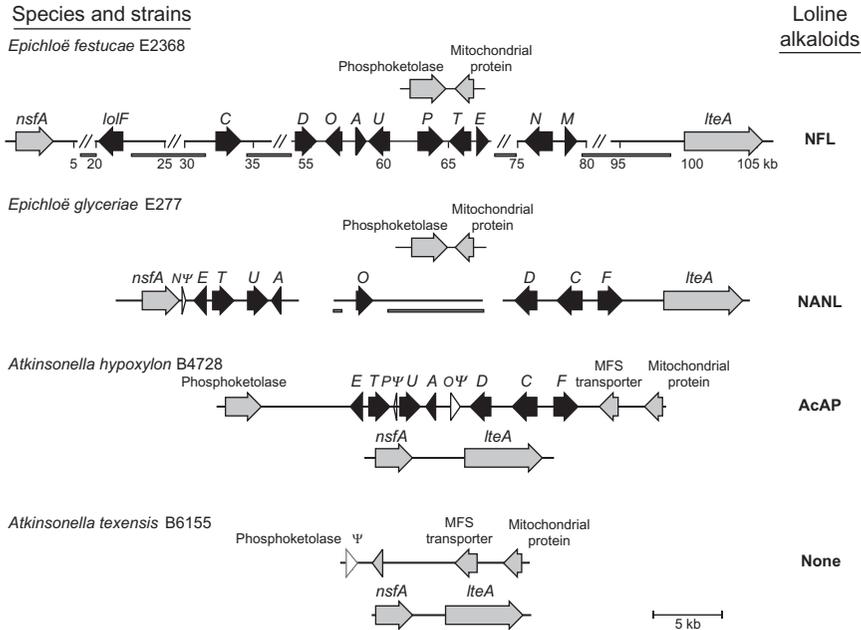
### 3.1.3 *Lolines and aminopyrrolizidines*

Loline alkaloids and related aminopyrrolizidines (Fig. 10.3A) are known from various cool-season grasses, legume genera *Adenocarpus* (Veen, Greinwald, Canto, Witte, & Czygan, 1992) and *Laburnum* (Tasso, Novelli, Sparatore, Fasoli, & Gotti, 2013), and a species of morning glory, *Argyrea mollis* (Tofern, Kaloga, Witte, Hartmann, & Eich, 1999). Orthologous clusters of genes, constituting the *LOL* locus, have been identified in several *Epiclloë* species (including several former *Neotyphodium* species) (Schardl, Young, Hesse, et al., 2013) and recently in an *At. hypoxylon* isolate that was confirmed to produce 1-acetamidopyrrolizidine (Pan et al., 2014). The full cluster (Fig. 10.6) has 11 genes, but inactivating mutations or deletions of certain late-pathway genes are associated with natural variations in aminopyrrolizidine profiles (Pan et al., 2014; Schardl, Young, Hesse, et al., 2013). Those that accumulate 1-acetamidopyrrolizidine as the end product lack functional *lolO* genes; those that accumulate *N*-acetylnorloline lack *lolN*, *lolM*, and *lolP*; and those with the full cluster accumulate *N*-formylloline.

Four strains have been identified that, due to the mutation of *lolO*, accumulate 1-acetamidopyrrolizidine as an end product (Pan et al., 2014). These are strains of *E. elymi*, *E. brachyelytri*, *E. canadensis*, and *A. hypoxylon*. All possess only the *EAS* genes for chanoclavine I production. It is intriguing to speculate that this pattern results from positive selection for the combination of 1-acetamidopyrrolizidine and chanoclavine I, perhaps by synergistic action against herbivores.

### 3.1.4 *Peramine*

Among known SMs, peramine (Fig. 10.3) has a nearly unique pyrrolopyrazine structure bearing some similarity to the pyrazinone linkers in



**Figure 10.6** Evidence that the *LOL* cluster moved after being constituted during the evolution of plant-associated Clavicipitaceae. The *LOL* clusters in *Epichloë festucae* E2368 and *Epichloë glyceriae* E277 are located between housekeeping genes *nsfA* and *lteA*, but in opposite orientations. The *LOL* cluster in *Atkinsonella hypoxylon* B4728 is located between genes encoding a phosphoketolase and a mitochondrial protein, which are linked in the *Epichloë* species. Pseudogenes, white arrows; flanking genes, grey arrows; and //, regions of the cluster containing repetitive elements.

drugmacidins and proline-containing diketopiperazines (cyclic dipeptides) that contribute to flavours in beer (Gautschi et al., 1997). Interestingly, the ring system of peramine also resembles the ring system formed from amino acids II and III of ergopeptine alkaloids (Schardl et al., 2006). However, unlike these other natural products, one oxo group has been reduced to enamine, probably by the action of a C-terminal reductase domain ( $R^*$ ) of the multifunctional peramine synthase (Tanaka, Tapper, Popay, Parker, & Scott, 2005). The functions of peramine synthase have been predicted by bioinformatic analysis of the *perA* gene sequence. Curiously, variants of *perA* in several *Epichloë* species lack the  $R^*$  domain but contain all of the others, and (at least in *E. festucae* E2368) this variant is expressed, suggesting that these *perA*- $\Delta R^*$  variants may actually determine synthesis of a related metabolite (Schardl, Young, Hesse, et al., 2013). If so, some condensation, esterase, reductase, or cyclase function needs to be provided in *trans* in order

to release the chain from the enzyme, and until such a subunit is identified, the nature of the metabolite might remain unknown.

### 3.2. Variation in gene order

In the *LOL* gene clusters (Fig. 10.6), the gene orders are well conserved between species and even genera. For example, the order in *E. festucae* E2368 is *lolF*, *C*, *D*, *O*, *A*, *U*, *P*, *T*, *E*, *N*, *M*, and although the *LOL* cluster in *A. hypoxylon* lacks *lolP*, *N*, and *M*, all of the other *LOL* genes are in the same order. What is curious is that, while the gene order in the *LOL* cluster has been conserved, the placement and orientation of the *LOL* cluster in the genome have not. Among the *Epichloë* species, the cluster has flipped orientation at least once. In E2368, the *lolF* gene is nearer to the housekeeping gene *nsfA*, and *lolE*, *N*, and *M* are on the side nearer to *lteA*, whereas in *E. glyceriae* E277, *lolE* and a remnant of *lolN* are nearer to *nsfA*, and *lolF* is nearer to *lteA*. Comparing between genera, it is apparent that the entire cluster has moved while remaining intact. Genes for a predicted mitochondrial protein and a predicted phosphoketolase flank the *LOL* cluster (plus a predicted MSF transporter gene) in *A. hypoxylon* but are adjacent to each other in *Epichloë* species. The converse is true concerning the adjacent *nsfA* and *lteA* genes in *Atkinsonella* species, which are orthologous to the housekeeping genes that flank *LOL* in *Epichloë* species. It is not apparent which are the ancestral and which are derived positions of the *LOL* cluster, and it is even possible that the positions in *Epichloë* and *At. hypoxylon* are both derived from an ancestral condition where the locus was in yet another position in the genome.

Unlike the *LOL* gene clusters, more variation is found within the *IDT* gene order and cluster locations (Fig. 10.5). Based on observations of various *IDT* clusters across diverse species, it appears the most likely ancestral order of genes is *idtQ*, *P*, *F*, *G*, *B*, *C*, *S*, *M*, *K*. However, there is a variation among the *Epichloë* species, whereby the gene order has undergone a rearrangement that is now represented by *idtP*, *Q*, *F*, *C*, *B*, *G*, *S*, *M*, *K*. The two additional genes, *ltmE* and *ltmJ*, that are found in *E. festucae* F11 and other lolitrem B-producing isolates appear to be derived from gene duplication and neofunctionalization, whereby *ltmJ* has arisen from *ltmK*, and *ltmE* is the likely fusion of *ltmC* and *ltmF* (Schardl, Young, Hesse, et al., 2013). Gene duplication or acquisition is not limited to the *Epichloë* species but also can be found in other genera. The *IDT* clusters of *M. acridum* and *M. anisopliae* harbour what appears to be a fairly recent duplication of *idtF*. The *E. aotearoae*

e899 *IDT* cluster harbours pseudogenes or remnants of *idtF* and *idtB* and lacks *idtQ* and *idtK* but has highly divergent paralogs evident in AT-rich regions of the genomes (Schardl, Young, Pan, et al., 2013).

Similar to the *LOL* clusters, it appears that the *IDT* clusters have also relocated (Fig. 10.5). It is clear that both *C. purpurea* and *A. take* share common flanking genes of *romB* and *ppdG*, but other species show more variation and tend to share only one flanking gene. Most commonly found in the *Epichloë* species is an association of the *IDT* clusters with a telomere, yet some of these clusters have flipped orientation. In F11, the telomere is most distal from *ltmK*, yet the *E. coenophiala* e4163 *IDT* cluster, which was derived from an *E. festucae* ancestor, has the telomere adjacent to *idtK*. Generally, a polyketide synthase gene or pseudogene and an associated hypothetical gene on one end and a telomere on the other end flank the *IDT* clusters of the *Epichloë* spp. However, *E. gansuensis* e7080 has a serine protein kinase adjacent to *idtP* that is also seen in the *At. texensis* *IDT* cluster, although the segment containing P450 monooxygenase genes *idtP* and *idtQ* has flipped in e7080. An *idoI* gene, so designated because it is similar to genes for indoleamine 2,3-dioxygenase, is adjacent to the *IDT* locus in *C. paspali* (flanks *idtM*), *Ch. alba* (flanks *idtK*), and *P. ipomoeae* (flanks *idtK*). The possibility has not been ruled out that the *idoI* gene product acts on indole-diterpenes either for further modification or as part of a catabolic pathway.

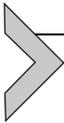
The ergot alkaloid (*EAS*) clusters (Fig. 10.4), like the *IDT* clusters, are also telomeric in *Epichloë* species and internal in others. Like the other clusters, gene content varies among species, but unlike *LOL* and *IDT* clusters, the order of *EAS* genes also varies considerably. In *Claviceps*, *Atkinsonella*, and *Balansia* species, the clusters are internal, and whether this is also the case for *P. ipomoeae* is unknown. In the sequenced members of these genera, the gene order is conserved, though only *P. ipomoeae* has all of the known *EAS* genes associated with this family of fungi.

### 3.3. Variations in gene expression

The *EAS* locus seems exceptional in that, even though some strains also lack it, others seem to have a full complement of *EAS* genes but do not produce detectable levels of ergot alkaloids (Schardl, Young, Hesse, et al., 2013). Some such *E. festucae* strains were subjected to RNA-seq analysis, results of which indicated that their *EAS* genes were not expressed. Why those genes are silent in some strains, whereas other strains such as *E. festucae* F11 express the genes and the alkaloids, remains unknown. In both cases,

the genes are interspersed with large blocks of AT-rich repeats and are located near telomeres, but there are some differences in the repeats and, surprisingly, the active locus in F11 is much closer to the telomere than is the inactive locus in strain E2368. It may be informative to conduct a detailed analysis of the effects of certain sequences in these regions as enhancers or silencers, as has been done for penicillin genes in *Aspergillus nidulans* (Shaaban et al., 2010).

Similar RNA-seq analysis indicated that genes of the partial *IDT* cluster in E2368 (which lacks key genes for indole-diterpene synthesis) are expressed very poorly compared to those in the complete *IDT* cluster of the lolitrem-producing strain, F11. In contrast, both the complete *perA* gene of F11 and the homologous *perA-ΔR\** gene of E2368 are expressed well, even though the latter lacks the terminal domain (R\*) required for peramine biosynthesis. Conceivably, *PerA-ΔR\** is involved in the synthesis of a related metabolite yet to be identified.



#### 4. FUTURE PERSPECTIVES

To date, our focus on sequencing multiple plant-associated Clavicipitaceae has been to understand the dynamics of defensive alkaloid diversity, such as how variation within an alkaloid gene cluster is responsible for the accumulation of diverse major pathway end products. This research has provided insight into genes responsible for specific pathway steps (Pan et al., 2014) and has been especially important to decipher alkaloid diversity of mutualistic endophytes found in cool-season grasses (Charlton et al., 2012; Schardl et al., 2012; Schardl, Young, Hesse, et al., 2013; Schardl, Young, Pan, et al., 2013; Takach et al., 2012; Takach & Young, 2014). Comparative genomics studies among the haploid *Epichloë* species have the potential to unravel the genome dynamics and development of polyploid *Epichloë* species that are so prevalent in nature (Schardl, 2010). Understanding alkaloid diversity, along with the identification of additional SM, can provide the molecular basis behind some of the ecological fitness benefits symbionts provide their host.

In-depth studies on the genome dynamics of plant-associated Clavicipitaceae can also provide insight into the life history traits of these fungi. Signalling between the host and symbiont is required to establish and maintain systemic infections, whereby the balance of this interaction is tightly regulated and synchronized with the growth and development of the host, regardless of transmission mode. The development of genomic

resources has already been employed to facilitate mapping signalling pathways and identifying physiological changes that occur in the interaction between symbiont and host (Schurmann, Buttermann, Herrmann, Giesbert, & Tudzynski, 2013; Tanaka, Takemoto, Chujo, & Scott, 2012). The availability of sequenced genomes for plant-associated Clavicipitaceae representing seven genera that span the symbiotic continuum will facilitate studies of the evolution and mechanisms underlying the diversity of metabolism, host interactions, and niche adaptation of plant-associated Clavicipitaceae.

## ACKNOWLEDGMENTS

This work was supported by USDA-NIFA grant 2012-67013-19384, USDA-CSREES grant 2010-34457-21269, National Institutes of Health grants R01GM086888 and 2 P20 RR-16481, and the Samuel Roberts Noble Foundation. Genome sequence analysis was conducted in the University of Kentucky Advanced Genetic Technologies Center. This is publication number 14-12-025 of the Kentucky Agricultural Experiment Station, published with approval of the director.

## REFERENCES

- Alderman, S. C. (2013). Survival, germination, and growth of *Epichloë typhina* and significance of leaf wounds and insects in infection of orchardgrass. *Plant Disease*, *97*, 323–328.
- Al-Reedy, R. M., Malireddy, R., Dillman, C. B., & Kennell, J. C. (2012). Comparative analysis of *Fusarium* mitochondrial genomes reveals a highly variable region that encodes an exceptionally large open reading frame. *Fungal Genetics and Biology*, *49*, 2–14.
- An, Z.-q., Siegel, M. R., Hollin, W., Tsai, H.-F., Schmidt, D., & Schardl, C. L. (1993). Relationships among non-*Acremonium* sp. fungal endophytes in five grass species. *Applied and Environmental Microbiology*, *59*, 1540–1548.
- Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E., et al. (2013). antiSMASH 2.0—A versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Research*, *41*, W204–W212.
- Bok, J. W., Hoffmeister, D., Maggio-Hall, L. A., Murillo, R., Glasner, J. D., & Keller, N. P. (2006). Genomic mining for *Aspergillus* natural products. *Chemistry & Biology*, *13*, 31–37.
- Bushley, K. E., Raja, R., Jaiswal, P., Cumbie, J. S., Nonogaki, M., Boyd, A. E., et al. (2013). The genome of *Tolyposcladium inflatum*: Evolution, organization, and expression of the cyclosporin biosynthetic gene cluster. *PLoS Genetics*, *9*, e1003496.
- Charlton, N. D., Craven, K. D., Mittal, S., Hopkins, A. A., & Young, C. A. (2012). *Epichloë canadensis*, a new interspecific epichloid hybrid symbiotic with Canada wildrye (*Elymus canadensis*). *Mycologia*, *104*, 1187–1199.
- Cheng, J. Z., Coyle, C. M., Panaccione, D. G., & O'Connor, S. E. (2010). A role for old yellow enzyme in ergot alkaloid biosynthesis. *Journal of the American Chemical Society*, *132*, 1776–1777.
- Clay, K. (1994). Hereditary symbiosis in the grass genus *Danthonia*. *New Phytologist*, *126*, 223–231.
- Clay, K., & Schardl, C. (2002). Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. *American Naturalist*, *160*, S99–S127.

- Cook, D., Beaulieu, W. T., Mott, I. W., Riet-Correa, F., Gardner, D. R., Grum, D., et al. (2013). Production of the alkaloid swainsonine by a fungal endosymbiont of the ascomycete order Chaetothyriales in the host *Ipomoea carnea*. *Journal of Agricultural and Food Chemistry*, *61*, 3797–3803.
- Coyle, C. M., Cheng, J. Z., O'Connor, S. E., & Panaccione, D. G. (2010). An old yellow enzyme gene controls the branch point between *Aspergillus fumigatus* and *Claviceps purpurea* ergot alkaloid pathways. *Applied and Environmental Microbiology*, *76*, 3898–3903.
- Coyle, C. M., & Panaccione, D. G. (2005). An ergot alkaloid biosynthesis gene and clustered hypothetical genes from *Aspergillus fumigatus*. *Applied and Environmental Microbiology*, *71*, 3112–3118.
- Cuomo, C. A., Güldener, U., Xu, J.-R., Trail, F., Turgeon, B. G., Di Pietro, A., et al. (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science*, *317*, 1400–1402.
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, *36*, W465–W469.
- Diehl, W. W. (1950). *Balansia and the Balansiae in America*. Washington, D.C.: United States Department of Agriculture.
- Edgar, R. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*, 113.
- Fleetwood, D. J., Khan, A. K., Johnson, R. D., Young, C. A., Mittal, S., Wrenn, R. E., et al. (2011). Abundant degenerate miniature inverted-repeat transposable elements in genomes of epichloid fungal endophytes of grasses. *Genome Biology and Evolution*, *3*, 1253–1264.
- Fleetwood, D. J., Scott, B., Lane, G. A., Tanaka, A., & Johnson, R. D. (2007). A complex ergovaline gene cluster in epichloë endophytes of grasses. *Applied and Environmental Microbiology*, *73*, 2571–2579.
- Formighieri, E. F., Tiburcio, R. A., Armas, E. D., Medrano, F. J., Shimo, H., Carels, N., et al. (2008). The mitochondrial genome of the phytopathogenic basidiomycete *Moniliophthora perniciosa* is 109 kb in size and contains a stable integrated plasmid. *Mycological Research*, *112*, 1136–1152.
- Freeman, E. M. (1904). The seed fungus of *Lolium temulentum* L., the darnel. *Philosophical Transactions of the Royal Society of London, Series B*, *196*, 1–27.
- Gao, Q., Jin, K., Ying, S.-H., Zhang, Y., Xiao, G., Shang, Y., et al. (2011). Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. *PLoS Genetics*, *7*, e1001264.
- Gautschi, M., Schmid, J. P., Peppard, T. L., Ryan, T. P., Tuorto, R. M., & Yang, X. (1997). Chemical characterization of diketopiperazines in beer. *Journal of Agricultural and Food Chemistry*, *45*, 3183–3189.
- Ghikas, D., Kouvelis, V., & Typas, M. (2006). The complete mitochondrial genome of the entomopathogenic fungus *Metarhizium anisopliae* var. *anisopliae*: Gene order and *trn* gene clusters reveal a common evolutionary course for all Sordariomycetes, while intergenic regions show variation. *Archives of Microbiology*, *185*, 393–401.
- Glenn, A. E., Rykard, D. M., Bacon, C. W., & Hanlin, R. T. (1998). Molecular characterization of *Myriogenospora atramentosa* and its occurrence on some new hosts. *Mycological Research*, *102*, 483–490.
- Gröger, D., & Floss, H. G. (1998). Biochemistry of ergot alkaloids—Achievements and challenges. *Alkaloids: Chemistry and Biology*, *50*, 171–218.
- Güldener, U., Mannhaupt, G., Munsterkötter, M., Haase, D., Oesterheld, M., Stumpflen, V., et al. (2006). FGDB: A comprehensive fungal genome resource on the plant pathogen *Fusarium graminearum*. *Nucleic Acids Research*, *34*, D456–D458.

- Haarmann, T., Lorenz, N., & Tudzynski, P. (2008). Use of a nonhomologous end joining deficient strain ( $\Delta ku70$ ) of the ergot fungus *Claviceps purpurea* for identification of a non-ribosomal peptide synthetase gene involved in ergotamine biosynthesis. *Fungal Genetics and Biology*, *45*, 35–44.
- Hofmann, A. (1978). Historical view on ergot alkaloids. *Pharmacology*, *16*, 1–11.
- Havemann, J., Vogel, D., Loll, B., & Keller, U. (2014). Cyclolization of D-lysergic acid alkaloid peptides. *Chemistry & Biology*, in press. <http://dx.doi.org/10.1016/j.chembiol.2013.11.008>.
- Kang, Y., Ji, Y. L., Zhu, K. R., Wang, H., Miao, H. M., & Wang, Z. W. (2011). A new *Epichloë* species with interspecific hybrid origins from *Poa pratensis* ssp *pratensis* in Liyang, China. *Mycologia*, *103*, 1341–1350.
- Kepler, R. M., Sung, G.-H., Harada, Y., Tanaka, K., Tanaka, E., Hosoya, T., et al. (2012). Host jumping onto close relatives and across kingdoms by *Tyrannicordyceps* (Clavicipitaceae) *gen. nov.* and *Ustilaginoides* (Clavicipitaceae). *American Journal of Botany*, *99*, 552–561.
- Khaldi, N., Seifuddin, F. T., Turner, G., Haft, D., Nierman, W. C., Wolfe, K. H., et al. (2010). SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genetics and Biology*, *47*, 736–741.
- Kuldau, G. A., Tsai, H.-F., & Schardl, C. L. (1999). Genome sizes of *Epichloë* species and anamorphic hybrids. *Mycologia*, *91*, 776–782.
- Leuchtmann, A., Bacon, C. W., Schardl, C. L., White, J. F., & Tadych, M. (2014). Nomenclatural realignment of *Neotyphodium* species with genus *Epichloë*. *Mycologia*, (in press).
- Leuchtmann, A., & Clay, K. (1989). Morphological, cultural and mating studies on *Atkinsonella*, including *A. texensis*. *Mycologia*, *81*, 692–701.
- Lowe, T. M., & Eddy, S. R. (1997). tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, *25*, 955–964.
- Martinez, D., Berka, R. M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S. E., et al. (2008). Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nature Biotechnology*, *26*, 553–560.
- Moon, C. D., Craven, K. D., Leuchtmann, A., Clement, S. L., & Schardl, C. L. (2004). Prevalence of interspecific hybrids amongst asexual fungal endophytes of grasses. *Molecular Ecology*, *13*, 1455–1467.
- Motoyama, T., Hayashi, T., Hirota, H., Ueki, M., & Osada, H. (2012). Terpendole E, a kinesin Eg5 inhibitor, is a key biosynthetic intermediate of indole-diterpenes in the producing fungus *Chaunopycnis alba*. *Chemistry & Biology*, *19*, 1611–1619.
- Munday-Finch, S. C., Wilkins, A. L., & Miles, C. O. (1998). Isolation of lolicine A, lolicine B, lolitriol, and lolitrem N from *Lolium perenne* infected with *Neotyphodium lolii* and evidence for the natural occurrence of 31-epilolitremin N and 31-epilolitremin F. *Journal of Agricultural and Food Chemistry*, *46*, 590–598.
- Ortel, I., & Keller, U. (2009). Combinatorial assembly of simple and complex d-lysergic acid alkaloid peptide classes in the ergot fungus *Claviceps purpurea*. *Journal of Biological Chemistry*, *284*, 6650–6660.
- Pan, J., Bhardwaj, M., Faulkner, J. R., Nagabhyru, P., Charlton, N. D., Higashi, R. M., et al. (2014). Ether bridge formation in loline alkaloid biosynthesis. *Phytochemistry*, *98*, 60–68.
- Pantou, M. P., Kouvelis, V. N., & Typas, M. A. (2008). The complete mitochondrial genome of *Fusarium oxysporum*: Insights into fungal mitochondrial evolution. *Gene*, *419*, 7–15.
- Pažoutová, S., Kolarik, M., & Kolínská, R. (2004). Pleomorphic conidiation in *Claviceps*. *Mycological Research*, *108*, 126–135.
- Pažoutová, S., Olšovská, J., Linka, M., Kolínská, R., & Flieger, M. (2000). Chemoraces and habitat specialization of *Claviceps purpurea* populations. *Applied and Environmental Microbiology*, *66*, 5419–5425.

- Philipson, M. N., & Christey, M. C. (1986). The relationship of host and endophyte during flowering, seed formation, and germination of *Lolium perenne*. *New Zealand Journal of Botany*, *24*, 125–134.
- Rehmeyer, C., Li, W., Kusaba, M., Kim, Y.-S., Brown, D., Staben, C., et al. (2006). Organization of chromosome ends in the rice blast fungus, *Magnaporthe oryzae*. *Nucleic Acids Research*, *34*, 4685–4701.
- Riederer, B., Han, M., & Keller, U. (1996). D-lysergyl peptide synthetase from the ergot fungus *Claviceps purpurea*. *Journal of Biological Chemistry*, *271*, 27524–27530.
- Rudgers, J. A., Holah, J., Orr, S. P., & Clay, K. (2007). Forest succession suppressed by an introduced plant-fungal symbiosis. *Ecology*, *88*, 18–25.
- Ryley, M. J. (2003). *Nigrocorpus scleroticus*, a common old world balansoid fungus. In J. F. White et al. (Ed.), *Clavicipitacean fungi: Evolutionary biology, chemistry, biocontrol and cultural impacts* (pp. 247–271). New York, NY: Marcel-Dekker, Inc.
- Saikia, S., Parker, E. J., Koulman, A., & Scott, B. (2006). Four gene products are required for the fungal synthesis of the indole-diterpene, paspaline. *FEBS Letters*, *580*, 1625–1630.
- Saikia, S., Takemoto, D., Tapper, B. A., Lane, G. A., Frazer, K., & Scott, B. (2012). Functional analysis of an indole-diterpene gene cluster for lolitrem B biosynthesis in the grass endosymbiont *Epichloë festucae*. *FEBS Letters*, *586*, 2563–2569.
- Sampson, K. (1937). Further observations on the systemic infection of *Lolium*. *Transactions of the British Mycological Society*, *21*, 84–97.
- Schardl, C. L. (2010). The epichloae, symbionts of the grass subfamily Poöideae. *Annals of the Missouri Botanical Garden*, *97*, 646–665.
- Schardl, C. L., Craven, K. D., Speakman, S., Stromberg, A., Lindstrom, A., & Yoshida, R. (2008). A novel test for host-symbiont codivergence indicates ancient origin of fungal endophytes in grasses. *Systematic Biology*, *57*, 483–498.
- Schardl, C. L., Leuchtman, A., & Spiering, M. J. (2004). Symbioses of grasses with seedborne fungal endophytes. *Annual Review of Plant Biology*, *55*, 315–340.
- Schardl, C. L., Panaccione, D. G., & Tudzynski, P. (2006). Ergot alkaloids—Biology and molecular biology. *Alkaloids: Chemistry and Biology*, *63*, 45–86.
- Schardl, C. L., Scott, B., Florea, S., & Zhang, D.-X. (2009). Epichloë endophytes: Clavicipitaceous symbionts of grasses. In H. B. Deising (Ed.), *The mycota V: Plant relationships* (pp. 275–306). Berlin: Springer-Verlag.
- Schardl, C. L., Young, C. A., Faulkner, J. R., Florea, S., & Pan, J. (2012). Chemotypic diversity of epichloae, fungal symbionts of grasses. *Fungal Ecology*, *5*, 331–344.
- Schardl, C. L., Young, C. A., Hesse, U., Amyotte, S. G., Andreeva, K., Calie, P. J., et al. (2013). Plant-symbiotic fungi as chemical engineers: Multi-genome analysis of the Clavicipitaceae reveals dynamics of alkaloid loci. *PLoS Genetics*, *9*, e1003323.
- Schardl, C. L., Young, C. A., Pan, J., Florea, S., Takach, J. E., Panaccione, D. G., et al. (2013). Currencies of mutualisms: Sources of alkaloid genes in vertically transmitted epichloae. *Toxins*, *5*, 1064–1088.
- Schurmann, J., Buttermann, D., Herrmann, A., Giesbert, S., & Tudzynski, P. (2013). Molecular characterization of the NADPH oxidase complex in the ergot fungus *Claviceps purpurea*: CpNox2 and CpPls1 are important for a balanced host-pathogen interaction. *Molecular Plant-Microbe Interactions*, *26*, 1151–1164.
- Selosse, M. A., & Schardl, C. L. (2007). Fungal endophytes of grasses: Hybrids rescued by vertical transmission? An evolutionary perspective. *New Phytologist*, *173*, 452–458.
- Shaaban, M., Palmer, J. M., El-Naggar, W. A., El-Sokkary, M. A., Habib, E. L. S. E., & Keller, N. P. (2010). Involvement of transposon-like elements in penicillin gene cluster regulation. *Fungal Genetics and Biology*, *47*, 423–432.
- Steiner, U., Leibner, S., Schardl, C. L., Leuchtman, A., & Leistner, E. (2011). *Periglandula*, a new fungal genus within the Clavicipitaceae and its association with Convolvulaceae. *Mycologia*, *103*, 1133–1145.

- Sung, G. H., Sung, J. M., Hywel Jones, N. L., & Spatafora, J. W. (2007). A multi-gene phylogeny of Clavicipitaceae (Ascomycota, Fungi): Identification of localized incongruence using a combinational bootstrap approach. *Molecular Phylogenetics and Evolution*, *44*, 1204–1223.
- Tadych, M., Ambrose, K., Bergen, M., Belanger, F., & White, J. (2012). Taxonomic placement of *Epichloë poae* sp. nov. and horizontal dissemination to seedlings via conidia. *Fungal Diversity*, *54*, 117–131.
- Tagami, K., Liu, C., Minami, A., Noike, M., Isaka, T., Fueki, S., et al. (2013). Reconstitution of biosynthetic machinery for indole-diterpene paxilline in *Aspergillus oryzae*. *Journal of the American Chemical Society*, *135*, 1260–1263.
- Takach, J. E., Mittal, S., Swoboda, G. A., Bright, S. K., Trammell, M. A., Hopkins, A. A., et al. (2012). Genotypic and chemotypic diversity of *Neotyphodium* endophytes in tall fescue from Greece. *Applied and Environmental Microbiology*, *78*, 5501–5510.
- Takach, J. E., & Young, C. A. (2014). Alkaloid genotype diversity of tall fescue endophytes. *Crop Science* (in press).
- Tanaka, A., Takemoto, D., Chujo, T., & Scott, B. (2012). Fungal endophytes of grasses. *Current Opinion in Plant Biology*, *15*, 462–468.
- Tanaka, E., & Tanaka, C. (2008). Phylogenetic study of clavicipitaceous fungi using acetaldehyde dehydrogenase gene sequences. *Mycoscience*, *49*, 115–125.
- Tanaka, A., Tapper, B. A., Popay, A., Parker, E. J., & Scott, B. (2005). A symbiosis expressed non-ribosomal peptide synthetase from a mutualistic fungal endophyte of perennial ryegrass confers protection to the symbiotum from insect herbivory. *Molecular Microbiology*, *57*, 1036–1050.
- Tasso, B., Novelli, F., Sparatore, F., Fasoli, F., & Gotti, C. (2013). (+)-Laburnamine, a natural selective ligand and partial agonist for the  $\alpha 4\beta 2$  nicotinic receptor subtype. *Journal of Natural Products*, *76*, 727–731.
- Tofern, B., Kaloga, M., Witte, L., Hartmann, T., & Eich, E. (1999). Phytochemistry and chemotaxonomy of the Convolvulaceae part 8—Occurrence of loline alkaloids in *Argyrea mollis* (Convolvulaceae). *Phytochemistry*, *51*, 1177–1180.
- Tudzynski, P., & Scheffer, J. (2004). *Claviceps purpurea*: Molecular aspects of a unique pathogenic lifestyle. *Molecular Plant Pathology*, *5*, 377–388.
- Turgeon, G., & Yoder, O. C. (2000). Proposed nomenclature for mating type genes of filamentous ascomycetes. *Fungal Genetics and Biology*, *31*, 1–5.
- Uhlig, S., Botha, C. J., Vrålstad, T., Rolén, E., & Miles, C. O. (2009). Indole-diterpenes and ergot alkaloids in *Cynodon dactylon* (Bermuda grass) infected with *Claviceps cynodontis* from an outbreak of tremors in cattle. *Journal of Agricultural and Food Chemistry*, *57*, 11112–11119.
- Vaillancourt, L., Du, M. Z., Wang, J., Rollins, J., & Hanau, R. (2000). Genetic analysis of cross fertility between two self-sterile strains of *Glomerella graminicola*. *Mycologia*, *92*, 430–435.
- Veen, G., Greinwald, R., Canto, P., Witte, L., & Czygan, F. C. (1992). Alkaloids of *Adenocarpus hispanicus* (Lam.) DC varieties. *Zeitschrift für Naturforschung C-A. Journal of Biosciences*, *47*, 341–345.
- Walker, J. (2004). *Claviceps phalaridis* in Australia: Biology, pathology and taxonomy with a description of the new genus *Cepsiclava* (Hypocreales, Clavicipitaceae). *Australasian Plant Pathology*, *33*, 211–239.
- White, J. F., Jr., & Reddy, P. V. (1998). Examination of structure and molecular phylogenetic relationships of some graminicolous symbionts in genera *Epichloë* and *Parepichloë*. *Mycologia*, *90*, 226–234.
- Wiemann, P., Sieber, C. M. K., von Bargen, K. W., Studt, L., Niehaus, E.-M., Espino, J. J., et al. (2013). Deciphering the cryptic genome: Genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathogens*, *9*, e1003475.
- Young, C., Bryant, M., Christensen, M., Tapper, B., Bryan, G., & Scott, B. (2005). Molecular cloning and genetic analysis of a symbiosis-expressed gene cluster for lolitrem

- biosynthesis from a mutualistic endophyte of perennial ryegrass. *Molecular Genetics and Genomics*, 274, 13–29.
- Young, C. A., Felitti, S., Shields, K., Spangenberg, G., Johnson, R. D., Bryan, G. T., et al. (2006). A complex gene cluster for indole-diterpene biosynthesis in the grass endophyte *Neotyphodium lolii*. *Fungal Genetics and Biology*, 43, 679–693.
- Young, C., McMillan, L., Telfer, E., & Scott, B. (2001). Molecular cloning and genetic analysis of an indole-diterpene gene cluster from *Penicillium paxilli*. *Molecular Microbiology*, 39, 754–764.
- Young, C. A., Tapper, B. A., May, K., Moon, C. D., Schardl, C. L., & Scott, B. (2009). Indole-diterpene biosynthetic capability of epichloë endophytes as predicted by *ltm* gene analysis. *Applied and Environmental Microbiology*, 75, 2200–2211.
- Zheng, P., Xia, Y. L., Xiao, G. H., Xiong, C. H., Hu, X., Zhang, S. W., et al. (2011). Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional chinese medicine. *Genome Biology*, 12, R116.