

Review

Genetics, Genomics and Evolution of Ergot Alkaloid Diversity

Carolyn A. Young ^{1,†,*}, Christopher L. Schardl ^{2,†}, Daniel G. Panaccione ³, Simona Florea ², Johanna E. Takach ¹, Nikki D. Charlton ¹, Neil Moore ⁴, Jennifer S. Webb ⁵ and Jolanta Jaromczyk ⁵

¹ Forage Improvement Division, The Samuel Roberts Noble Foundation, Ardmore, OK 73401, USA; E-Mails: jtakach@greenandgrow.com (J.E.T.); ndcharlton@noble.org (N.D.C.)

² Department of Plant Pathology, University of Kentucky, Lexington, KY 40546, USA; E-Mails: schardl@uky.edu (C.L.S.); sflor2@uky.edu (S.F.)

³ Division of Plant and Soil Sciences, West Virginia University, Morgantown, WV 26506, USA; E-Mail: danpan@wvu.edu

⁴ Computer Science Department, University of Kentucky, Lexington, KY 40546, USA; E-Mail: neil@uky.edu

⁵ Advanced Genetic Technologies Center, University of Kentucky, Lexington, KY 40546, USA; E-Mails: jswebb2@uky.edu (J.S.W.); jolanta.jaromczyk@uky.edu (J.J.)

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: cayoung@noble.org; Tel.: +1-580-224-6860; Fax: +1-580-224-6802.

Academic Editor: Richard A. Manderville

Received: 6 March 2015 / Accepted: 8 April 2015 / Published: 14 April 2015

Abstract: The ergot alkaloid biosynthesis system has become an excellent model to study evolutionary diversification of specialized (secondary) metabolites. This is a very diverse class of alkaloids with various neurotropic activities, produced by fungi in several orders of the phylum Ascomycota, including plant pathogens and protective plant symbionts in the family Clavicipitaceae. Results of comparative genomics and phylogenomic analyses reveal multiple examples of three evolutionary processes that have generated ergot-alkaloid diversity: gene gains, gene losses, and gene sequence changes that have led to altered substrates or product specificities of the enzymes that they encode (neofunctionalization). The chromosome ends appear to be particularly effective engines for gene gains, losses and rearrangements, but not necessarily for neofunctionalization. Changes in gene expression

could lead to accumulation of various pathway intermediates and affect levels of different ergot alkaloids. Genetic alterations associated with interspecific hybrids of *Epichloë* species suggest that such variation is also selectively favored. The huge structural diversity of ergot alkaloids probably represents adaptations to a wide variety of ecological situations by affecting the biological spectra and mechanisms of defense against herbivores, as evidenced by the diverse pharmacological effects of ergot alkaloids used in medicine.

Keywords: *Claviceps*; *Epichloë*; *Periglandula*; Clavicipitaceae; gene clusters; chanoclavine; ergopeptine; subterminal; natural products; secondary metabolism

1. Importance of Ergot Alkaloids

Ergot alkaloids are well known mycotoxins that can contaminate food and feed but also can serve as starting materials for important pharmaceuticals. The ergot fungi, for which these alkaloids are named, have been responsible for historic episodes of mass poisoning. In Middle-Age Europe, ingestion of rye grain or flour that was contaminated with ergots—the resting stage (sclerotia) of *Claviceps purpurea*—led to multiple episodes of disfiguring and deadly poisoning of local populations. Historic events associated with ergot poisoning include the first Crusade [1], the Salem witch trials (and others) [2–4], and the interrupted 1722 Russian campaign (under Peter the Great) against the Ottoman empire [5]. In modern times, ergot-alkaloid poisoning occasionally occurs through their medicinal use [6], but mass ergot poisonings of humans are rare, having last been reported in the 1970s [7,8]. Ergot alkaloid toxicity is still a significant problem with livestock, both due to ergot contaminated feed and naturally infested forage or rangeland grasses—such as tall fescue (*Lolium arundinaceum*), sleepygrass (*Achnatherum robustum*) and drunken horse grass (*Achnatherum inebrians*)—which can be symbiotic with seed-transmissible *Epichloë* species that are capable of producing ergot alkaloids [9–11].

Over the past two decades, ergot alkaloids have been tapped for increasingly diverse medical uses. The ergopeptine ergotamine is used to treat migraines [12], and other natural and semisynthetic ergopeptines and dihydroergopeptines have been used for diseases of the brain. For example, bromocriptine (2-bromo-ergocryptine) is used as a component in treatment of Parkinsonism [13] and bromocriptine, or the extensively substituted dihydrolysergic acid amide, cabergoline, is used in treatment of prolactinoma, a benign adenoma of the pituitary gland [14,15]. Ergot alkaloids are also of social relevance because a semisynthetic alkaloid, lysergic acid diethylamide (LSD), is an illicit drug that is by far the most potent hallucinogen known. LSD had a major impact on the countercultural and hippie movements of the 1960s, since Albert Hofmann first produced it and noted its properties [16].

2. Structural Content of Ergot Alkaloid Biosynthesis (EAS) Loci Define Alkaloid Potential

The ergot alkaloids represent a diverse class of natural products that are generally divided into three subclasses: the simpler clavines, lysergic acid and its simple amides, and the highly complex ergopeptines. Figure 1 shows the simpler clavines in blue, green, and purple, and the lysergic acid and simple amides and ergopeptines both in red, reflective of the orders of biosynthetic steps: blue for early,

green for middle, and purple and red for late steps in different fungal families. Clavines are tricyclic or tetracyclic compounds known from the fungal families Clavicipitaceae (order Hypocreales) and Trichocomaceae (Eurotiales), with the latter showing more variation due to hydroxylation, prenylation and acetylation (reviewed in [17]). The presence of *EAS* clusters in the Arthrodermataceae (Onygenales) suggests that they also may produce clavines, a possibility that is supported by the demonstration of chanoclavine I dehydrogenase activity of the *Arthroderma benhamiae* EasD ortholog (GenBank accession EFE37118.1) [18]. Lysergic acid and the lysergic acid amides are tetracyclic compounds with a common ergolene core, whereas ergopeptines are lysergic acid-tripeptide derivatives. These more complex ergot alkaloids are almost exclusively known from Clavicipitaceae, although an ergopeptine has also been reported from a *Dicyma* sp. (Xylariaceae, Xylariales) [19].

All known ergot alkaloid biosynthesis genes are present in the genomes of ascomycetous fungi grouped either in a single ergot alkaloid synthesis (*EAS*) cluster or divided into two *EAS* clusters (Figure 2). The particular forms and overall profile of ergot alkaloids produced by a fungus are determined by the presence or absence of pathway genes and, for several *EAS* genes, the particular enzyme isoforms they encode. The basic functions have been determined for all known *EAS* genes in the Clavicipitaceae, and most in the Trichocomaceae (reviewed in [20]), so that detecting gene presence or absence provides considerable power to predict alkaloid profiles, but there is also important variation in substrate and product specificities of some *EAS* gene-encoded enzymes, which we are not yet able to infer from the gene sequences [17].

2.1. Genes Encoding the Early Pathway Steps

The four conserved early pathway steps for all natural ergot alkaloids produce chanoclavine I (CC) (recently reviewed [20]), and are catalyzed by enzymes encoded in the four genes: *dmaW*, encoding dimethylallyltryptophan synthase, *easF*, encoding dimethylallyltryptophan *N*-methyltransferase, *easC*, encoding a catalase, and *easE*, encoding chanoclavine-I synthase. Some strains, such as *E. elymi* E56, produce CC as the pathway end-product because they contain functional copies only of these four genes in an *EAS* cluster, which we designate as *EAS*^{CC} (Figure 2; Table 1). Based on similar complements in their genomes, we would predict that *E. brachyelytri* E4804 and *Atkinsonella hypoxylon* B4728 are also CC producers, though CC has not been detected in plants infected with B4728. Strains capable of making complex ergot alkaloids also can generate substantial levels of CC and other intermediates and spur products as a result of inherent pathway inefficiency, a property suggested to have selective advantage [21,22].

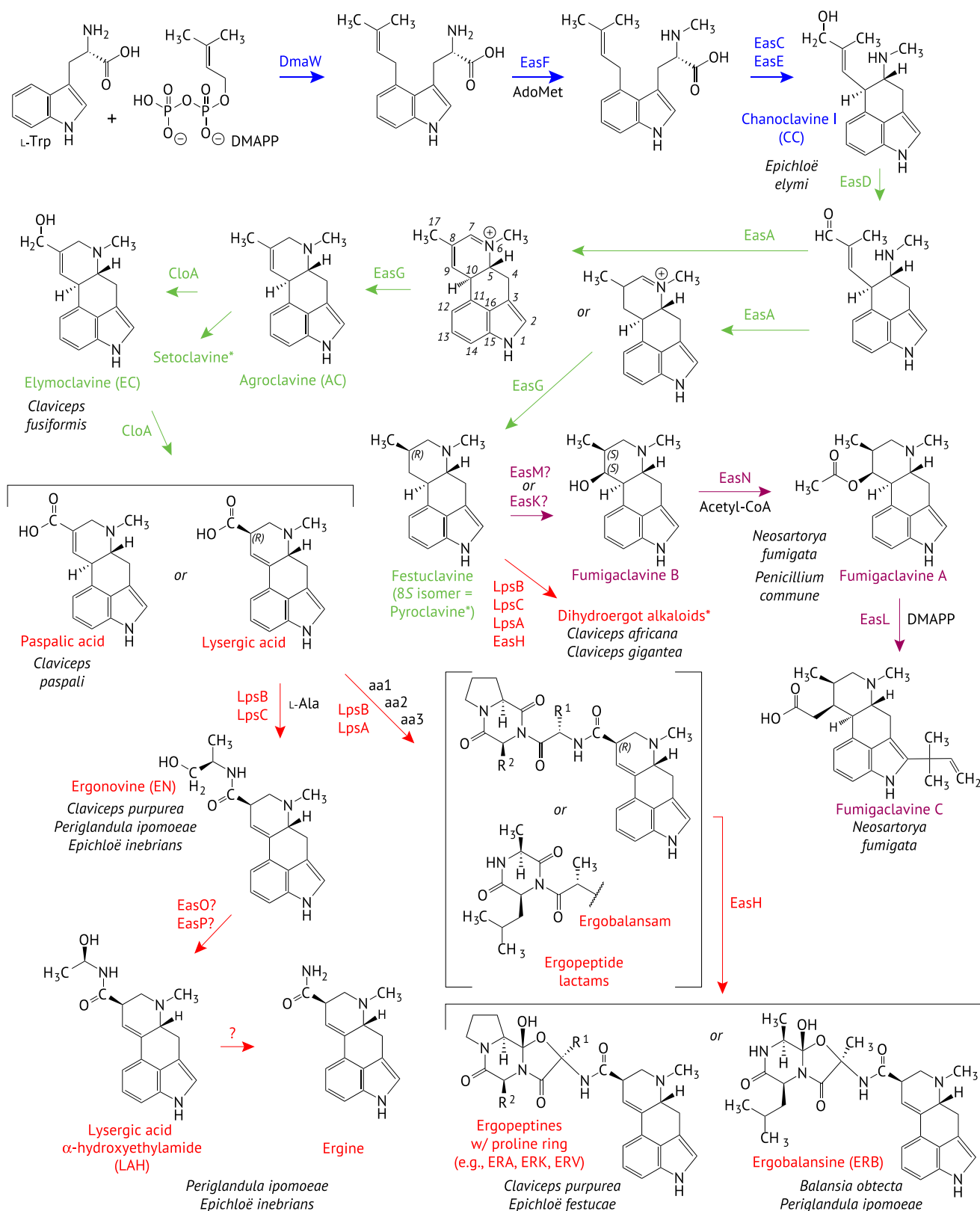


Figure 1. The ergot alkaloid pathway showing steps that result in diversification of compounds. Pathway steps are color-coded based on the position or diversification within the pathway, Blue = early steps to the intermediate chanoclavine, Green = mid steps leading to the tetracyclic clavines, Red = late steps represented by the lysergic acid amides and the complex ergopeptines, Purple = steps to fumigaclavines produced by Trichocomaceae.

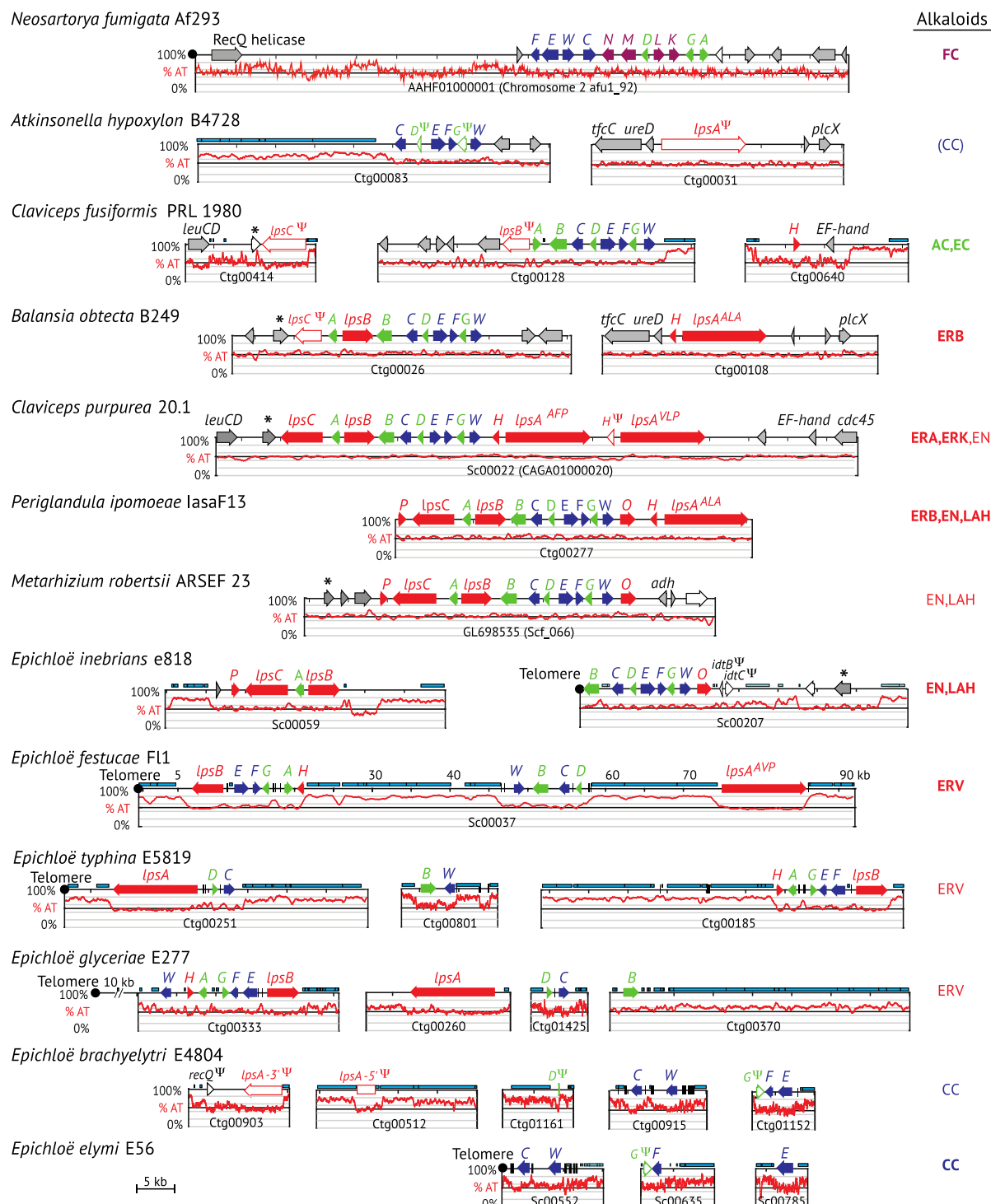


Figure 2. Relative adenine and thymine (%AT) DNA content of ergot alkaloid synthesis (*EAS*) loci. Gene name abbreviations are as follow: all *eas* genes = last letter, *cloA* = *B* and *dmaW* = *W*. Gene names are colored to represent the stage of the pathway for the encoded product (see Figure 1). Pseudogenes are represented by Ψ and white-filled arrows. The major pathway end product of each strain is indicated on the right in bold face (product produced) or regular type (product predicted but not yet tested), or in parentheses (product predicted but undetected). Arrows marked with * represent orthologues of *C. purpurea* AET79176 (GenBank). Cyan bars indicate repeats, and vertical black bars indicate miniature inverted-repeat transposable elements (MITEs). Where present, telomeres are positioned at left.

Table 1. Ergot alkaloid synthesis (*EAS*) gene names and encoded functions.

Gene name ^a	Enzyme	<i>EAS^{CC}</i>	<i>EAS^{EC}</i>	<i>EAS^{ERP}</i>	<i>EAS^{LAH}</i>	<i>EAS^{EN/ERP}</i>	<i>EAS^{LAH/ERP}</i>	<i>EAS^{FC}</i>
<i>dmaW</i>	Dimethylallyltryptophan synthase	+	+	+	+	+	+	+
<i>easF</i>	Dimethylallyltryptophan <i>N</i> -methylase	+	+	+	+	+	+	+
<i>easC</i>	Catalase	+	+	+	+	+	+	+
<i>easE</i>	Chanoclavine-I synthase	+	+	+	+	+	+	+
<i>easD</i>	Chanoclavine-I dehydrogenase		+	+	+	+	+	+
<i>easA</i> ^b	Chanoclavine-1 aldehyde oxidoreductase		+ iso	+ iso	+ iso	+ iso	+ iso	+ red
<i>easG</i>	agroclavine, festuclavine or pyroclavine dehydrogenase		+	+	+	+	+	+
<i>cloA</i> ^b	agroclavine, festuclavine, or elymoclavine monooxygenase		+	+	+	+	+	
<i>lpsB</i>	lysergyl peptide synthetase subunit 2			+	+	+	+	
<i>lpsA</i> ^b	lysergyl peptide synthetase subunit 1			+		+	+	
<i>easH</i>	ergopeptide lactam hydroxylase			+		+	+	
<i>lpsC</i>	lysergyl peptide synthetase and reductase subunit				+	+	+	
<i>easO</i> ^c	putative ergonovine oxygenase				+		+	
<i>easP</i> ^c	putative LAH synthase				+		+	
<i>easM</i> ^c	possible festuclavine 9-monooxygenase							+
<i>easK</i> ^c	possible festuclavine 9-monooxygenase							+
<i>easN</i>	fumigaclavine acetylase							+
<i>easL</i>	fumigaclavine reverse prenyltransferase							+
Fungal species	Examples of ergot alkaloid-producing fungi	<i>Epichloë elymi</i>	<i>Claviceps fusiformis</i>	<i>Epichloë festucae, Balansia oblecta</i>	<i>Epichloë inebrians</i>	<i>Claviceps purpurea</i>	<i>Periglandula ipomoeae</i>	<i>Neosartorya fumigata</i>

^a Pathway steps are color-coded based on the positions within the pathway as shown in Figure 1; ^b Specificity of encoded gene can vary. *EasA* functions as either an isomerase “+ iso” or reductase “+ red”; ^c Actual role not confirmed.

2.2. Diversification of the EAS Pathways

Beyond the production of CC, multiple biosynthetic pathways begin to branch and diverge (Figure 1), and the chemotypic variation between and even within species is reflected in the gene content of each *EAS* locus known or predicted to direct biosynthesis of such pathway end-products as elymoclavine (EC), lysergic acid α -hydroxyethylamide (LAH), ergonovine (EN), and ergopeptines such as ergovaline (ERV), ergotamine (ERA), ergocryptine (ERK) or ergobalansine (ERB). Where clarification is needed, we will designate the various *EAS* clusters with superscripts reflecting end products, as *EAS^{EC}*, *EAS^{LAH}* or *EAS^{EN}*, as well as *EAS^{ERP}* for ergopeptine producers, *EAS^{EN/ERP}* for producers of EN and ergopeptines, and *EAS^{LAH/ERP}* for producers of LAH and ergopeptines. Compared to *EAS^{EN}*, *EAS^{LAH}* has two additional genes, *easO* and *easP*, suggesting that EN may be the LAH precursor. Fungi with *EAS^{LAH}* clusters also tend to accumulate substantial levels of ergine, probably by spontaneous hydrolysis of LAH [23]. The *EAS^{EN/ERP}* cluster in the most infamous ergot fungus, *Claviceps purpurea*, and the *EAS^{LAH/ERP}* cluster in the morning-glory symbiont, *Periglandula ipomoeae*, are the only ones identified to date that determine synthesis of three different ergot alkaloid subclasses [24].

2.2.1. Completion of the Tetracyclic Ergolene Common Core

The EAS pathway diversifies at multiple steps depending, not only on presence or absence of genes, but also on the substrate- and product-specificities of several of the encoded enzymes [17,20] (Figure 1). Once CC is oxidized by the action of the EasD enzyme to form chanoclavine aldehyde, EasA then catalyzes a reduction step that allows rotation around the C8-C9 bond so that an iminium ion (*i.e.*, Schiff base) can form as the first step in the synthesis of the D-ring of the ergolene core common to most ergot alkaloids. Surprisingly, this is one of the steps at which different pathways can diverge to give either ergot alkaloids or dihydroergot alkaloids [17,25]. EasA proteins that follow the reduction step with reoxidation are effectively isomerases, typically found in *Claviceps purpurea* and many species of *Balansia*, *Epichloë* and *Periglandula*. In contrast, EasA isoforms that only reduce the C8-C9 bond direct the pathway toward dihydroergot alkaloids, such as those found in *Claviceps africana* and *Claviceps gigantea* as well as in the fumigaclavine producer, *Neosartorya fumigata*.

The ergolene D-ring is completed by a reduction catalyzed by EasG, to yield agroclavine (for ergot alkaloids) or festuclavine (for dihydroergot alkaloids) [20] (Figure 1). After the D-ring is closed, the C8-linked methyl group can be oxidized by the action of a cytochrome P450 monooxygenase. This enzyme, designated CloA, apparently represents another point in the pathway where variation in an enzyme can affect the alkaloid profile [17]. It appears likely that different isoforms of CloA determine the level of oxidation, such that CloA of *C. fusiformis* catalyzes the 2-electron oxidation of agroclavine (AC) to EC, whereas CloA of *C. purpurea* and many other Clavicipitaceae catalyzes a 6-electron oxidation of agroclavine to paspalic acid or lysergic acid (LA). Whether LA is generated spontaneously or enzymatically from paspalic acid remains unclear.

2.2.2. Formation of Lysergic Acid, Lysergic Acid Amides and Complex Ergopeptines

Despite its fame as a starting material for laboratory synthesis of LSD, LA does not generally occur in appreciable concentrations in natural systems [22]. This is because fungi that make LA are usually

capable of converting it to any of a multitude of lysergic acid amides, ranging from the simplest (ergine = lysergic acid amide) to complex ergopeptines in which LA has an amide linkage with a tricyclic moiety derived from three additional amino acids (Figure 1) (reviewed in [17]). This divergence point involves a remarkable system that centers on the enzyme subunit LpsB (=LPS2) plus one or both of its partner subunits LpsA (=LPS1) or LpsC (=LPS3) depending on whether functional *lpsA* or *lpsC* genes are present [26]. Each of the Lps subunits contains modules that contribute specific catalytic activities and, in combination with other Lps subunits, comprise multi-enzyme complexes called nonribosomal peptide synthetases (NRPSs). Each module AMPylates and thio-esterifies an amino acid, and then condenses it with the similarly processed amino acid on the adjacent module. LpsB specifies LA, whereas LpsC specifies L-alanine (Ala). LpsC also has a C-terminal “R*” domain (cd05235: SDR_e1) that catalyzes reductive release of the LA-Ala conjugate as EN. Alternatively, if LpsB partners with LpsA they form the enzyme complex required for formation of ergopeptide lactams, which are then oxidatively cyclized by the action of EasH [27] to form ergopeptines or their 8(*R*) isomers, the ergopeptinines. The particular composition of each ergopeptide is determined by specificity of each of the three modules in LpsA for its cognate amino acid [28], so far giving 21 different combinations (Table 2) [19,29,30]. For example, *Epichloë* strains that produce ergovaline (ERV) have LpsA^{AVP} (where the superscripts are single-letter codes for the amino acids specified, in order, by the first, second and third modules of LpsA). In contrast, *C. purpurea* strain 20.1 has genes for the LpsA isoforms LpsA^{AFP} and LpsA^{VL}P, which determine production of ergotamine (ERA) and ergocryptine (ERK), respectively. There are 19 ergopeptines known, most with corresponding ergopeptinines, which are their 8(*S*) stereoisomers. Two additional ergopeptinines have no known 8(*R*) isomers (reviewed in [30]). Additionally, *Claviceps africana* produces dihydroergosine, which is similar to ergosine except that it has a saturated D-ring, presumably because it is derived from festuclavine rather than agroclavine (reviewed in [17]).

2.2.3. Fumigaclavine Production by the Trichocomaceae

The Trichocomaceae produce various clavines derived from festuclavine (reviewed in [17]) (Figure 1). They lack agroclavine or its derivatives because they have a reducing rather than isomerizing EasA isoform. They may also produce pyroclavine, the 8(*S*) stereoisomer of festuclavine, due to functional differences in EasG [31]. Further modifications are catalyzed by enzymes encoded in the *EAS* cluster of *N. fumigata*, for which orthologs are not found in other fungi, giving rise to the fumigaclavines. The 9-hydroxylation is probably catalyzed by either EasM or EasK, both predicted to be cytochrome P450. Then *O*-acetylation is catalyzed by EasN, and the “reverse” prenylation step is catalyzed by EasL.

Table 2. Lysergic acid-linked substituents of natural ergopeptines ^a.

Ergopeptine	AA1	R ¹	AA2	R ²	AA3	R ³
Ergotamine (ERA)	Ala	Me	Phe	CH ₂ Ph	Pro	prolyl (CH ₂) ₃
Ergovaline (ERV)	Ala	Me	Val	<i>i</i> -Pr	Pro	prolyl (CH ₂) ₃
Ergosine	Ala	Me	Leu	<i>i</i> -Bu	Pro	prolyl (CH ₂) ₃
Dihydroergosine ^b	Ala	Me	Leu	<i>i</i> -Bu	Pro	prolyl (CH ₂) ₃
β-Ergosine	Ala	Me	Ile	<i>sec</i> -Bu	Pro	prolyl (CH ₂) ₃
Ergosedmine	Ile	<i>sec</i> -Bu	Leu	<i>i</i> -Bu	Pro	Prolyl (CH ₂) ₃
Ergobine	Ala	Me	ABA	Et	Pro	prolyl (CH ₂) ₃
Ergocristine	Val	<i>i</i> -Pr	Phe	CH ₂ Ph	Pro	prolyl (CH ₂) ₃
Ergocornine	Val	<i>i</i> -Pr	Val	<i>i</i> -Pr	Pro	prolyl (CH ₂) ₃
Ergocryptine ^{c,d} (ERK)	Val	<i>i</i> -Pr	Leu	<i>i</i> -Bu	Pro	prolyl (CH ₂) ₃
β-Ergocryptine ^d	Val	<i>i</i> -Pr	Ile	<i>sec</i> -Bu	Pro	prolyl (CH ₂) ₃
γ-Ergocryptinine ^{c,e}	Val	<i>i</i> -Pr	norLeu	<i>n</i> -Bu	Pro	prolyl (CH ₂) ₃
Ergobutyryne	Val	<i>i</i> -Pr	ABA	Et	Pro	prolyl (CH ₂) ₃
Ergoladinine ^e	Val	<i>i</i> -Pr	Met	EtSCH ₃	Pro	prolyl (CH ₂) ₃
Ergogaline	Val	<i>i</i> -Pr	homoIle	2-Me- <i>n</i> -Bu	Pro	prolyl (CH ₂) ₃
Ergostine	ABA	Et	Phe	CH ₂ Ph	Pro	prolyl (CH ₂) ₃
Ergonine	ABA	Et	Val	<i>i</i> -Pr	Pro	prolyl (CH ₂) ₃
Ergoptine ^c	ABA	Et	Leu	<i>i</i> -Bu	Pro	prolyl (CH ₂) ₃
β-ergoptine	ABA	Et	Ile	<i>sec</i> -Bu	Pro	prolyl (CH ₂) ₃
Ergobutine	ABA	Et	ABA	Et	Pro	prolyl (CH ₂) ₃
Ergobalansine (ERB)	Ala	Me	Leu	<i>i</i> -Bu	Ala	Me
Unnamed, from <i>Dicyma</i> sp.	Ala	Me	Leu	<i>i</i> -Bu	Phe	CH ₂ Ph

^a Abbreviations: AA= amino acid position; ABA = 2-aminobutyric acid, norLeu = L-norleucine; homoIle = L-homoisoleucine. Other L-amino acids and R-groups are abbreviated as standard; ^b Dihydroergosine has a saturated D-ring, whereas others listed here have a 9-10 double bond; ^c Synonyms: ergosine = α-ergosine, ergocryptine = α-ergocryptine, ergoptine = α-ergoptine; ^d Synonyms: α-, β-, or γ-ergocryptine = α-, β-, or γ-ergokryptine, respectively; ^e Only the 8(*R*) (=isolysergyl) isomers—namely, ergoladinine and γ-ergocryptinine—have been reported to date.

2.3. Contents of the *EAS* Loci

The *EAS* pathway specificity mentioned above is reflected in the structural content of *EAS* loci across the Clavicipitaceae, which varies based on presence or absence of genes that correspond to ergot alkaloid biosynthetic capability within a given strain. The most *EAS* genes (14) are present in *P. ipomoeae*, which produces ergobalansine (ERB), EN and LAH, whereas only four *EAS* genes are common to all strains that are capable or predicted to produce CC [24]. Interestingly, those four early-pathway genes are grouped together in the *N. fumigata* cluster, whereas the mid-pathway genes for festuclavine biosynthesis are interspersed with late-pathway genes for fumigaclavine [32,33] (Figure 2). The gene arrangements differ between *N. fumigata* and the Clavicipitaceae, and also differ considerably between *Epichloë* spp. and other Clavicipitaceae. In *At. hypoxylon*, *Balansia obtecta*, the *Claviceps* spp., *Epichloë inebrians* (formerly *Epichloë gansuensis* var. *inebrians*; [34]) and *P. ipomoeae*, early and mid-pathway genes (except for *easA*) are interspersed, whereas the late genes (*lps* genes, *easH*, *easO* and *easP*) are at the *EAS*-cluster periphery, separated from the main *EAS* cluster, fatally mutated or

missing. The *easA* gene is the exception among mid-pathway genes, being located between late-pathway genes *lpsB* and *lpsC* in these species. In the other *Epichloë* species, there is similar variation in gene content and functionality, but greater variation in gene arrangement even for early- and late-pathway genes. Clearly, strains that produce only CC are derived by losses of mid- and late-pathway genes, because remnants and pseudogenes often remain in their genomes.

Putative *EAS* genes can be identified in fungal genome sequences using bioinformatics pipelines developed to identify core secondary metabolite genes encoding NRPSs, polyketide synthases (PKS), prenyltransferases (*dmaW* homologues) or terpene cyclases and mixed or hybrid versions [35–37]. The *EAS* gene clusters such as *EAS^{CC}* can sometimes be identified within the prenyltransferase or terpene cyclase group, whereas a more complex *EAS* cluster, such as *EAS^{ERP}*, will fall into a hybrid or mixed category since they contain sequences for both prenyltransferases and NRPSs [38]. In genomes of several *Trichophyton* and *Arthroderma* species (Arthrodermataceae), apparent orthologues of *dmaW*, *easF*, *easC*, *easE* and *easD* are identifiable in a cluster [18]. Genes flanking this cluster match signatures of various biosynthetic functions, so it may well be that Arthrodermataceae produce one or more alkaloid subclasses yet to be identified. The *EAS* gene complements and structural arrangements identified in *Metarhizium robertsii* [39] are very similar to *E. inebrians* and, as such, we would predict that *M. robertsii* strain ARSEF 23 could be capable of producing EN and LAH, assuming there is similar specificity of *LpsB* with LA and *LpsC* with alanine.

3. Phylogenetic Relationships of *EAS* Genes

3.1. Comparison of *EAS* Gene Phylogenies

We have inferred phylogenetic trees for *EAS* genes of known and suspected ergot alkaloid-producing Clavicipitaceae, plus *N. fumigata*, which we presumed to be an outgroup in keeping with housekeeping gene relationships. Phylogenies of the core genes for the first four steps in ergot alkaloid biosynthesis—namely, *dmaW*, *easF*, *easC* and *easE*—were congruent, so a concatenated data set (*WFCE*) was constructed, from which a maximum likelihood (ML) tree was inferred with PhyML at phylogeny.fr [40] (Figure 3). The *lpsB* tree (Figure 4) was also congruent with the corresponding *WFCE* subtree (*i.e.*, the tree pruned of taxa lacking *lpsB*). The *WFCE* and *lpsB* phylogenies were significantly supported at all nodes. The *WFCE* phylogeny grouped sequences from seven out of eight *Epichloë* species in a clade that was basal in the Clavicipitaceae, whereas the sequences from the eighth representative, *E. inebrians*, appeared as the sister to the *Claviceps* clade. Comparing this phylogeny to that of the housekeeping gene, *tefA* (Figure 5), the only significant disparity was the basal placement of the main clade of *Epichloë* *EAS* genes, which contrasted with the *tefA* phylogeny that placed all of the *Epichloë* species together in a clade with a sister relationship to the *Claviceps* clade. There was also a possible disparity in placement of *P. ipomoeae*, but this cannot be considered significant because the relevant branch in the *tefA* tree lacked statistical support. Therefore, with the glaring exception of the main *Epichloë* *EAS* clade, evolution of *EAS* genes in the Clavicipitaceae appears to have been by direct descent without duplication (paralogy) or horizontal gene transfer.

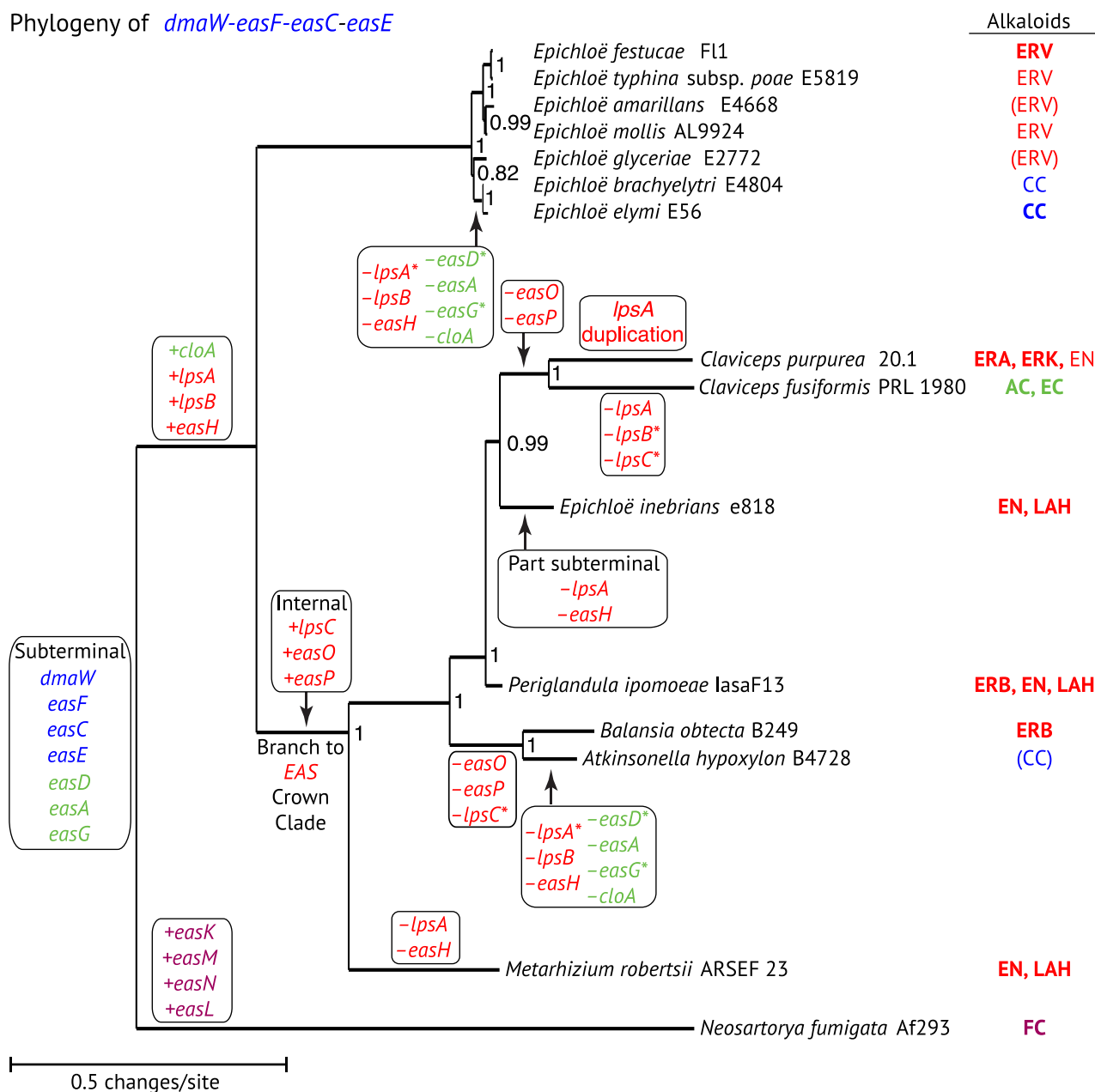
Phylogeny of *dmaW-easF-easC-easE*

Figure 3. Phylogeny of concatenated *dmaW-easF-easC-easE* genes. The phylogenetic tree is based on a nucleotide alignment of coding sequences of the core genes for the first four steps in ergot alkaloid biosynthesis available from sequenced genomes. Sequences were aligned with MUSCLE [41], and trees were inferred by maximum likelihood with PhyML implemented by Phylogeny.fr [40]. Node support was determined by the approximate likelihood ratio test [42]. Gene gains and losses are indicated by + and −, respectively, and asterisks (*) indicate that remnants or pseudogenes can be found in one or more members of the clade. Genes are color-coded based on position of the encoded step within the pathway. The major pathway end product of each strain is indicated on the right in bold face (product produced) or regular type (product predicted but not yet tested), or in parentheses (product predicted but undetected).

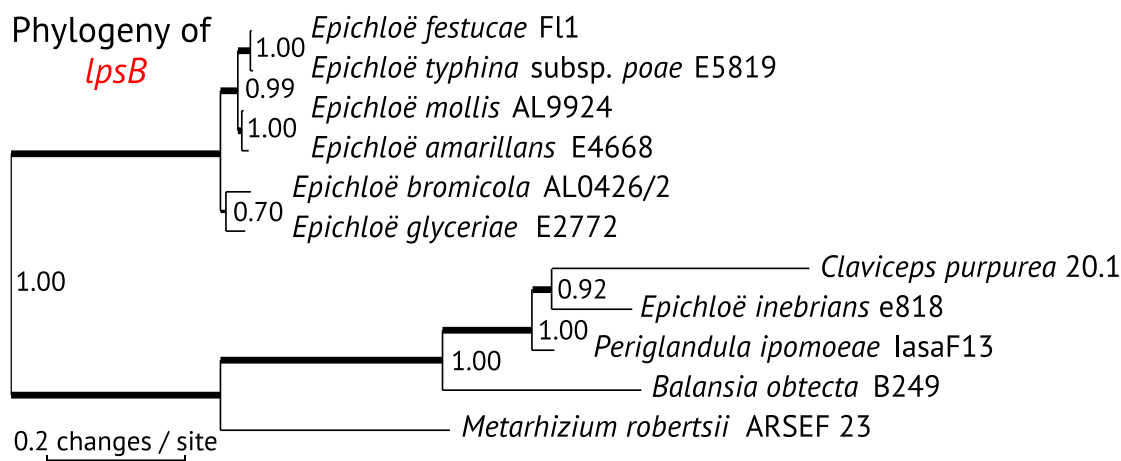


Figure 4. Phylogeny of lysergyl peptide synthetase subunit 2 (*lpsB*). The phylogenetic tree was inferred by maximum likelihood on a nucleotide alignment of coding sequences. Methods are as in Figure 3. The left edge is placed to correspond to the root inferred in Figure 3 with *Neosartorya fumigata* *EAS* genes as the outgroup; *N. fumigata* lacks *lpsB*.

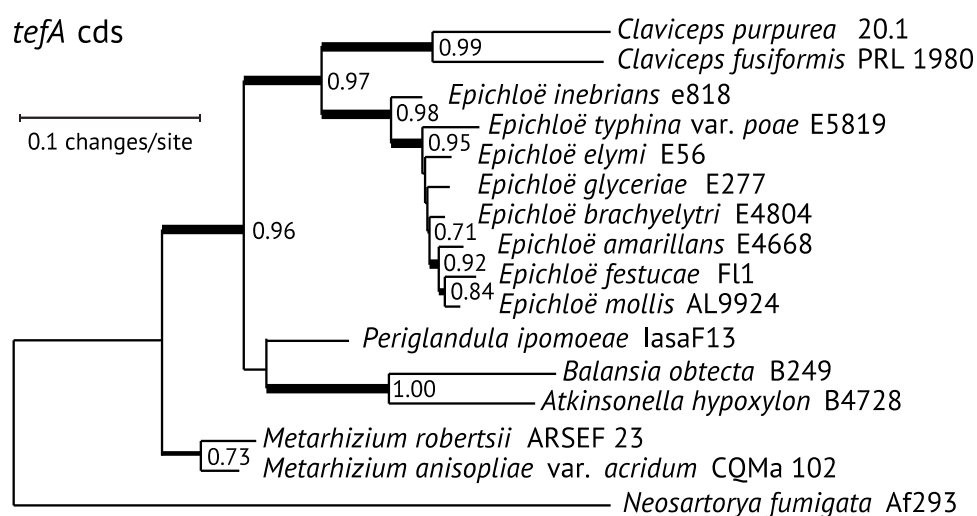


Figure 5. Phylogeny of *tefA*, encoding translation elongation factor 1- α . The phylogenetic tree inferred by maximum likelihood on a nucleotide alignment of coding sequences. Methods are as in Figure 3.

The well-supported position of *EAS* genes from most *Epichloë* species, being basal among the Clavicipitaceae (Figures 3 and 4) indicates a deviation from strict orthology because it differs dramatically from housekeeping gene phylogenies (Figure 5). Possible causes of this deviation are trans-species polymorphism, paralogy or horizontal gene transfer. In a BLASTp search of *dmaW* against available sequences at GenBank, no homologues were closely related to this clade except those of other *Epichloë* species, so there was no obvious source for a horizontal gene transfer event. Comparing the genomic context, sequences nearest the *EAS* clusters differed between *E. festucae* and *E. inebrians* (assemblies of other *Epichloë* species did not link *EAS* clusters with other genes), so trans-species polymorphism also was unsupported. This leaves, as our favored possibility, that the *E. inebrians* and other *Epichloë* *EAS* clusters were derived from paralogous copies that arose from duplication of the *EAS* cluster in an ancestor to most or all of the Clavicipitaceae. In this regard, the *tefA* phylogeny (Figure 5)

placed *E. inebrians* basal in genus *Epichloë*, supporting the possibility that the cluster common to most *Epichloë* species was lost on that basal branch to *E. inebrians*. However, a puzzle remains in that no genus other than *Epichloë* showed indications of paralogous *EAS* clusters. This may be just a matter of limited sampling, and we predict that a wider and deeper survey of the Clavicipitaceae will reveal paralogs related to the basal group of *Epichloë EAS* sequences.

Paralogous *dmaW* genes are in fact evident in some of the Clavicipitaceae. Specifically, *C. purpurea* 20.1 has two *dmaW* copies flanking a paralogous *easF* and located 94 kb and 98 kb from the *EAS*-cluster *dmaW*. *Epichloë mollis* also has a paralogous *dmaW*, which appears to be a pseudogene. However, these paralogues are due to relatively recent duplications and group with the respective *Claviceps* and *Epichloë dmaW* genes in phylogenetic analysis [43].

3.2. Mapping *EAS* Gene Gains and Losses

Mapping genomic alternations onto the *WFCE* phylogeny (Figure 3) revealed repeated instances in which multiple *EAS* genes have been lost. The most extensive losses were eight genes in two separate instances resulting in gene sets for CC production: The branch to *E. elymi* and *E. brachyelytri*, and the branch to *At. hypoxylon*. The identification of remnant or pseudogene copies of other *EAS* genes supported the scenario of extensive gene loss, as inferred from the phylogeny. Losses of multiple *EAS* genes on numerous lineages have given rise to at least four distinct chemotypes in addition to the variations in ergopeptines. These gene losses add to the potential for ergot alkaloid diversification, together with neofunctionalization of *lpsA* (Table 2), and with *easA* variations to give agroclavine or festuclavine as precursors of ergot alkaloids and dihydroergot alkaloids, respectively, and *easG* variations to yield 8(*S*)-dihydroergot alkaloids (Figure 1) [17].

Interestingly, in almost all cases of *EAS* gene loss (Figure 3), all genes were lost or inactivated for a branch of the pathway, leaving only those *EAS* genes required for biosynthesis of the observed metabolites. In two instances, this process has given CC as the end product (or presumed end product), in one instance it has given EN rather than LAH, and in three instances it has eliminated the ergopeptine pathway. The only exception to this pattern is in *C. fusiformis*, which has apparently retained *easH* despite losing all other genes for late pathway steps both to ergopeptines and to EN and LAH. However, whether *easH* is transcribed or gives an active (but presumably useless) product in *C. fusiformis* is unknown. Among many natural strains, gene losses are similarly evident in indole-diterpene (*IDT*) and loline alkaloid (*LOL*) gene clusters with loss or inactivation of all other genes for downstream steps of the affected pathway branch [24,44,45]. Such common patterns suggest that there is selection against expression of enzymes in strains that lack their normal substrate. We speculate that expression of such enzymes may be directly harmful because the enzymes may catalyze reactions with substrates other than the missing natural substrate, thereby generating toxic products.

3.3. Positional Changes of Clusters with Respect to Telomeres

Also mapped to the *WFCE* phylogeny (Figure 3) are three changes in the *EAS* cluster positions relative to telomeres. Considering that the *N. fumigata* and basal *Epichloë EAS* genes are near telomeres (“subterminal”), it is parsimonious to propose this as the ancestral state. If so, internalization would have occurred on the branch that separates most of the Clavicipitaceae from the basal *Epichloë EAS* genes.

An interesting translocation event is evident on the branch to *E. inebrians*, whereby the *EAS* cluster was divided and at least one of the two resulting portions returned to a subterminal position. Interestingly, the *cloA* gene (designated *B* on the map in Figure 2) is situated such that its stop codon is a single base away from the telomere repeat array, implying that the telomere serves as a transcription terminator for that gene. The translocation event on the *E. inebrians* branch nicely illustrates the dynamics of subterminal regions of the chromosomes. Recalling that the *IDT* cluster—directing biosynthesis of the indole-diterpene, paxilline—is subterminal in the closely related *E. gansuensis* strain E7080 [24], it is particularly interesting to find two remnant *IDT* genes flanking the centromeric side of the subterminal *EAS* cluster in *E. inebrians*. It appears likely that the *EAS* cluster displaced most of the *IDT* cluster as it moved into the subterminal position. This supports an important role of subterminal regions in arrangements and rearrangements of secondary metabolism gene clusters. With that in mind, we would predict that most *EAS* cluster rearrangements occur in association with subterminal clusters, whereas internal *EAS* clusters are more stable. Evidence pertinent to that prediction is discussed later, in Section 5.

3.4. Evolution of *LpsC*

Assuming that the common ancestor of Clavicipitaceae and Trichocomaceae possessed the seven genes for early and intermediate biosynthetic steps, a parsimonious scenario would identify three branches for gene gains: one to *N. fumigata*, one to all Clavicipitaceae, and the third to Clavicipitaceae after splitting from the basal *Epichloë* *EAS* clade (Figure 3). In the context of our proposed paralogy, the last of these would have involved acquisition of the new genes, *lpsC*, *easO* and *easP*, on one of the branches after that split in the Clavicipitaceae. Interestingly, *LpsC* serves as an alternative to the *LpsA* subunit in interacting with *LpsB* as a component of a lysergyl peptide synthetase complex (Figure 1). A phylogenetic analysis of *Lps* subunit A domains (Figure 6) indicates that *lpsC* may be derived from a copy of the first module of *lpsA* fused to an R* (reductase) coding sequence. Thus, biosynthesis of EN and LAH seems to have evolved later than biosynthesis of the much more complex ergopeptides.

3.5. Evolution of Module Specificity in *LpsA*

The evolution of module specificity in *LpsA* subunits can also be mapped onto the *WFCE* (Figure 3) and *lpsA* A-domain (Figure 6) phylogenies. However, a crucial difference between *WFCE* and *lpsA* phylogenies is that *lpsA* A-domains consistently group *P. ipomoeae* and *B. obtecta* in a terminal clade, whereas *WFCE* groups *P. ipomoeae* with *Claviceps* spp. and places *B. obtecta* in a relatively basal position. This may be the result of *lpsA* paralogy due to duplications and losses (a scenario reminiscent of the tandemly duplicated *lpsA* genes in *C. purpurea*), or may simply be due to lineage sorting effects in the evolution of these lineages (as suggested by the short and poorly supported branch in the *tefA* phylogeny). To discuss *LpsA* module evolution, we now use single-letter abbreviations for the three specified amino acids in order of the modules 1, 2 and 3. The ancestral state could have been ALP or AVP, with module A2 changing specificity either on the basal *Epichloë* branch or on the lineage leading to the *Claviceps*/*Periglandula*/*Balansia* clade, respectively. During evolution of the *Periglandula*/*Balansia* clade, specificity of module 3 switched from P to A. What is especially interesting is that, within *Claviceps*, there is an accelerated diversification of modules 1 and 2, such that

The figure displays a phylogenetic tree and a corresponding domain architecture diagram for Lysophosphatidylserine (LPS) synthase (Lps) proteins. The phylogenetic tree is rooted at the bottom left, showing relationships between various species and their Lps variants. Bootstrap values are indicated at the nodes. The tree is divided into several major clades, each associated with a specific Lps protein or group of proteins. The domain architecture diagram on the right shows the organization of these proteins, with domains represented by colored boxes and arrows indicating the direction of the protein sequence.

Lps AMPylation domains

- Metarhizium robertsii* ARSEF 23
- Epichloë inebrians* e818
- Periglandula ipomoeae* IasaF13
- Claviceps purpurea* 20.1
- Claviceps paspali* RRC-1481
- Epichloë festucae* Fl1
- Balsania oblecta* B249
- Periglandula ipomoeae* IasaF13
- Claviceps purpurea* 20.1
- Claviceps purpurea* 20.1
- Claviceps africana* Cla9
- Epichloë festucae* Fl1
- Balsania oblecta* B249
- Periglandula ipomoeae* IasaF13
- Claviceps purpurea* 20.1
- Claviceps purpurea* 20.1
- Claviceps africana* Cla9
- Epichloë festucae* Fl1
- Balsania oblecta* B249
- Periglandula ipomoeae* IasaF13
- Claviceps purpurea* 20.1
- Claviceps purpurea* 20.1
- Claviceps africana* Cla9
- Epichloë festucae* Fl1
- Metarhizium robertsii* ARSEF 23
- Balsania oblecta* B249
- Epichloë inebrians* e818
- Periglandula ipomoeae* IasaF13
- Claviceps purpurea* 20.1
- Claviceps paspali* RRC-1481
- Claviceps africana* Cla9

LpsC

- LpsC (Ala)
- LpsC (Ala)
- LpsC (Ala)
- LpsC (Ala)
- LpsA^{AVP} (Ala)
- LpsA^{ALA} (Ala)
- LpsA^{ALA} (Ala)
- LpsA^{AFP} (Ala)
- LpsA^{VLP} (Val)
- LpsA^{ALP} (Ala)
- LpsA^{AVP} (Pro)
- LpsA^{ALA} (Ala)
- LpsA^{ALA} (Ala)
- LpsA^{AFP} (Pro)
- LpsA^{VLP} (Pro)
- LpsA^{ALP} (Pro)
- LpsA^{AVP} (Val)
- LpsA^{ALA} (Leu)
- LpsA^{ALA} (Leu)
- LpsA^{AFP} (Phe)
- LpsA^{VLP} (Leu)
- LpsA^{ALP} (Leu)
- LpsB (LA)
- LpsB
- LpsB (LA)
- LpsB (LA)
- LpsB (LA)
- LpsB (LA)
- LpsB (LA)
- LpsB (dihydroLA)

Domain Architecture Diagram

The diagram illustrates the domain architecture of the Lps proteins. It shows a series of domains represented by colored boxes (red, green, blue, purple) and arrows indicating the direction of the protein sequence. The domains are labeled as follows:

- A1**: LpsA^{AVP} (Ala), LpsA^{ALA} (Ala), LpsA^{ALA} (Ala), LpsA^{AFP} (Ala), LpsA^{VLP} (Val), LpsA^{ALP} (Ala)
- A3**: LpsA^{AVP} (Pro), LpsA^{ALA} (Ala), LpsA^{ALA} (Ala), LpsA^{AFP} (Pro), LpsA^{VLP} (Pro), LpsA^{ALP} (Pro)
- A2**: LpsA^{AVP} (Val), LpsA^{ALA} (Leu), LpsA^{ALA} (Leu), LpsA^{AFP} (Phe), LpsA^{VLP} (Leu), LpsA^{ALP} (Leu)
- LpsB**: LpsB (LA), LpsB, LpsB (LA), LpsB (LA), LpsB (LA), LpsB (LA), LpsB (LA), LpsB (dihydroLA)

Scale bar: 0.20 changes/site

Figure 6. Phylogeny of the Lps subunit AMPylation domains. The phylogenetic tree was inferred by maximum likelihood on a nucleotide alignment of coding sequences. Methods are as in Figure 3. The specified substrates are given in parentheses as LA = lysergic acid, dihydroLA = dihydrolysergic acid, and standard abbreviations for common L-amino acids. The LpsA superscripts indicate single-letter codes for the amino acids specified by AMPylation domains of module 1, 2 and 3 (A1, A2 and A3), respectively. Functionality and specificity of *M. robertsii* LpsB and LpsC are unknown. The *P. ipomoeae* EAS cluster is shown at right with Lps genes and modules color-coded.

4.1. Distribution of EAS Genes across *Epichloë* Species

The distribution of the *EAS* locus is well understood within *Epichloë* species due to the availability of genomic and genotyping data (www.endophyte.uky.edu). Genome sequences (including draft genomes) for 54 isolates representing 10 described *Epichloë* species, two varieties, and five undescribed species have enabled comparisons of the *EAS* locus across a diverse collection of isolates with varying capabilities to produce ergot alkaloids [24,34,38,43,45,46]. Moreover, in recent studies, markers developed from the genome sequences have been used for PCR-based genotyping of endophyte-infected grass

collections for the presence of genes encoding key alkaloid biosynthesis pathway steps [10,47–52]. Compared to genome sequencing, PCR-based genotyping is less reliable in determining if a gene is functional, but the presence of full-length genes typically associated with *EAS* clusters for the metabolites CC, ERV or EN usually corresponds well to the production of the corresponding metabolite. The predictive power of this method is helped by the tendency for complete losses or large deletions in the nonfunctional alkaloid biosynthesis genes, and for the downstream genes to be partly or completely deleted as well [24].

The presence of *EAS* genes within *Epichloë* species has been observed in strains of at least 19 (9 nonhybrid and 10 hybrid species) of the 35 taxa tested (including varieties and undescribed species) (Table 3). The *EAS* locus has a discontinuous distribution within *Epichloë* species. For example, not all *E. festucae* isolates are capable of producing ergovaline, since some isolates (e.g., *E. festucae* E434) lack the *EAS* locus. However, in addition to their discontinuous distribution, the *EAS* loci can differ in structural content. Strains capable of producing CC (e.g., *E. elymi* E56) only contain functional genes encoding the first four *EAS* pathway steps (DmaW, EasF, EasC and EasE) (Figure 1; Table 1), whereas ergovaline producers (e.g., *E. festucae* F11) have more complex *EAS* loci (*EAS^{ERP}*) with at least 11 functional genes present (Table 1).

Many of the chemotype differences identified within the *Epichloë* species can be attributed to the presence or absence of genes encoding the key pathway steps, but there are some anomalies. Isolates, such as *E. festucae* strains E2368 and E189, *E. amarillans* E4668 and *E. glyceriae* E2772 appear to have complete *EAS^{ERP}* clusters with no obvious deleterious mutations, yet ergot alkaloids have not been detected in host plants symbiotic with these isolates. Transcriptome (RNA-seq) analysis of E2368-infected meadow fescue (*Lolium pratense*) and tall fescue show that the *EAS* genes are not expressed in this strain, thus providing a clue to why the predicted alkaloid, ERV, is not produced [24].

Many endophytes can readily produce ergot alkaloids *in planta* but have a more limited and unreliable production in non-symbiotic culture conditions [53,54]. Histone methylation apparently helps repress expression of *EAS* and other alkaloid gene clusters when *Epichloë* species are grown in culture, since the *EAS* and *LTM* genes of *E. festucae* strain F11 were de-repressed under non-symbiotic culture conditions when two genes that encode histone H3 methylases were deleted [55].

Table 3. *EAS* gene distribution within and between *Epichloë* species.

<i>Epichloë</i> species ^a	Host species	Detection method ^b	<i>EAS</i> gene variations (strains observed) ^c	Reference
<i>Epichloë amarillans</i>	<i>Agrostis hyemalis</i>	GT, DG, G	0* (4), (ERV) (1)	[24]
<i>E. aotearoae</i>	<i>Echinopogon ovatus</i>	G	0 (1)	[24,43]
<i>E. baconii</i>	<i>Agrostis tenuis</i> , <i>Calamagrostis villosa</i>	GT, G	0* (3)	[43]
<i>E. brachyelytri</i>	<i>Brachyelytrum erectum</i>	GT, G	0 (1), CC (3)	[24]
<i>E. bromicola</i>	<i>Bromus erectus</i> , <i>Bromus benekenii</i> , <i>Bromus tomentellus</i> , <i>Agropyron hispidus</i>	GT, DG, G	0* (5)	[43]
<i>E. cabralii</i> (H)	<i>Phleum alpinum</i> <i>Bromus laevipes</i>	G, GT	0 (1), (ERV) (2)	[50]
<i>E. canadensis</i> (H)	<i>Elymus canadensis</i>	GT, DG	CC (1), ERV (1)	[43,47]
<i>E. chisosa</i> (H)	<i>Achnatherum eminens</i>	DG	0 (1)	[43]
<i>E. coenophiala</i> (H)	<i>Lolium arundinaceum</i>	GT, DG	0* (11), ERV (12) , ERV (39)	[43,49,51]
<i>E. elymi</i>	<i>Elymus virginicus</i>	GT, G	0 (1), CC (1)	[24]
<i>E. festucae</i>	<i>Festuca trachyphylla</i> , <i>Festuca rubra</i> subsp. <i>rubra</i> , <i>Lolium giganteum</i>	GT, G	0 (1), ERV (1) , (ERV) (2)	[24]
<i>E. festucae</i> var. <i>lolii</i>	<i>Lolium perenne</i>	GT, G	ERV (2) , (ERV) (1)	[56,57]
<i>E. festucae</i> var. <i>lolii</i> x <i>E. typhina</i> (H)	<i>Lolium perenne</i>	DG	ERV (1)	[43,58]
<i>E. funkii</i> (H)	<i>Achnatherum robustum</i>	GT, DG	CC (1)	[43]
<i>E. gansuensis</i>	<i>Achnatherum inebrians</i>	G	0 (1)	[24]
<i>E. inebrians</i>	<i>Achnatherum inebrians</i>	G	EN, LAH (1)	[24]
<i>E. glyceriae</i>	<i>Glyceria striata</i>	GT	(ERV) (2)	[24]
<i>E. mollis</i>	<i>Holcus mollis</i>	G	ERV (1)	[43]
<i>E. occultans</i> (H)	<i>Lolium</i> sp. (2x)	GT	0 (3)	[43]
<i>E. schardlii</i> (H)	<i>Cinna arundinacea</i>	GT	0 (1)	[59]
<i>E. siegelii</i> (H)	<i>Lolium pratense</i>	DG	0 (1)	[43]

Table 3. Cont.

<i>Epichloë</i> species ^a	Host species	Detection method ^b	<i>EAS</i> gene variations (strains observed) ^c	Reference
<i>E. sylvatica</i>	<i>Brachypodium sylvaticum</i>	GT	0 (2)	[34]
<i>E. typhina</i>	<i>Lolium perenne</i> , <i>Dactylis glomerata</i>	G, GT	0 (3)	[24,43]
<i>E. typhina</i> ssp. <i>clarkii</i>	<i>Holcus lanatus</i>	GT	ERV (1)	unpublished
<i>E. typhina</i> ssp. <i>poae</i>	<i>Poa nemoralis</i> , <i>Bromus laevipes</i>	GT, G	0 (3), ERV (1)	[24,50]
<i>E. uncinata</i> (H)	<i>Lolium pratense</i>	DG	0 (1)	[43]
<i>E. sp.</i> AroTG-2 (H)	<i>Achnatherum robustum</i>	GT	EN (1)	[10]
<i>E. sp.</i> BlaTG-3 (H)	<i>Bromus laevipes</i>	GT	0* (1), CC (2)	[50]
<i>E. sp.</i> FaTG-2 (H)	<i>Lolium sp.</i> (6x)	GT, DG	ERV (10) , ERV (33)	[43,49,51,60]
<i>E. sp.</i> FaTG-3 (H)	<i>Lolium sp.</i> (6x), (8X)	GT, DG	0 (11)	[43,51,60]
<i>E. sp.</i> FaTG-4 (H)	<i>Lolium sp.</i> (10x)	GT, DG	ERV (1) , ERV (11)	[43,51]
<i>E. sp.</i> FcaTG-1 (H)	<i>Festuca campestris</i>	GT	0 (3)	unpublished
<i>E. sp.</i> FveTG-1 (H)	<i>Festuca versuta</i>	GT	0 (2)	unpublished
<i>E. sp.</i> PalTG-1 (H)	<i>Poa alsodes</i>	GT	0* (1)	unpublished
<i>E. sp.</i> PauTG-1 (H)	<i>Poa autumnalis</i>	GT	0 (1)	unpublished

^a Endophytes that are known hybrids = (H); ^b Detection methods for the *EAS* genes DG = draft genome, G = genome, GT = PCR-based genotyping; ^c Number of independent strains evaluated. Alkaloids are abbreviated CC = chanoclavine I, ERV = ergovaline, EN = ergonovine, LAH = lysergic acid α -hydroxyethylamide, and are in bold face (product produced) or regular type (product predicted but not yet tested), or in parentheses (product predicted but undetected). 0 = No *EAS* genes identified, 0* = contains only remnants of *EAS* clusters; 0 and 0* are unable to produce ergot alkaloids.

4.2. Pseudogenes and Gene Remnants within the EAS Locus

Pseudogenes and gene remnants have been identified within or adjacent to many of the described *EAS* clusters within the Clavicipitaceae. Most *Epichloë* species that are unable to produce ergot alkaloids lack functional *EAS* genes or only retain remnants of some *EAS* genes. For example, a remnant *lpsA* sequence is present in the genome sequences of *E. amarillans* strain E57, *E. bromicola* strain AL0434, and *Epichloë coenophiala* strains e4509, AR542 and AR584. Typically these *lpsA* remnants have multiple frameshifts, stop codons or both, and are flanked or even disrupted by AT-rich repeat sequences; or, in the case of E57, the gene is truncated at the telomere [24]. In comparison to *E. bromicola* strain AL0434, which has only an *lpsA* remnant, strain AL0426/2 has retained a greater number of *EAS* genes (Figure 7). The *E. bromicola* AL0426/2 genome assembly (www.endophyte.uky.edu) contains *lpsB*, *easE*, *easF* and *easG* (contig 634), but *easE* is a pseudogene in which part of the coding sequence is missing. Since *dmaW* and *easC* are also missing, the AL0426/2 isolate is not expected to produce ergot alkaloids. The identification of two independent *EAS* gene loss events in *E. bromicola* (strains AL0434 and AL0426/2) suggests that this species may have greater *EAS* diversity than is currently recognized, but *E. bromicola* strains capable of producing ergot alkaloids have not yet been identified [61].

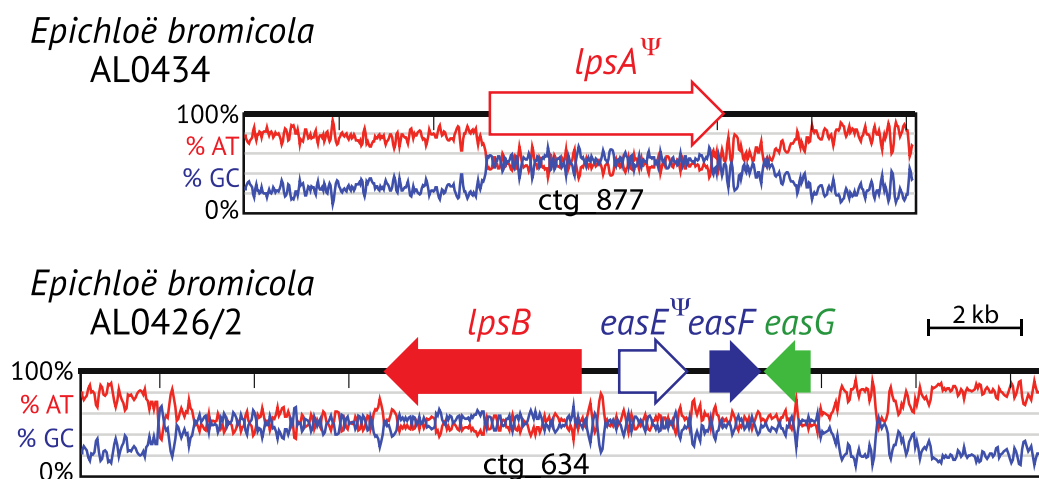


Figure 7. Remaining *EAS* genes and pseudogenes after independent losses in two *E. bromicola* isolates, AL0434 and AL0426/2. The AT-GC DNA contents are shown under the maps. Pseudogenes are represented by Ψ.

4.3. Hybrids: *EAS* Gene Cluster Variations

Compared to sexual strains or other haploids, the hybrid *Epichloë* species have a greater potential to contain *EAS* genes because each of the ancestors contributes genetic information. In addition, variations of the *EAS* locus within a hybrid can be due to copy number or differences within the inherited *EAS* cluster (*EAS*^{CC} vs *EAS*^{ERP} clusters). Some, but not all isolates from the hybrid species, *Epichloë canadensis*, *E. coenophiala*, and *E. sp.* FaTG-2, contain two *EAS* copies (Figure 8). The *E. canadensis* isolate e4815 contains two *EAS* clusters—*EAS*^{ERP} for ERV and *EAS*^{CC}—that are representative of the two contributing ancestors, *E. amarillans* and *E. elymi*, respectively. In contrast, *E. canadensis* isolate CWR34 contains a single *EAS*^{CC} cluster, contributed by its *E. elymi* ancestor. Mating-type gene differences between *E. canadensis* isolates e4815 and CWR34 clearly indicate that

they are the result of independent hybridizations, but it is unclear if CWR34 subsequently lost the *E. amarillans* *EAS* cluster, or if its particular ancestral strain of *E. amarillans* lacked *EAS* genes.

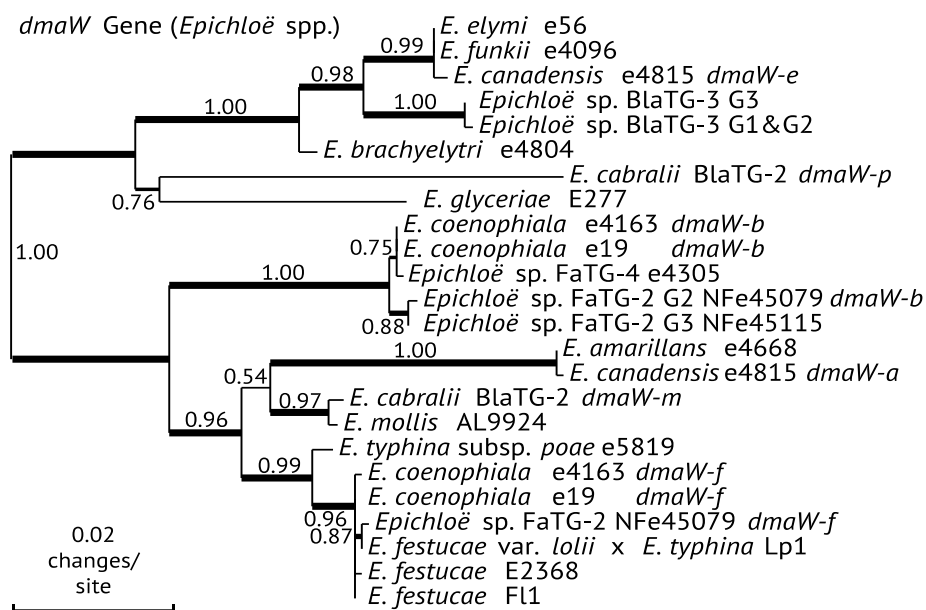


Figure 8. Phylogeny of *dmaW* genes of *Epichloë* strains. The phylogenetic tree was inferred by maximum likelihood on a nucleotide alignment of coding sequences. Methods are as in Figure 3. The *dmaW* alleles are distinguished in hybrids that possess more than one copy with a letter that refers to the ancestral progenitor (*a* = *E. amarillans*, *b* = *E. baconii*-related *Lolium* associated *Epichloë* subclade, *e* = *E. elymi*, *f* = *E. festucae*, *m* = *E. mollis*-related and *p* = *E. typhina* subsp. *poae*). The *dmaW* gene of *E. inebrians* has been omitted in this analysis because the gene and *EAS* locus is more similar to that of *P. ipomoeae* than to those of other *Epichloë* species (see Figure 3).

4.4. Gene Losses in Hybrids with Multiple Copies

Some isolates of *E. coenophiala* are unable to produce ergot alkaloids because they only contain a remnant *EAS* locus (e.g., e4309, AR542 and AR584; Table 3) [51]. The *E. coenophiala* isolates known to produce ERV, such as isolates e19 and e4163, have two copies of the *EAS* clusters from two of the three contributing ancestors: *E. festucae* and the *Lolium*-associated *Epichloë* (LAE) subclade. Interestingly, the *EAS* clusters of both *E. coenophiala* isolates e19 and e4163 are structurally very similar with respect to AT content and repeat sequences, though they differ in which *EAS* genes are absent or inactive pseudogenes in each *EAS* cluster (Figure 9). The *lpsB2* gene in e19 (from the *E. festucae* ancestor) is nonfunctional due to a frame shift within the coding region. The situation in e4163 is more complex in that neither *EAS* cluster is complete; *EAS1* (from LAE) lacks *lpsB1* and *easE1*, and *EAS2* (from *E. festucae*) includes a nonfunctional *easG2*. Therefore, since e4163 is able to produce ERV, each *EAS* cluster must functionally complement the genetic deficiencies of the other cluster. *Epichloë* sp. FaTG-2 isolates NFe45115 and NFe45079 are both capable of producing ERV, but *EAS* copy numbers differ; NFe45115 has one copy, whereas NFe45079 has two copies [49]. Recent genome sequence data have also revealed that NFe45079 contains only a single copy of *lpsA*, but all other genes are present in two copies.

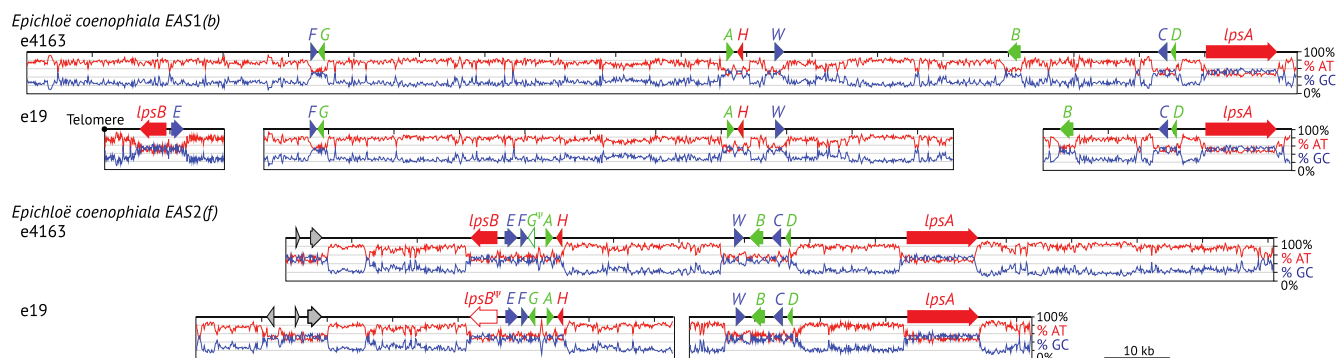


Figure 9. Structures of the *EAS* clusters from two *E. coenophiala* strains, e19 and e4163. The AT-GC contents are shown under the maps. Gene names are abbreviated as in Figure 2.

4.5. Endophyte Genetic Variation within a Single Host Species

The associations of endophytes and hosts represent co-evolving associations as evidenced by the tendency for evolution of the *Epichloë* spp. to track evolution of their hosts [62]. It is interesting to note that sometimes a single host species has symbiotic associations with more than one endophyte species, but each individual plant will only host one endophyte genet. Tall fescue can be symbiotic with *E. coenophiala*, *Epichloë* sp. FaTG-2, FaTG-3 and FaTG-4, and strains representing each of these endophytes have different alkaloid profiles [43,51,60]. *Elymus canadensis* can be symbiotic with *E. elymi* and *E. canadensis*, and each of these endophytes also exhibits chemotype variation [47]. As we develop and refine high throughput methods to explore large host collections more thoroughly, more endophyte variation may be identified within a single host species.

Recently *Bromus laevipes*, a bunchgrass native to California, was found to have independently formed symbiotic associations with three *Epichloë* species, the nonhybrid *E. typhina* subsp. *poae* (designated *Bromus laevipes* Taxonomic Group 1; BlaTG-1), a hybrid designated BlaTG-2, which appears to be phylogenetically similar to *E. cabralii* from *Phleum alpinum*, and another hybrid designated BlaTG-3 [50]. Endophyte diversity within this host collection was identified by PCR-based genotyping of the *EAS*, *LOL*, indole-diterpene/lolitre (IDT/LTM) and peramine (*PER*) loci. The BlaTG-3 isolates could be separated into three genotypes (G1, G2, and G3) based on the *EAS* and mating type complement. Isolates considered BlaTG-3 G1 only contained *dmaW*, whereas BlaTG-3, G2, and G3, had different mating-type genes, but shared the same *EAS* gene complement (*EAS^{CC}*) in keeping with its ability to produce CC. In contrast, the complete *EAS* gene complement required for ERV production (*EAS^{ERP}*) was identified within BlaTG-2, yet ergot alkaloids were not detected in plants with BlaTG-2. RT-PCR expression analyses of the *EAS* genes from BlaTG-2-infected plants indicated they were not expressed. Interestingly, although BlaTG-2 and *E. cabralii* from *Phleum alpinum* are hybrids with the same two ancestral *Epichloë* species, no *EAS* genes are present in characterized *E. cabralii* strains.

Sleepygrass plants growing in the vicinity of Cloudcroft, New Mexico, U.S.A., are renowned for their narcotic effects on livestock because they can have high levels of the ergot alkaloids ergine and EN [63]. It is now clear that two endophyte species can be symbiotic with sleepygrass, *E. funkii* [64] and a so-far undescribed hybrid *Epichloë* species designated AroTG-2 [10]. Each of these endophytes has the *EAS*

complement and ability to produce different ergot alkaloids, either CC (EAS^{CC}) or ergine and EN (EAS^{EN}), respectively [10]. Phylogenetic analysis of *dmaW* groups the *E. funkii* gene in a clade with *dmaW* of the *E. festucae* clade (Figure 8). In contrast, AroTG-2 has a *dmaW* sequence more similar to that of *E. inebrians* from drunken horse grass than to that of other *Epichloë* species (data not shown), which is in keeping with the similarity of the alkaloids produced by these two endophyte species and their strong stupor-inducing effects on grazing livestock [11,63].

5. Synteny and Rearrangements in the *EAS* Loci

5.1. Syntenic Regions of the *EAS* Loci

Although functional *EAS* genes are always clustered, they are not always in a single cluster, arrangements of the genes are not highly conserved, and locations of the clusters can vary, some being subterminal (near chromosome ends), and others being internal and flanked on both sides by long regions rich in housekeeping genes (Figure 10). With the exception of the *Epichloë* species (discussed below), the degree of divergence in *EAS* gene arrangements, gene contents, and the pathway end-products generally relates to the degree of divergence between species. Thus, it is unsurprising that the *EAS* cluster arrangements and gene contents differ greatly in comparison of *N. fumigata* to the Clavicipitaceae. A surprisingly consistent feature is the arrangement and close linkage of *easE* and *easF* in all except *E. elymi* E56; even their *Ar. benhamiae* orthologues, respectively designated ARB_4648 and ARB_4647, are adjacent but arranged tail to tail [18]. However, *EAS* gene arrangements and orientations are very similar in what we will call the “crown *EAS* clade” (Figure 3): *Metarhizium* spp. (including *Metarhizium acridum*, which is not shown), *P. ipomoeae*, *B. obtecta*, *At. hypoxylon*, *Claviceps* spp. and *E. inebrians*. Differences within the crown *EAS* clade were as follows: (1) an inversion of *lpsB-easA* segment in *C. fusiformis* relative to the others; (2) separation of the *easH-lpsA* segment from others in *B. obtecta*, due to an event that (based on remnant genes) occurred in a common ancestor of *B. obtecta* and *At. hypoxylon*; (3) breakage of the cluster in *E. inebrians* by a telomere introduced immediately downstream of *cloA*; and (4) several gene losses or inactivations that have resulted in changes in ergot alkaloid profiles, as discussed above.

Despite the conserved arrangement of *EAS* genes in the aforementioned crown *EAS* clade, there is very limited synteny of flanking genes (Figure 2), except within the *Claviceps* subclade and, separately, within the *B. obtecta/At. hypoxylon* subclade. The only group of orthologous genes that flanks most members of the crown *EAS* clade is represented in *C. purpurea* 20.1 by AET79176 (GenBank accession number; labeled with asterisks in Figure 2) and is similarly positioned in all except *E. inebrians*, which has its AET79176 orthologue at the opposite end of the cluster. With the possible exceptions of *At. hypoxylon* B4728 and *E. festucae* E2368, the AET79176 orthologue was linked to *EAS* in every strain that had one. In B4728, it is also possible that the gene is linked to the *EAS* cluster, but at >75 kb from *easC* (the AET79176 orthologue being near the middle of the 123,479-bp contig00086, which is otherwise very AT-rich and lacks other identifiable genes). So far, there is no putative function or conserved signature domain for AET79176, but it might warrant future investigation for a possible role in the regulation of *EAS* genes or in a biosynthetic role yet to be identified.

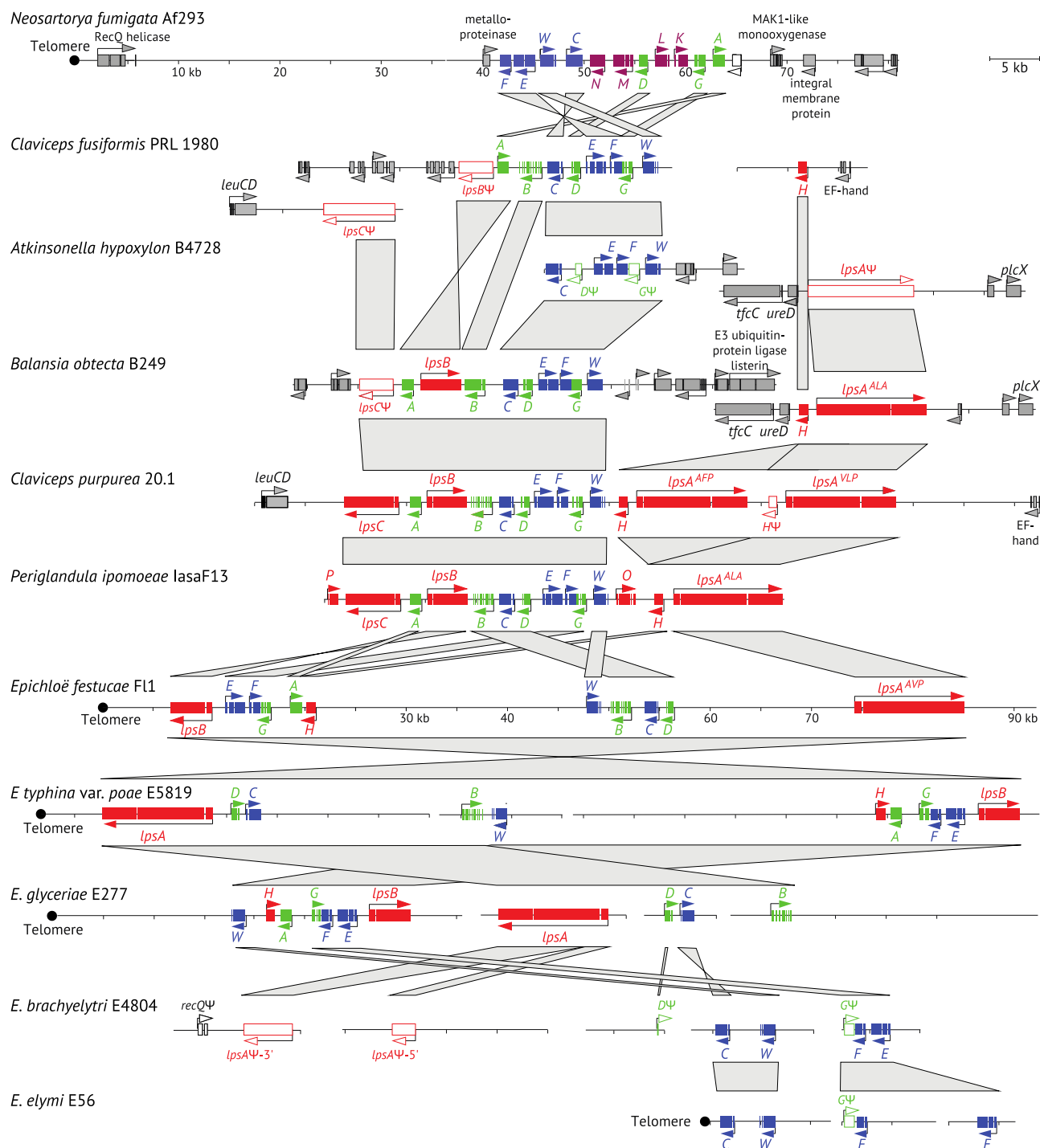


Figure 10. Structures of representative *EAS* loci showing synteny of *EAS* genes between species. Genes are colored to represent the stage of the pathway for the encoded product (see Figures 1 and 2). Pseudogenes are represented by Ψ and white-filled arrows. Gray polygons link orthologous genes and gene blocks but are not meant to imply particular phylogenetic relationships. The *EAS* crown clade includes clusters from *At. hypoxylon*, *B. obtecta*, *C. purpurea*, *C. fusiformis* and *P. ipomoeae*.

5.2. *Epichloë* Species Have More *EAS* Loci Rearrangements

The *Epichloë* species show the greatest variation in cluster organization (Figure 10). Even *EAS*^{ERP} clusters, which possess apparently functional forms of 11 *EAS* genes, show extreme rearrangements of

gene positions relative to each other and to the telomeres. (The actual linkage and order of the contigs containing *EAS* genes was established for *E. festucae* strains E2368 and F11 [24] but could not be determined for other genome sequences when not assembled into scaffolds.) Also highly variable among *Epichloë* *EAS* clusters was the organization of their extensive, AT-rich, repetitive sequences, which may directly facilitate cluster instability and rearrangements, and in some cases cause partial or entire deletion of genes giving, for example, *EAS^{CC}* clusters (Figures 2 and 9) [24]. In addition, within the *Epichloë* species, there is a strong tendency for the *EAS* locus to be retained at the subterminal region, even though the order of genes relative to the telomere (chromosome end) can differ.

Miniature inverted-repeat transposable elements (MITEs) are prevalent in *Epichloë* species [57], and have been identified in the promoters of some *EAS* genes with expansion in the *EAS^{CC}* clusters associated with *dmaW* and *easC*. The repetitive AT-rich sequences and prevalence of MITEs are also associated with the gene clusters for biosynthesis of other alkaloid classes; namely, *IDT/LTM* (indole-diterpenes) and *LOL* (lolines) [24]. The presence of long transposon-derived repeat blocks seems consistent with the subterminal location of telomere-linked clusters like *EAS* and *IDT/LTM*, but the fact that they also feature prominently in the *Epichloë* *LOL* clusters, which are typically internal with extensive genic regions flanking both ends, suggests that this is a more general feature of alkaloid clusters [24].

5.3. The Complex History of *EAS* Loci

Although not by any means restricted to subtelomeric and subterminal regions, blocks of transposon-derived repeats are features of these genomic regions in fungi [65,66], and such regions are prone to considerable instability and gene duplication events [67,68]. A subterminal location appears to be the ancestral state of the *EAS* cluster, considering that it is a shared feature of *N. fumigata* and the basal *EAS* clade in Clavicipitaceae; namely, the clade comprised of the majority of *Epichloë* *EAS* clusters (Figure 3). Thus, increased stability, particularly of the *EAS* core containing early- and mid-pathway genes, seems to be in keeping with the shift from subterminal to internal location in the common ancestor of the crown *EAS* clade. Nevertheless, the differences in flanking housekeeping genes between the *Claviceps* subclade and the *B. obtecta*/*At. hypoxylon* subclade, plus indications of gene duplication in *C. purpurea*, indicate additional complexity in the history of the *EAS* clusters. In *C. purpurea*, the duplication of *lpsA* has enabled its neofunctionalization to greatly enhance the diversity of ergopeptine products (Figure 10). Interestingly, 55–73 kb downstream of *lpsC* in the *C. purpurea* genome are two additional copies of *dmaW* and *easF* (Figure 11), suggesting that gene duplication has been a particularly dynamic evolutionary process in *C. purpurea* in addition to providing an attractive explanation for the fact that most of the known ergopeptin(in)es are reported from this species. Furthermore, the *dmaW* and *easF* duplications are close to a *recQ* helicase pseudogene, and reflecting their typical locations, *recQ* helicase genes are also called telomere-linked helicase genes (*TLH*). (For example, a *recQ* helicase gene is located near the *EAS*-linked telomere of *N. fumigata*, as shown in Figures 2 and 9). In *C. purpurea*, the association of a *recQ* helicase pseudogene with duplicated *EAS* genes and in the vicinity of the *EAS* cluster suggests more recent evolutionary history associated with a chromosome end than implied in our phylogenetic inferences (Figure 3). Thus, repeated shifts between subterminal and internal locations may have characterized *EAS* clusters in the Clavicipitaceae, and perhaps especially in *C. purpurea* as a driver of ergopeptine diversification.

Acknowledgments

We thank Ginger A. Swoboda (Noble Foundation), Juan Pan (UKY) and Li Chen (UKY) for genotyping of *Epichloë* species. We thank Jennifer A. Rudgers for plant material, Leopoldo J. Iannone for *E. cabralii* and Martina Oberhofer and Stanley H. Faeth for the sleepygrass endophyte. This work was supported by USDA-CSREES grant 2009-34457-20125, USDA-CSREES Grant 2010-34457-21269, USDA-NIFA grant 2012-67013-19384, NSF grant EPS-0814194, 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 15-12-037 of the Kentucky Agricultural Experiment Station, published with approval of the director.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Billings, M. *The Crusades: Five Centuries of Holy Wars*; Sterling Publishing Company: New York, NY, USA, 1996.
2. Alm, T. The witch trials of Finnmark, northern Norway, during the 17th century: Evidence for ergotism as a contributing factor. *Econ. Bot.* **2003**, *57*, 403–416.
3. Caporael, L.R. Ergotism: the Satan loosed in Salem? *Science* **1976**, *192*, 21–26.
4. Woolf, A. Witchcraft or mycotoxin? The Salem witch trials. *J. Toxicol.: Clin. Toxicol.* **2000**, *38*, 457–460.
5. Zadoks, J.C. *On the Political Economy of Plant Disease Epidemics: Capita Selecta in Historical Epidemiology*; Wageningen Academic Publishers: Wageningen, The Netherlands, 2008.
6. Tribble, M.A.; Gregg, C.R.; Margolis, D.M.; Amirkhan, R.; Smith, J.W. Fatal ergotism induced by an HIV protease inhibitor. *Headache* **2002**, *42*, 694–695.
7. Urga, K.; Debella, A.; W’Medihn, Y.; Bayu, A.; Zewdie, W. Laboratory studies on the outbreak of gangrenous ergotism associated with consumption of contaminated barley in Arsi, Ethiopia. *Ethiop. J. Health Dev.* **2002**, *16*, 317–323.
8. Scott, P. Ergot alkaloids: Extent of human and animal exposure. *World Mycotoxin J.* **2009**, *2*, 141–149.
9. Lyons, P.C.; Plattner, R.D.; Bacon, C.W. Occurrence of peptide and clavine ergot alkaloids in tall fescue grass. *Science* **1986**, *232*, 487–489.
10. Shymanovich, T.; Saari, S.; Lovin, M.; Jarmusch, A.; Jarmusch, S.; Musso, A.; Charlton, N.; Young, C.; Cech, N.; Faeth, S. Alkaloid variation among epichloid endophytes of sleepygrass (*Achnatherum robustum*) and consequences for resistance to insect herbivores. *J. Chem. Ecol.* **2015**, *41*, 93–104.
11. Miles, C.O.; Lane, G.A.; di Menna, M.E.; Garthwaite, I.; Piper, E.L.; Ball, O.J.P.; Latch, G.C.M.; Allen, J.M.; Hunt, M.B.; Bush, L.P.; *et al.* High levels of ergonovine and lysergic acid amide in toxic *Achnatherum inebrians* accompany infection by an *Acremonium*-like endophytic fungus. *J. Agric. Food Chem.* **1996**, *44*, 1285–1290.

12. De Groot, A.N.; van Dongen, P.W.; Vree, T.B.; Hekster, Y.A.; van Roosmalen, J. Ergot alkaloids. Current status and review of clinical pharmacology and therapeutic use compared with other oxytocics in obstetrics and gynaecology. *Drugs* **1998**, *56*, 523–535.
13. Burn, D. Neurology (2) Parkinson's disease: Treatment. *Pharm. J.* **2000**, *264*, 476–479.
14. Crosignani, P.G. Current treatment issues in female hyperprolactinaemia. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **2006**, *125*, 152–164.
15. Micale, V.; Incognito, T.; Ignoto, A.; Rampello, L.; Sparta, M.; Drago, F. Dopaminergic drugs may counteract behavioral and biochemical changes induced by models of brain injury. *Eur. Neuropsychopharmacol.* **2006**, *16*, 195–203.
16. Hofmann, A. *LSD: Mein Sorgenkind. Die Entdeckung einer "Wunderdroge"*; Deutscher Taschenbuch Verlag: München, Germany, 2006.
17. Robinson, S.L.; Panaccione, D.G. Diversification of ergot alkaloids in natural and modified fungi. *Toxins* **2015**, *7*, 201–218.
18. Wallwey, C.; Heddergott, C.; Xie, X.; Brakhage, A.A.; Li, S.-M. Genome mining reveals the presence of a conserved biosynthetic gene cluster for the biosynthesis of ergot alkaloid precursors in the fungal family Arthrodermataceae. *Microbiology* **2012**, *158*, 1634–1644.
19. Vazquez, M.J.; Roa, A.M.; Reyes, F.; Vega, A.; Rivera-Sagredo, A.; Thomas, D.R.; Diez, E.; Hueso-Rodriguez, J.A. A novel ergot alkaloid as a 5-HT(1A) inhibitor produced by *Dicyma* sp. *J. Med. Chem.* **2003**, *46*, 5117–5120.
20. Gerhards, N.; Neubauer, L.; Tudzynski, P.; Li, S.-M. Biosynthetic pathways of ergot alkaloids. *Toxins* **2014**, *6*, 3281–3295.
21. Panaccione, D.G., Origins and significance of ergot alkaloid diversity in fungi. *FEMS Microbiol. Lett.* **2005**, *251*, 9–17.
22. Panaccione, D.G.; Schardl, C.L.; Coyle, C.M. Pathways to diverse ergot alkaloid profiles in fungi. In *Recent Advances in Phytochemistry*; Romeo, J.T., Ed.; Elsevier: Amsterdam, The Netherlands, 2006; Volume 40, pp. 23–52.
23. Flieger, M.; Sedmera, P.; Vokoun, J.; Řičicová, A.; Řeháček, Z. Separation of four isomers of lysergic acid α -hydroxyethylamide by liquid chromatography and their spectroscopic identification. *J. Chromatogr. A* **1982**, *236*, 441–452.
24. Schardl, C.L.; Young, C.A.; Hesse, U.; Amyotte, S.G.; Andreeva, K.; Calie, P.J.; Fleetwood, D.J.; Haws, D.C.; Moore, N.; Oeser, B.; *et al.* Plant-symbiotic fungi as chemical engineers: Multi-genome analysis of the Clavicipitaceae reveals dynamics of alkaloid loci. *PLoS Genet.* **2013**, *9*, e1003323.
25. Coyle, C.M.; Cheng, J.Z.; O'Connor, S.E.; Panaccione, D.G. An old yellow enzyme gene controls the branch point between *Aspergillus fumigatus* and *Claviceps purpurea* ergot alkaloid pathways. *Appl. Environ. Microbiol.* **2010**, *76*, 3898–3903.
26. Ortel, I.; Keller, U. Combinatorial assembly of simple and complex d-lysergic acid alkaloid peptide classes in the ergot fungus *Claviceps purpurea*. *J. Biol. Chem.* **2009**, *284*, 6650–6660.
27. Havemann, J.; Vogel, D.; Loll, B.; Keller, U. Cyclolization of D-lysergic acid alkaloid peptides. *Chem. Biol.* **2014**, *21*, 146–155.
28. Haarmann, T.; Lorenz, N.; Tudzynski, P., Use of a nonhomologous end joining deficient strain ($\Delta ku70$) of the ergot fungus *Claviceps purpurea* for identification of a nonribosomal peptide synthetase gene involved in ergotamine biosynthesis. *Fungal Genet. Biol.* **2008**, *45*, 35–44.

29. Uhlig, S.; Petersen, D.; Rolén, E.; Egge-Jacobsen, W.; Vrålstad, T. Ergosedmine, a new peptide ergot alkaloid (ergopeptine) from the ergot fungus, *Claviceps purpurea* parasitizing *Calamagrostis arundinacea*. *Phytochem. Lett.* **2011**, *4*, 79–85.
30. Schardl, C.L.; Panaccione, D.G.; Tudzynski, P. Ergot alkaloids—Biology and molecular biology. *Alkaloids* **2006**, *63*, 45–86.
31. Matuschek, M.; Wallwey, C.; Wollinsky, B.; Xie, X.; Li, S.-M. *In vitro* conversion of chanoclavine-I aldehyde to the stereoisomers festuclavine and pyroclavine controlled by the second reduction step. *RSC Adv.* **2012**, *2*, 3662.
32. Unsöld, I.A.; Li, S.M. Overproduction, purification and characterization of FgaPT2, a dimethylallyltryptophan synthase from *Aspergillus fumigatus*. *Microbiol. SGM* **2005**, *151*, 1499–1505.
33. Coyle, C.M.; Panaccione, D.G. An ergot alkaloid biosynthesis gene and clustered hypothetical genes from *Aspergillus fumigatus*. *Appl. Environ. Microbiol.* **2005**, *71*, 3112–3118.
34. Chen, L.; Li, X.Z.; Li, C.J.; Swoboda, G.A.; Young, C.A.; Sugawara, K.; Leuchtmann, A.; Schardl, C.L. Two distinct *Epichloë* species symbiotic with *Achnatherum inebrians*, drunken horse grass. *Mycologia* **2015**, in press.
35. Blin, K.; Medema, M.H.; Kazempour, D.; Fischbach, M.A.; Breitling, R.; Takano, E.; Weber, T. antiSMASH 2.0—A versatile platform for genome mining of secondary metabolite producers. *Nucl. Acids Res.* **2013**, *41*, W204–W212.
36. Khaldi, N.; Seifuddin, F.T.; Turner, G.; Haft, D.; Nierman, W.C.; Wolfe, K.H.; Fedorova, N.D. SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* **2010**, *47*, 736–741.
37. Keller, N.P.; Turner, G.; Bennett, J.W. Fungal secondary metabolism—From biochemistry to genomics. *Nat. Rev. Microbiol.* **2005**, *3*, 937–947.
38. Schardl, C.L.; Young, C.A.; Moore, N.; Krom, N.; Dupont, P.-Y.; Pan, J.; Florea, S.; Webb, J.S.; Jaromczyk, J.; Jaromczyk, J.W.; *et al.* Genomes of plant-associated Clavicipitaceae. *Adv. Bot. Res.* **2014**, *70*, 291–327.
39. Gao, Q.; Jin, K.; Ying, S.-H.; Zhang, Y.; Xiao, G.; Shang, Y.; Duan, Z.; Hu, X.; Xie, X.-Q.; Zhou, G.; *et al.* Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. *PLoS Genet.* **2011**, *7*, e1001264.
40. Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J.-F.; Guindon, S.; Lefort, V.; Lescot, M.; *et al.* Phylogeny.fr: Robust phylogenetic analysis for the non-specialist. *Nucl. Acids Res.* **2008**, *36*, W465–W469.
41. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* **2004**, *32*, 1792–1797.
42. Anisimova, M.; Gascuel, O. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* **2006**, *55*, 539–552.
43. Schardl, C.L.; Young, C.A.; Pan, J.; Florea, S.; Takach, J.E.; Panaccione, D.G.; Farman, M.L.; Webb, J.S.; Jaromczyk, J.; Charlton, N.D.; *et al.* Currencies of mutualisms: Sources of alkaloid genes in vertically transmitted epichloae. *Toxins* **2013**, *5*, 1064–1088.
44. Pan, J.; Bhardwaj, M.; Faulkner, J.R.; Nagabhyru, P.; Charlton, N.D.; Higashi, R.M.; Miller, A.-F.; Young, C.A.; Grossman, R.B.; Schardl, C.L. Ether bridge formation in loline alkaloid biosynthesis. *Phytochemistry* **2014**, *98*, 60–68.

45. Pan, J.; Bhardwaj, M.; Nagabhyru, P.; Grossman, R.B.; Schardl, C.L. Enzymes from fungal and plant origin required for chemical diversification of insecticidal loline alkaloids in grass-*Epichloë* symbiota. *PLoS One* **2014**, *9*, e115590.
46. Schardl, C.L.; Farman, M.L.; Jaromczyk, J.W.; Young, C.A.; Webb, J.S.; Jaromczyk, J.; Moore, N.; Bec, S.; Calie, P.J.; Charlton, N.; *et al.* Genome sequences of 56 plant-associated Clavicipitaceae. **2015**, unpublished.
47. Charlton, N.D.; Craven, K.D.; Mittal, S.; Hopkins, A.A.; Young, C.A. *Epichloë canadensis*, a new interspecific epichloid hybrid symbiotic with Canada wildrye (*Elymus canadensis*). *Mycologia* **2012**, *104*, 1187–1199.
48. Schardl, C.L.; Young, C.A.; Faulkner, J.R.; Florea, S.; Pan, J. Chemotypic diversity of epichloae, fungal symbionts of grasses. *Fungal Ecol.* **2012**, *5*, 331–344.
49. Takach, J.E.; Mittal, S.; Swoboda, G.A.; Bright, S.K.; Trammell, M.A.; Hopkins, A.A.; Young, C.A. Genotypic and chemotypic diversity of *Neotyphodium* endophytes in tall fescue from Greece. *Appl. Environ. Microbiol.* **2012**, *78*, 5501–5510.
50. Charlton, N.D.; Craven, K.D.; Afkhami, M.E.; Hall, B.A.; Ghimire, S.R.; Young, C.A. Interspecific hybridization and bioactive alkaloid variation increases diversity in endophytic *Epichloë* species of *Bromus laevipes*. *FEMS Microbiol. Ecol.* **2014**, *90*, 276–289.
51. Takach, J.E.; Young, C.A. Alkaloid genotype diversity of tall fescue endophytes. *Crop. Sci.* **2014**, *54*, 667–678.
52. Young, C.A.; Charlton, N.D.; Takach, J.E.; Swoboda, G.A.; Trammell, M.A.; Huhman, D.V.; Hopkins, A.A. Characterization of *Epichloë coenophiala* within the US: Are all tall fescue endophytes created equal? *Front. Chem.* **2014**, *2*, 95.
53. Bacon, C.W. Procedure for isolating the endophyte from tall fescue and screening isolates for ergot alkaloids. *Appl. Env. Microbiol.* **1988**, *54*, 2615–2618.
54. Blankenship, J.D.; Spiering, M.J.; Wilkinson, H.H.; Fannin, F.F.; Bush, L.P.; Schardl, C.L. Production of loline alkaloids by the grass endophyte, *Neotyphodium uncinatum*, in defined media. *Phytochemistry* **2001**, *58*, 395–401.
55. Chujo, T.; Scott, B. Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a fungal endophyte–plant symbiosis. *Mol. Microbiol.* **2014**, *92*, 413–434.
56. Fleetwood, D.J.; Scott, B.; Lane, G.A.; Tanaka, A.; Johnson, R.D. A complex ergovaline gene cluster in epichloë endophytes of grasses. *Appl. Environ. Microbiol.* **2007**, *73*, 2571–2579.
57. Fleetwood, D.J.; Khan, A.K.; Johnson, R.D.; Young, C.A.; Mittal, S.; Wrenn, R.E.; Hesse, U.; Foster, S.J.; Schardl, C.L.; Scott, B. Abundant degenerate miniature inverted-repeat transposable elements in genomes of epichloid fungal endophytes of grasses. *Genome Biol. Evol.* **2011**, *3*, 1253–1264.
58. Panaccione, D.G.; Johnson, R.D.; Wang, J.H.; Young, C.A.; Damrongkool, P.; Scott, B.; Schardl, C.L. Elimination of ergovaline from a grass-*Neotyphodium* endophyte symbiosis by genetic modification of the endophyte. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12820–12825.
59. Ghimire, S.R.; Rudgers, J.A.; Charlton, N.D.; Young, C.; Craven, K.D. Prevalence of an intraspecific *Neotyphodium* hybrid in natural populations of stout wood reed (*Cinna arundinacea* L.) from eastern North America. *Mycologia* **2011**, *103*, 75–84.

60. Christensen, M.J.; Leuchtman, A.; Rowan, D.D.; Tapper, B.A. Taxonomy of *Acremonium* endophytes of tall fescue (*Festuca arundinacea*), meadow fescue (*F. pratensis*), and perennial rye-grass (*Lolium perenne*). *Mycol. Res.* **1993**, *97*, 1083–1092.
61. Leuchtman, A.; Schmidt, D.; Bush, L.P. Different levels of protective alkaloids in grasses with stroma-forming and seed-transmitted *Epichloë/Neotyphodium* endophytes. *J. Chem. Ecol.* **2000**, *26*, 1025–1036.
62. Schardl, C.L.; Craven, K.D.; Speakman, S.; Stromberg, A.; Lindstrom, A.; Yoshida, R. A novel test for host-symbiont codivergence indicates ancient origin of fungal endophytes in grasses. *Syst. Biol.* **2008**, *57*, 483–498.
63. Petroski, R.; Powell, R.G.; Clay, K. Alkaloids of *Stipa robusta* (sleepygrass) infected with an *Acremonium* endophyte. *Nat. Toxins* **1992**, *1*, 84–88.
64. Moon, C.D.; Guillaumin, J.-J.; Ravel, C.; Li, C.; Craven, K.D.; Schardl, C.L. New *Neotyphodium* endophyte species from the grass tribes Stipeae and Meliceae. *Mycologia* **2007**, *99*, 895–905.
65. Farman, M.L. Telomeres in the rice blast fungus *Magnaporthe oryzae*: The world of the end as we know it. *FEMS Microbiol. Lett.* **2007**, *273*, 125–132.
66. Starnes, J.H.; Thornbury, D.W.; Novikova, O.S.; Rehmeier, C.J.; Farman, M.L. Telomere-targeted retrotransposons in the rice blast fungus *Magnaporthe oryzae*: Agents of telomere instability. *Genetics* **2012**, *191*, 389–406.
67. Wu, C.; Kim, Y.-S.; Smith, K.M.; Li, W.; Hood, H.M.; Staben, C.; Selker, E.U.; Sachs, M.S.; Farman, M.L. Characterization of chromosome ends in the filamentous fungus *Neurospora crassa*. *Genetics* **2009**, *181*, 1129–1145.
68. Stajich, J.E.; Wilke, S.K.; Ahren, D.; Au, C.H.; Birren, B.W.; Borodovsky, M.; Burns, C.; Canback, B.; Casselton, L.A.; Cheng, C.K.; *et al.* Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11889–11894.