Mechanisms of reproductive isolation in sexual Epichloë endophytes

Author(s): Schirrmann, Melanie K.
Publication Date: 2015
Permanent Link: https://doi.org/10.3929/ethz-a-010531497

Rights / License: In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Mechanisms of reproductive isolation in sexual *Epichloë* endophytes

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

**Melanie Katharina Schirrmann**

Dipl.-Biol., Karlsruhe Institute of Technology
born on 08.08.1986
citizen of Germany

accepted on the recommendation of

Prof. Dr. Adrian Leuchtmann, examiner
Prof. Dr. Ottmar Holdenrieder, co-examiner
Prof. Dr. Florian Schiestl, co-examiner

2015
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
</tr>
<tr>
<td>Zusammenfassung</td>
</tr>
<tr>
<td>General Introduction</td>
</tr>
<tr>
<td>Chapter 1</td>
</tr>
<tr>
<td>Chapter 2</td>
</tr>
<tr>
<td>Chapter 3</td>
</tr>
<tr>
<td>General Discussion</td>
</tr>
<tr>
<td>Acknowledgements</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
</tr>
</tbody>
</table>
SUMMARY

Understanding the mechanisms underlying reproductive isolation in eukaryotes is of fundamental interest in speciation. The mechanisms driving reproductive isolation can be divided into pre- and postzygotic barriers, depending on whether they occur before or after fertilization. Sexual species of *Epichloë* endophytes (Ascomycota, Clavicipitaceae) that infect different pooid grasses (Poaceae) are an interesting study system for investigating the complex interplay of potential pre- and postzygotic barriers. *Epichloë* can easily be cultured in the laboratory on agar plates, as well as *in planta* in the greenhouse, which allows multiple experimental techniques, for example inoculation tests and mating trials, to be used in testing for reproductive isolation. Host-associated taxa within the *Epichloë typhina* complex are particularly useful for studying reproductive isolation, because they are interfertile and produce viable hybrids in artificial experiments of cross-fertilization.

In the first chapter, the genetic population structure of three host-associated *Epichloë* taxa, i.e. *E. typhina* subsp. *typhina* infecting *Dactylis glomerata*, *E. typhina* subsp. *clarkii* infecting *Holcus lanatus* and *E. typhina* subsp. *poae* infecting *Poa nemoralis*, from a site with sympatric populations in Switzerland was assessed with a newly developed set of 15 microsatellites. The results showed high genetic differentiation between the taxa, indicating that reproductive barriers keep these populations genetically distinct. Furthermore, no grasses infected with hybrid endophytes could be detected, although previously hybrid ascospores were found to be present at the sampling site.

In the second chapter, reciprocal infections were conducted with four host-associated taxa of the *E. typhina* complex, i.e. *E. typhina* subsp. *typhina* (*D. glomerata*), *E. typhina* subsp. *clarkii* (*H. lanatus*), *E. typhina* subsp. *poae* (*P. nemoralis*) and *E. typhina* subsp. *typhina* (*Poa trivialis*), to better understand the mechanisms underlying reproductive isolation. By comparing the infection frequencies of parental strains and hybrids in each host plant, an inferior performance of hybrid strains was detected, suggesting postzygotic reproductive barriers. Furthermore, parental strains showed strict host-specificity, which may explain the maladaptation of hybrids to the parental host grasses. Adaptation to one host could therefore be sufficient to restrict gene flow between different host-associated *Epichloë* taxa.

In the third chapter, conidia competition between *E. typhina* subsp. *typhina* (*D. glomerata*) and *E. typhina* subsp. *clarkii* (*H. lanatus*) was tested *in planta* to study potential prezygotic barriers during sexual reproduction. Among successful matings mainly within-taxon progeny were found, whereas hybrid progeny were only found in small proportions. This may indicate that prezygotic barriers restrict the formation of hybrids between host-associated taxa because of genetic conflicts. Furthermore, in order to identify genes that are involved in host adaptation driving the divergence between the two taxa, whole genomes of an *E. typhina* subsp. *typhina* (*D. glomerata*) and an *E. typhina* subsp. *clarkii* (*H. lanatus*)
population were re-sequenced. The analysis indicated high genetic differentiation between the taxa, which is congruent with the study based on only a few microsatellites. By identifying genes that show elevated genetic differentiation ($F_{ST}$ and $D_{xy}$) relative to the genome-wide average, several putative secreted proteins were found to be under positive selection ($dN/dS$ ratio and McDonald-Kreitman test). Candidate genes were mainly encoding for proteins, which are known to be important during the pathogenic stage of the *Epichloë* life cycle. This suggests that endophytes closely interact with the host, and therefore these genes may be involved in host adaptation and eventually in reproductive isolation.

This PhD thesis provides insights into the different mechanisms underlying pre- and postzygotic barriers, which keep natural *Epichloë* populations genetically distinct. In this context, host adaptation seems to play an important role, because it affects mating success as well as successful infections of different host grasses. Future experiments should aim to shed more light on the candidate genes that are involved in the reproductive isolation of endophytes.
ZUSAMMENFASSUNG


Im dritten Kapitel wurde die Konkurrenzfähigkeit von Conidien zwischen *E. typhina* subsp. *typhina* (*D. glomerata*) und *E. typhina* subsp. *clarkii* (*H. lanatus*) *in planta* getestet, um potentielle präzygotische Barrieren während der sexuellen Vermehrung zu testen. Unter
Zusammenfassung


GENERAL INTRODUCTION

Speciation
Speciation is one of the most fundamental processes in biology that leads to biological diversity. To define species, alternative concepts are used that utilise different properties of the organisms, e.g. the morphology, ecology, evolution or phylogenetics (De Queiroz, 2007). One of the commonly used criteria to define biological species is the lack of interbreeding between species. Mayr (1969) proposed that ‘species are groups of interbreeding natural populations that are reproducively isolated from other such groups’. However, the mechanisms of speciation are variable in different organisms and still incompletely known.

Extrinsic geographic barriers appear obvious impediments to gene flow, and therefore it has long been believed that species originate through allopatric divergence (Mayr, 1963). Fungi seemed to be exceptions, because like other microorganisms they were considered to have global geographic ranges (Finlay, 2002). However, the inferred geographic range of a fungal species depends upon the method to define species (Taylor et al., 2006). While some fungal species defined by morphology show global geographic ranges, fungal species defined by phylogenetic species concepts typically have very narrow ranges, suggesting that morphospecies may encompass several to many cryptic species (Cai et al., 2011). Therefore, among complexes of sibling species speciation may be consistent with allopatric divergence, because the cryptic species occupy non-overlapping areas separated by geographic barriers (Dettman et al., 2003; O’Donnell et al., 2004; Taylor et al., 2006).

The possibility of sympatric speciation in sexual populations that live in the same geographic area has also long been dismissed in fungi, as a very low level of gene flow, such as one migrant per generation, is sufficient to prevent differentiation (Slatkin, 1987). However, speciation in sympatry may indeed occur following the formation of isolating barriers, and in this context, adaptation to different hosts and habitats has been suggested to be a strong driver of this process (Fournier and Giraud, 2008; McCoy, 2003). Fungi are unable to disperse between development on a host or habitat, and therefore mutations involved in adaptation to a new habitat can pleiotropically affect both the fitness on the habitat and the ability to mate in this habitat (Giraud et al., 2008). Adaptation to a host can thus be sufficient to restrict gene flow in sympaty, if mating occurs within the specialized host (Giraud, 2006). For example, speciation of Mycosphaerella graminicola has been associated with strong host specialization, for which frequent sexual recombination has been instrumental by driving rapid adaptive evolution (Stukenbrock et al., 2011). Another example of sympatric divergence mediated by the host is provided by Ascochyta pathogens (Peever, 2007). The host specificity shown by experimental inoculations, coupled with the inability of hybrid progeny to colonize and reproduce on a host, may constitute strong reproductive barriers in these fungi. However, compelling evidence for sympatric divergence
is extremely difficult to provide, because excluding a past period of allopatry is almost always impossible (Coyne and Orr, 2004).

**Reproductive Isolation**

Traits that mediate reductions in interspecific gene flow, known as reproductive isolating barriers, are usually divided into prezygotic barriers that occur before fertilization, and extrinsic and intrinsic postzygotic barriers that occur after fertilization (Coyne and Orr, 2004). In comparison to extrinsic forms of postzygotic isolation, which depend on the interaction with either the environment or with other individuals, intrinsic forms are due to genetic incompatibilities that are independent of the environment (Kohn, 2005).

Prezygotic barriers may include: (1) specialization of a biotic vector that can prevent contact between two populations, yielding ecological isolation (Bultman and Leuchtmann, 2003), (2) specialization to a habitat or host, allowing for ecological premating isolation (Giraud, 2006), (3) differences in the time of reproduction, or (4) assortative mating due to mate recognition, if individuals or gametes are able to discriminate between conspecific and heterospecific spores (Büker et al., 2013; Le Gac and Giraud, 2008).

Intrinsic postzygotic barriers are associated with hybrid inviability and sterility, and are expected to arise as a result of the divergence of incipient species (Orr and Turelli, 2001). Hybrid inviability or reduced fitness can occur due to genetic incompatibilities of diverging strains. Mutations that become fixed in divergent lineages can display negative epistatic interactions when brought together in the same individual, a phenomenon known as Dobzhansky-Muller (DM) incompatibilities (Dobzhansky, 1937). A characteristic feature of this mechanism is that severe incompatibilities are expressed only in hybrids and not within taxa. For example, heterospecific crosses among *Microbotryum* species produce fewer viable mycelia *in vitro* than conspecific ones (Le Gac et al., 2007), and crosses among *Neurospora* species lead to few or abnormal perithecia or few viable ascospores (Dettman et al., 2003).

In contrast, extrinsic postzygotic barriers are linked to ecological factors, whereby hybrids can be viable and fertile under *in vitro* conditions, but unfit in a natural environment. This can be the case if hybrids display intermediate traits between parental phenotypes and are, therefore, poor competitors in either parental environment (Coyne and Orr, 2004; Kohn, 2005). In *M. violaceum*, hybrids between close species were inoculated into parental host species, *Silene latifolia* and *S. dioica*. In one host species, hybrids performed as well as the parental species specialized for this host, whereas when inoculated in the reciprocal host, hybrids performed worse than the parental species, showing that the same hybrids had a lower viability (Le Gac et al., 2007).
Hybridization
Many fungal species can potentially hybridize, as many do not exhibit complete interspecific sterility. Hybridization has been described as the process of sexual recombination between genetically distinct individuals with different potential outcomes (Harrison, 2012). If hybrids are fertile, hybridization can lead to speciation or introgressive hybridization, potent evolutionary forces that can promote the emergence of new species and allow adaptation to new ecological niches (Rieseberg and Carney, 1998; Schardl and Craven, 2003). Introgressive hybridization can be divided in to allopolyploid hybridization when hybrids have a chromosomal number that sums that of the parental species, and allodiploid or homoploid hybridization when the ploidy of the hybrids is identical to that of the parents (Butlin et al., 2012).

By combining traits, allopolyploid hybrids can have selective advantages over their parental species, because alloplody can simultaneously promote reproductive isolation due to triploidy in backcrosses and a new ecological niche (Giraud et al., 2008). Indeed, allopolyploids of plants exhibit a wider range of phenotypes allowing them able to exploit vacant ecological niches (Abbott et al., 2013; Rieseberg and Willis, 2007). An example of fungal allopolyploids can be found in Cryptococcus neoformans, a human pathogen causing meningoencephalitis. Diploid hybrids have been found to show enhanced fitness compared to the less virulent parental haploid forms, which is hypothesized to have contributed to their worldwide distribution (Lin et al., 2007).

In contrast to allopolyploids, which are reproductively isolated from their parents, homoploid hybrids are in competition with their parental species and with backcrossed individuals (Kohn, 2005). Stable allodiploid or homoploid hybrids are therefore much more unlikely to persist in nature than polyploid ones (Giraud et al., 2008). However, homoploid speciation is observed in the rust Melampsora x columbiana that emerged from hybridization between Melampsora medusa and M. occidenta that are parasites of Populus deltoids and P. trichocarpa, respectively (Newcombe et al., 2000). The hybrid rust could infect a hybrid poplar, which was resistant to the two parental rust species and therefore widely grown in California. By occupying this novel ecological niche, its maintenance and persistence as a new species was facilitated.

Mutualistic Interactions
Mutualistic interactions, the way two organisms of different species exist in a reciprocally positive relationship, play a key role in ecology (Boucher et al., 1982; Bronstein, 2009), and are thought to have driven the evolution of much of the biological diversity, e.g. flower forms and coevolution between groups of different species (Leigh, 2010). However, mutualism has received less attention than other symbiotic interactions such as predation and parasitism, where one species benefits at the expense of the other (Bronstein, 1994).

The most familiar mutualisms are between plants and pollinators (Thompson and
Pellmyr, 1992), plants and nitrogen fixing bacteria (Denison and Kiers, 2004), plants and mycorrhizae (Smith and Read, 1997), and endophytic fungi that form close associations with their host plants (Clay, 1990; Gundel et al., 2013; Leuchtmann, 2003). In contrast to plant-pollinator mutualisms, in which the behavior of a mobile species, such as the pollination of plants or the dispersal of seeds, is rewarded with food from a sedentary species, many associations between species involve only the acquisition of limiting nutrients (Bronstein, 2009). For example, rhizobia bacteria fix nitrogen in nodules of roots for leguminous plants (Fabaceae) and receive in return energy-containing carbohydrates and various mineral nutrients (Kiers et al., 2003). In mycorrhizal associations between plant roots and fungi, the plant provides carbohydrates to the fungus and receives primarily phosphate but also nitrogenous compounds in exchange (Remy et al., 1994). Another form of mutualism exists between plants and endophytic fungi, in which fungi are thought to benefit their hosts through direct defense against natural enemies of the plant (Clay, 1988; Faeth, 2002). By extensive colonization of the plant tissue the endophytes may outcompete and prevent pathogenic organisms from taking hold. Some endophytes also produce chemicals, which can inhibit the growth of competitors and protect the plants from herbivores (Caroll, 1992). Associations between fungal endophytes and plants have been the focus of much research, particularly among the grass endophytes (Gundel et al., 2012; Rudgers et al., 2012).

**Epichloë Endophytes**

*Epichloë* species (Ascomycota, Clavicipitaceae) are widespread endophytes of pooid grasses (Poaceae) and can be characterized by two different life cycles (Fig. 1) (Leuchtmann, 1992; Schardl et al., 2009). They are generally viewed as plant mutualists, because of the action of mycotoxins, e.g. alkaloids, in infected grasses, which protect the host plant from herbivores (Clay, 2009; Schardl et al., 2013; Tanaka et al., 2005). Furthermore, they colonize only the apoplast between plant cells without entering them, and can enhance plant growth and resistance to stress such as drought (Clay, 1988; Schardl and Phillips, 1997). In return, the endophytes profit from these interactions through receiving photosynthetically derived carbohydrates from the host plant.

*Epichloë* species form external reproductive structures on inflorescences, called stromata, which partially or completely inhibit host flowering and seed production, known as ‘choke disease’ (Clay and Schardl, 2002; Western and Cavett, 1959). They are heterothallic and possess a bipolar self-incompatibility mating system (White and Bultman, 1987), which is based on the mating type gene *MAT-1* with two idiomorphs *MAT-1-1* and *MAT-1-2* (Turgeon and Yoder, 2000; Yokoyama et al., 2005). Successful mating therefore requires transfer of spermata (male gametes) between stromata of opposite mating types, which is usually accomplished by specialized flies of the genus *Botanophila* in a process similar to pollination (Bultman et al., 1998). *Epichloë* fungi benefit from these flies, which are attracted by volatile odors, as reliable vectors of spermata for cross-fertilization, whereas fly larvae feed and
develop on fertilized stromata by consuming fungal tissues (Kohlmeyer and Kohlmeyer, 1974; Schardl et al., 2004; Steinebrunner et al., 2008).

On the stroma, haploid spermatia and receptive hyphae are simultaneously formed (White and Bultman, 1987). After the transfer of spermatia of the opposite mating type, spermatia germinate and start proliferating on the surface (Chung and Schardl, 1997). Within a peripheral layer of cavities, which later develop into perithecia, fusion (plasmogamy) between male and female structures (gametangia) takes place. If partners are sexually compatible, young asci develop within perithecia and karyogamy and meiosis occurs (Bultman et al., 2011). Asci contain eight haploid ascospores that are ejected, wind-dispersed and mediate horizontal transmission through infections of host florets and then seed (Bultman and Leuchtmann, 2009). Furthermore, ascospores may also serve as spermatia and fertilize stromata in the absence of Botanophila flies (Alderman and Rao, 2008; Górzyńska et al., 2010).

**Figure 1**: Asexual and sexual life cycle of *Epichloë* endophytes (Leuchtmann and Schardl, 1998). During the asymptomatic asexual life cycle the endophyte grows intercellularly in vegetative plants (Sampson, 1933). At the onset of inflorescence development, it infects developing seeds and later upon germination the embryo, and is thus transmitted vertically through successive host generation. During the sexual life cycle, the fungus grows asymptotically in leaf sheaths and meristems, but then forms external stromata with spermatia on flowering tillers. Spermatia function as male gametes and are transferred by *Botanophila* flies to stromata of opposite mating type. After successful mating, perithecia containing asci develop and filamentous ascospores are ejected. Germinating ascospores may cause infection of host florets and then seeds.
Diversity and Mating Compatibility of *Epichloë*

Previously, *Epichloë* species have been classified under two different genera depending on differences in their mode of reproduction. Asexual species that are only transmitted maternally by growth into the ovules and seeds of infected plants were referred to *Neotyphodium* (Glenn et al., 1996), whereas sexual species that form cylindrical stromata with asci and ascospores on the surface of culms were placed in the genus *Epichloë*. Recent changes of nomenclatural rules (McNeill et al., 2012) demand that both, anamorphic *Neotyphodium* and sexual *Epichloë* species, are included in one genus *Epichloë*. Following a realignment of the genus by Leuchtmann et al. (2014), 10 teleomorph-typified species and 24 anamorph-typified species (including three subspecies and six varieties) are currently accepted in the genus *Epichloë*. Sexual species have been described based on morphological characters, which include size and disarticulation patterns of ascospores, size and density of perithecia, and host species preferences (Leuchtmann et al., 2014; Leuchtmann and Schardl, 1998). Using such criteria, ten sexual morphospecies of *Epichloë* were previously characterized. In Table 1, the morphospecies *E. typhina* and *E. clarkii*, which were used in this study, are listed with corresponding hosts and mating population (MP). Besides their morphological characters, most species can be distinguished by apparent sexual incompatibility among them established by experimental mating tests (Leuchtmann et al., 1994; Schardl et al., 1997). Based on such reproductive incompatibility, Schardl et al. (1997) identified nine mating populations. Furthermore, an additional sexual *Epichloë* species, *E. liyangensis* has recently been described without assigning it to a mating population (Kang et al., 2011).

**Table 1**: Mating populations (MP) of *Epichloë* on host grasses as identified by mating tests adapted from Schardl et al. (1997).

<table>
<thead>
<tr>
<th>Mating population</th>
<th>Morphospecies</th>
<th>Host species</th>
<th>Host tribe</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP-I</td>
<td><em>E. typhina</em></td>
<td><em>Athoxanthum odoratum</em></td>
<td>Aveneae</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Phleum pratense</em></td>
<td>Aveneae</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Dactylis glomerata</em></td>
<td>Poeae</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Lolium perenne</em></td>
<td>Poeae</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Poa pratensis</em></td>
<td>Poeae</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>P. trivialis</em></td>
<td>Poeae</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>P. nemoralis</em></td>
<td>Poeae</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Brachypodium pinnatum</em></td>
<td>Brachypodieae</td>
</tr>
<tr>
<td>MP-I</td>
<td><em>E. clarkii</em></td>
<td><em>Holcus lanatus</em></td>
<td>Poeae</td>
</tr>
</tbody>
</table>

* Host species that were used in this thesis
Most mating populations show rather narrow host ranges restricted to grass species of one genus or closely related host genera with the exception of MP-1 (Tab. 1), including the two interfertile morphospecies *E. typhina* and *E. clarkii* (now treated as *E. typhina* subsp. *clarkii*). This mating population has a very broad host range, including isolates from nine different host species in seven genera from three different tribes of Pooideae (Leuchtmann and Schardl, 1998; Schardl et al., 1997). However, mating populations defined by experimental mating tests may be more comprehensive than natural mating populations, where a number of mechanisms can effectively prevent gene flow and thus genetically isolate populations associated with different hosts (Leuchtmann and Schardl, 1998). A case of sexual isolation within MP-I may be identified in the morphospecies *E. typhina* subsp. *clarkii* on *Holcus lanatus*. Although hybrid progeny can easily be obtained by artificial cross-fertilizations between *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *typhina* (*D. glomerata*), morphologically distinguishable hybrid ascospores have rarely been observed (Bultman et al., 2011; Schardl et al., 1997). In fact, interfertility can be retained long after speciation without the development of incompatibility, because host specialization can pleiotropically generate reproductive isolation (Giraud et al., 2008; Le Gac and Giraud, 2008).

**Significance of endophytes and applied importance**

The mechanisms of speciation leading to biodiversity are still incompletely known, especially in Fungi, one of the most species rich kingdoms. Sexual species of *Epichloë* endophytes offer an ideal model system to study eukaryotic mechanisms of speciation. They include species complexes with several host-associated taxa and both pre- and postzygotic reproductive isolation may potentially be involved in the maintenance of species barriers.

Understanding the evolutionary mechanisms leading to species diversity in *Epichloë* is not only relevant to better understand speciation in Fungi, but also for applied aspects of endophyte research. *Epichloë* species are of importance in agriculture and turf grass industry, where they can cause severe economic losses to seed production of grasses and no efficient control options are currently available (Leyronas and Raynal, 2008). However, these endophytes can also enhance their hosts’ survival through protection from abiotic and biotic stresses, and exclusive grass-endophyte associations have been developed and commercialized (Johnson et al., 2013). The development of efficient control strategies therefore relies on investigating the species diversity of *Epichloë* in natural populations, as well as elucidating mechanisms of compatibility between hosts and endophytes.
Thesis Outline
The aim of this research is to investigate potential isolation mechanisms among members of the *E. typhina* species complex, including the two interfertile taxa *E. typhina* subsp. *typhina* infecting *D. glomerata* and *E. typhina* subsp. *clarkii* infecting *H. lanatus* and other host-associated cryptic species so far classified in *E. typhina*. Previous research has shown a lack of complete specialization of the fly allegedly responsible for transferring gametes among different stromata of opposite mating type (Bultman et al., 2011). Therefore, it is likely that other pre- and postzygotic mechanisms are acting together and contribute to various degrees to the reproductive isolation observed between sympatrically growing *Epichloë* endophytes. The main research questions of my thesis are as follows:

1) **Are sympatric *Epichloë* species genotypically differentiated at a natural field site?**
   *Epichloë typhina* subsp. *typhina* (*D. glomerata*) and *E. typhina* subsp. *clarkii* have been shown to be interfertile experimentally, although they seem to be reproductively isolated in natural populations. In an extended field study of two sympatric populations at the Aubonne Arboretum, Vaud, in Switzerland, a small fraction (~9%) of the fertilized stromata were found to be hybrid ascospores (Bultman et al., 2011), suggesting that matings between the two host-associated taxa can potentially occur under natural field conditions. However, capability of these spores to infect hosts seemed to be virtually null, as grasses infected with hybrid *Epichloë* were never found. Starting from a representative sampling of stromata from the same location at the Aubonne Arboretum, I aimed to confirm the genetic differentiation of the parental species by means of molecular markers. This was achieved by developing a new set of 15 microsatellites markers using whole genome DNA sequences of *E. typhina* subsp. *poae* infecting *P. nemoralis*.

2) **What is the role of host specificity in the reproductive isolation of *Epichloë***?
   Potential mechanisms of reproductive isolation among natural populations of *Epichloë* may be associated with host specificity and/or a reduced fitness of hybrids (Giraud et al., 2008). To analyze the compatibility of different *Epichloë* strains of the *E. typhina* complex and their F1 progeny with host grasses, I mimicked the infection of new hosts in the sexual life cycle by the inoculation of wounded seedlings (Latch and Christensen, 1985). Furthermore, because the sexual reproduction of *Epichloë* species involves stroma formation on flowering tillers of the host grass, their reproductive success is directly linked to the fitness of the host. In particular, tiller numbers may be increased for infected plants as a result of the plant-fungus mutualistic interaction (Groppe et al., 1999). Using greenhouse experiments, I therefore assessed the number of stromata and symptomless tillers of host plants, which were positively infected after inoculation with different *Epichloë* strains.
3) Are spermatia from different host-specific taxa recognized or competing among each other on the stroma? Can we identify strongly diverged candidate genes that are involved in the reproductive isolation between natural Epichloë populations?

Using a set of 15 microsatellites, I found high genetic differentiation between two sympatric Epichloë populations, i.e. *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *typhina* (*D. glomerata*), possibly maintained by very restricted gene flow. These results suggest that, although prezygotic barriers might be incomplete because of apparent interfertility, hindrance of hybrid formation is supported by other mechanisms of reproductive isolation. By testing for assortative mating on stromata, I aimed to get a more complete picture of prezygotic barriers among Epichloë. Furthermore, to find genomic regions with functions related to the maintenance of reproductive isolation between host-associated taxa, I re-sequenced whole genomes of a pooled subset of individuals from those two sympatric populations. I examined patterns of genetic diversity between and within populations to identify strongly diverged genes, which may be involved in host adaptation and eventually reproductive isolation (Stukenbrock, 2013).
References


121–133. doi:10.1007/s13225-011-0127-8


**General Introduction**


CHAPTER 1

GENETIC EVIDENCE FOR REPRODUCTIVE ISOLATION AMONG SYMPATRIC *EPICHLOË* ENDOPHYTES AS INFERRED FROM NEWLY DEVELOPED MICROSATellite MARKERS

Published as
Schirrmann, M.K.¹, Zoller, S.², Fior, S.¹, & Leuchtmann, A.¹ (2014). Genetic evidence for reproductive isolation among sympatric *Epichloë* endophytes as inferred from newly developed microsatellite markers. Microbial Ecology, 70(1), 51-60.

¹ETH Zurich, Institute of Integrative Biology, Plant Ecological Genetics, Zürich, Switzerland
²ETH Zurich, Genetic Diversity Centre (GDC), Zürich, Switzerland
Abstract
Reproductive isolation is central to the maintenance of species, and especially in sympatry effective barriers to prevent interspecific crosses are expected. Host-specificity is thought to constitute an effective mechanism for the formation of barriers in different genera of Fungi, but evidence for endophytes is so far lacking. Sexual Epichloë species (Ascomycota, Clavicipitaceae) represent an ideal study system to investigate the mechanisms underlying speciation as mediated by host-specificity, because they include species complexes with several host-specific taxa. Here, we studied genetic differentiation of three host-specific Epichloë species using microsatellite markers that were newly in silico identified on the genome of E. typhina subsp. poae. Among these, 15 were experimentally tested and applied to study an extensive sampling of isolates representing E. typhina subsp. typhina\(^1\) infecting Dactylis glomerata and E. typhina subsp. clarkii\(^2\) infecting Holcus lanatus from a site with sympatric populations in Switzerland, as well as a reduced sampling of E. typhina subsp. poae\(^3\) infecting Poa nemoralis to create a three-taxon dataset. Both principal coordinate analysis and Bayesian clustering algorithm showed three genetically distinct groups representing the three host-specific species. High pairwise \(F_{ST}\) values among the three species, as well as sequencing data of the \(tefA\) gene revealing diagnostic single nucleotide polymorphisms (SNPs) further support the hypothesis of genetic discontinuities among the taxa. These results provide genotypic evidence of the maintenance of reproductive isolation of the species in a context of sympatry. In silico testing of 885 discovered microsatellites on the genome of E. festucae extend their applicability to a wider taxonomic range of Epichloë.

Introduction
Speciation in fungi can occur through numerous mechanisms resulting in the formation of isolating barriers among divergent lineages. Compared to speciation in plants and animals, where geographic separation has long been hypothesised to be a strong driver in the formation of current biodiversity, the mechanisms underlying speciation in fungi have been less amenable to a general consensus (Giraud et al., 2008; Kohn, 2005; Restrepo et al., 2014). Indeed, as eukaryotic microorganisms typically have global geographic ranges, geographic isolation has been questioned as a promoter of divergence (Finlay, 2002). Similarly, the process of sympatric speciation has often been regarded as unlikely, because minimal gene flow is sufficient to prevent differentiation (Slatkin, 1987). However, recent advances facilitated by the use of genome-wide data in well-established model systems (Ellison et al., 2011; Liti et al., 2009) are greatly contributing to the understanding of speciation in fungi, and evidence is emerging for both types of speciation processes. Several instances have been described where divergence originated in a context of allopatry, but

---

1 Name in the publication: E. typhina
2 Name in the publication: E. clarkii
3 Name in the publication: E. poae
speciation remained cryptic because of undistinguishable phenotypes within species complexes (Dettman et al., 2003; Koufopanou et al., 1997; O'Donnell et al., 2004). On the other hand, speciation in sympathy may indeed occur following the formation of isolating barriers, and in this context, adaptation to different hosts has been proven to be a strong driver of this process (Fournier and Giraud, 2008; Peever, 2007; Stukenbrock et al., 2011). Because of the difficulty to separate between development and mating, mutations providing adaptation to a new habitat can pleiotropically affect both the fitness and the ability to mate in this habitat, which can thus be sufficient to restrict gene flow in sympathy (Giraud, 2006; Giraud et al., 2008). However, whether host-specificity is indicative of different species in a context of sympathy depends on the efficiency of the isolating barriers that prevent admixtures of genotypes, which should be proven with population genetic data.

Sexual species of Epichloë (Ascomycota, Clavicipitaceae) constitute an ideal study system to investigate the genetics behind reproductive isolation, as they form host adapted species complexes with pooid grasses (Poaceae) that are maintained in sympatric populations under a strong depletion of natural hybrids (Bultman and Leuchtmann, 2009; Leuchtmann, 1992; Schardl et al., 2009). The endophytes grow in plant apoplasts primarily in shoot apical meristems and leaf sheaths (Scott, 2001) with little or no negative effects on the vegetative tissues of their hosts. However, during their sexual life cycle they may form external reproductive structures, i.e. stromata, that partially or completely inhibit host flowering and seed production. The effect of such interaction on the plant is known as ‘choke disease’ (Western and Cavett, 1959). Sexual species of Epichloë are heterothallic and possess a bipolar self-incompatibility mating system (White and Bultman, 1987). Successful mating, therefore, requires transfer of spermia (male gametes) between stromata of opposite mating types, usually mediated by specialized flies of the genus Botanophila (Anthomyiidae) in a process similar to pollination (Bultman et al., 1998).

Most sexual Epichloë species are circumscribed based on their reproductive incompatibilities with other species following a biological species concept (Craven et al., 2001). Based on intersterility tests between species, nine different mating populations corresponding usually to a single taxonomic Epichloë species were identified (Schardl et al., 1997). Each species typically shows a rather narrow host range restricted to grass species of one genus or closely related host genera, with the exception of E. typhina that forms a complex with several host-specific taxa (Craven et al., 2001). Some of the species of the E. typhina complex have been demonstrated to be able to hybridize in artificial experiments of cross-fertilization, but appear to remain reproductively isolated in natural populations (Leuchtmann and Schardl, 1998; White, 1993).

A particular case where this is evident to a large extent includes E. typhina subsp. clarkii infecting Holcus lanatus and E. typhina subsp. typhina infecting Dactylis glomerata. These host-specific species are morphologically well defined (Leuchtmann and Clay, 1997) and occur scattered at multiple locations of their distribution range throughout temperate
Chapter 1

Europe and western Asia. In a recent field study, Bultman et al. (2011) described sympatric populations of the two species infecting hosts growing intermixed at the site of Aubonne in Canton Vaud, Switzerland. Here, hybrid ascospores were reported in a small fraction (~9%) of the fertilized stromata, however, capability of these spores to infect hosts seemed to be virtually null, as grasses infected with hybrid Epichloë were never found. This scenario suggests that E. typhina subsp. clarkii and E. typhina subsp. typhina have developed strong isolating barriers that maintain the species boundaries. However, the mechanisms that underlie the maintenance of the species barriers are largely unknown.

Addressing questions of speciation in sexual Epichloë species involves careful examination of the structure and hybridization of natural populations for which efficient molecular tools, such as microsatellites, are necessary. Previously, microsatellite markers have been developed to study the ecology and diversity of E. bromicola and Neotyphodium spp. (Groppe et al., 1995; van Zijl de Jong et al., 2003), and for identifying different sexual and asexual Epichloë endophytes in planta (Groppe and Boller, 1997; Moon et al., 1999; Simpson et al., 2013). Neotyphodium species, today classified in the genus Epichloë (Leuchtmann et al., 2014), are thought to be derived from ancestral Epichloë species, often as a result of ancient hybridization events (Moon et al., 2004; Schardl and Craven, 2003). However, the available microsatellites yielded unsatisfactory results when tested on our samples, probably because of partial primer mismatches among Epichloë species.

In this study, we aimed to investigate the genetic signature of species delimitation in the case of the two sympatric populations of E. typhina subsp. typhina and E. typhina subsp. clarkii at the Aubonne site in order to gain insight in the genetic differentiation of the fungal species and provide genetic evidence of their effective reproductive isolation. To achieve this goal, we developed a new set of microsatellite markers starting from whole genome sequences of E. typhina subsp. poae (Schardl et al., 2013) to detect markers that are potentially applicable to a wide range of species in the genus. To our knowledge, this study is the first that addresses the genetics behind reproductive isolation among sympatric Epichloë species.

Material and methods

Computational Workflow for Microsatellite Design

A workflow was developed to find, in silico test and filter microsatellites in the genomic DNA sequences of E. typhina subsp. poae strain ES819 downloaded from Genome Projects at University of Kentucky (http://www.endophyte.uky.edu). The workflow employed in-house Perl scripts calling the Tandem Repeats Finder v. 4.0.4 (Benson, 1999) to find repeats, Primer3 v. 2.3.0 (Rozen and Skaletsky, 2000) to construct primers in the flanking regions, and BLAST+ v. 2.2.26 (Camacho et al., 2009) for in silico primer testing on the genome.

Primer pairs were designed in order to obtain PCR products between 100 and 500 bp long. The workflow allowed for filtering of the microsatellites into PCR fragment size classes.
and motif length classes. Primers with more than one potential PCR product were discarded. Furthermore, the primers were in silico tested on the genome of the grass species *Brachypodium distachyon* strain Bd21-1 downloaded from the *Brachypodium* Genome Database (http://www.brachypodium.org) to reduce the potential risk of cross-amplification on the genome of the host plants. This should allow better use of the primers for experiments targeting the endophytes in planta. An additional primer blasting against *E. festucae* strain E894 genome (Genome Projects at University of Kentucky) with up to two mismatches was performed to identify shared repeats with a closely related species and thus create a large dataset of markers potentially applicable to a broad taxonomic range within the genus.

**Primer Testing**

Lab protocols of newly developed microsatellites were initially tested on a set of samples, hereafter named ‘lab strains’, which included pure cultures of *E. typhina* subsp. *poae*, *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii*, *E. bromicola* and F1 progeny of four crosses (i.e., *E. typhina* subsp. *clarkii* x *E. typhina* subsp. *clarkii*, *E. typhina* subsp. *typhina* x *E. typhina* subsp. *typhina*, *E. typhina* subsp. *typhina* x *E. typhina* subsp. *clarkii*; Tab. 1) grown on potato dextrose agar (PDA; BD Comp., Sparks, Maryland) to test the detectability of hybrids, should they occur in natural populations. Crosses were performed in the greenhouse on freshly emerging stromata on the grasses from which the parental fungal strains were collected.

**Table 1**: Species with their respective hosts and experimental hybrid progeny with numbers of isolates used in this study, and the number of microsatellites out of 38 tested that successfully amplified with M13.

<table>
<thead>
<tr>
<th>Fungal species</th>
<th>Host</th>
<th>No. of isolates</th>
<th>No. of microsatellites</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. typhina</em> subsp. <em>poae</em></td>
<td><em>Poa nemoralis</em></td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>Dactylis glomerata</em></td>
<td>4</td>
<td>37</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>Holcus lanatus</em></td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td><em>E. bromicola</em></td>
<td><em>Bromus erectus</em></td>
<td>2</td>
<td>31</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em> x progeny</td>
<td></td>
<td>6</td>
<td>35</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td>progeny</td>
<td>4</td>
<td>38</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>typhina</em> x progeny</td>
<td></td>
<td>4</td>
<td>37</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em> x progeny</td>
<td></td>
<td>4</td>
<td>37</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>typhina</em> x progeny</td>
<td></td>
<td>4</td>
<td>37</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Microsatellite sequences were *in silico* screened for 100% perfect di-, tri-, tetra-, penta- and hexanucleotides on the genome of *E. typhina* subsp. *poae* with an optimal annealing temperature between 58°C and 60°C (Gardner et al., 2011). In total, 75,432 potential primer pairs amplifying 8,451 microsatellites were generated by our workflow from *E. typhina* subsp. *poae* whole genome sequences. Out of these, 8,136 primer pairs covered 1,084 microsatellites with motifs repeated 5-20 times. Fifty-one primer pairs were randomly chosen across repeat types and labelled with cost-efficient M13 (Schuelke, 2000) to sort out non-working microsatellites in preliminary tests performed on the lab strains. The PCR volume of 10 μl contained approximately 1 ng of genomic DNA, with 5x PCR Buffer (Promega), 5 u/μl Go-Taq Polymerase (Promega, Madison, Wisconsin), 25 mM MgCl₂, 2.5 mM of each dNTP, 2.5 μM of forward primer, and 2.5 μM of each reverse and universal FAM-labelled M13 primers. Amplification was carried out separately for each locus on Sensoquest labcycler (Witec AG, Luzern, Switzerland) with initial denaturing at 94°C for 3 min, followed by 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 30 s, followed by 8 cycles of 94°C for 30 s, 52°C for 30 s, and 72°C for 30 s with a final extension of 60°C for 10 min. All PCR products were run with GeneScan-500 LIZ as size standard on a 3130 Genetic Analyzer (Applied Biosystems, Foster City, California), and electropherograms were analyzed using Geneious 6.1.8 (Drummond et al., 2013).

Out of 51 tested M13 primer pairs, 38 successfully amplifying were retained (Tab. 1). Sixteen of the microsatellites consisted of dinucleotide repeats, 12 of trinucleotide repeats, eight of tetranucleotide repeats, one of a pentanucleotide motif, and one of a hexanucleotide motif. Successful amplification of these primers across our lab strains ranged from 82% to 100% (Tab. 1). The targeted PCR fragments containing microsatellites averaged 294 bp in length and ranged between 107 and 478 bp.

Eventually, 15 polymorphic markers were chosen based on clean and consistent peaks without stuttering in order to decrease the error rate. To create PCR multiplex sets, primer pairs were selected according to their different product sizes and labelled with one of three fluorescent labels (6-FAM, HEX and ATTO0550). Finally, markers were arranged in four multiplex sets (Tab. 2) and tested on the lab strains using a PCR volume of 10 μl containing approximately 1 ng of genomic DNA, 2x Multiplex Mastermix (Multiplex PCR Kit, Quiagen), and 10 x Primermix (2 μM). Amplification was carried out on Sensoquest labcycler (Witec AG) with initial denaturing at 95°C for 5 min, followed by 28 cycles of 95°C for 30 s, 58°C for 90 s, and 72°C for 30 s, with a final extension of 60°C for 30 min.
Table 2: Repeat motifs, primer sequences and size ranges of 15 selected microsatellite loci from *Epichloë* spp. used in four (M1-M4) fluorescent labelled multiplexes.

<table>
<thead>
<tr>
<th>Mix</th>
<th>Loci</th>
<th>Dye†</th>
<th>Repeat motif</th>
<th>Primer sequence (5’ – 3’)</th>
<th>Size range (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>E29</td>
<td>H</td>
<td>(AGC)$_9$</td>
<td>F: TTCCAGCAGCTCTCAATACC</td>
<td>129-201</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: ACAGTGGTTCCTGAGGTTGA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E8</td>
<td>F</td>
<td>(AC)$_{14}$</td>
<td>F: CAGTTGACAAATGTTGAGACC</td>
<td>226-266</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: AGCAAGTTCGTCAACGGTCTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E50</td>
<td>H</td>
<td>(TTG)$_{12}$</td>
<td>F: TCGTCTTGGACTTTGCTTT</td>
<td>315-378</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: TTGAGGTTGTCAGATAACAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E39</td>
<td>F</td>
<td>(GTTTC)$_{12}$</td>
<td>F: GTGCACATGCAATGACAGAG</td>
<td>425-500</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: ACCCACAAGACGGATGACA</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>E47</td>
<td>H</td>
<td>(CTCA)$_9$</td>
<td>F: GCCTGTGAGAAAGACGTGAT</td>
<td>286-354</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: GATCGAACAACCGGATCATAc</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E32</td>
<td>A</td>
<td>(CAG)$_{11}$</td>
<td>F: AGATGAATGGTCAGCAGTTCC</td>
<td>326-347</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: GGACCATCTTCTCAACCTC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E45</td>
<td>A</td>
<td>(GT)$_{15}$</td>
<td>F: TTAGCTGGGAGGATGAGTA</td>
<td>376-466</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: CTGGTACGGAAGCCGAGATA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E13</td>
<td>F</td>
<td>(GA)$_{11}$</td>
<td>F: GTTCCTCCAAGGCTTCAATT</td>
<td>485-551</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: GAGAAACGATATTCGCATTG</td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>E4</td>
<td>F</td>
<td>(AG)$_{9}$</td>
<td>F: ATTCAGCTGTAGCGAGTAGA</td>
<td>126-170</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: CAGAACAATGTCAGATCCTC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E33</td>
<td>F</td>
<td>(TCG)$_{11}$</td>
<td>F: TGCCAGATGTGGTTCAATGCT</td>
<td>329-338</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: AACCCATCTACGCTTTGAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E36</td>
<td>H</td>
<td>(TGC)$_{7}$</td>
<td>F: ATTCGAGAATGGATGACCTG</td>
<td>393-417</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: AAGAAAGGAAATGGGATTGCT</td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td>E22</td>
<td>F</td>
<td>(TGGA)$_{10}$</td>
<td>F: GCAAGGATTTTGGTGATGAA</td>
<td>124-152</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: GCGGATCCTGCTGGGATGAA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E27</td>
<td>H</td>
<td>(GA)$_{8}$</td>
<td>F: TATATTGACGCAGTGCTTT</td>
<td>370-428</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: TGGCATTGCAGAAGCCGATG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E11</td>
<td>F</td>
<td>(CT)$_{11}$</td>
<td>F: GTCAGAGGACGATGACGAG</td>
<td>265-289</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: ATGTAATGCTGCTGCTGCTC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E46</td>
<td>A</td>
<td>(AG)$_{9}$</td>
<td>F: TCGTGACACCTTCCTCTGTAT</td>
<td>382-418</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R: AGAGGTGTCGAGACCATCAT</td>
<td></td>
</tr>
</tbody>
</table>

† H - Hex, F - 6-Fam, A - ATTO0550

**Genotyping of Natural Populations**

Sympatric host grasses infected by *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* were sampled in spring 2013 at a natural grassland site of approximately 2ha in Aubonne, Canton Vaud, Switzerland. One stroma of each infected plant occurring along eight transects across the population were collected every 5m resulting in 27 stromata of *E. typhina* subsp. *typhina* and 50 of *E. typhina* subsp. *clarkii*. Approximately 10 % of the estimated number of
individuals of *E. typhina* subsp. *typhina* and 25% of those of *E. typhina* subsp. *clarkii* were thus sampled. Additionally, in order to compare genetic distances between the two sympatric species to those of a third taxon, 22 samples of *E. typhina* subsp. *poae* from three different locations were included in the analysis. The host of *E. typhina* subsp. *poae* occurs in sympatry with *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* at one edge of the population in Aubonne, but the rare occurrence of the fungus at this site allowed to collect only three samples. Eighteen additional samples came from the Botanical Garden in Zurich, and one from Zollikerberg in Zurich. DNA was extracted from cultured fungi, and in particular from each stroma using mycelium from the interior part of the stroma split open under sterile conditions. DNA extraction followed the cetyl-trimethyl ammonium bromide (CTAB) standard protocol (Doyle and Doyle, 1987). Extracted DNA was subjected to electrophoresis to confirm the presence of template for PCR and quantified using a Nanodrop spectrophotometer (Thermo Scientific, Inc., Waltham, Massachusetts).

The 15 selected markers in four multiplexes described above were used to study the population structure of *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae* at the Aubonne site. In order to test whether genetic diversity in the three species was also detectable in gene sequences, a portion of *tefA* nuclear region was sequenced for ten, 15 and seven representative samples of *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*, respectively. The *tefA* gene is commonly used for distinguishing described sexual *Epichloë* species, as it is rich in polymorphic introns allowing for sufficient resolution (Craven et al., 2001). PCR reactions were performed using primers 5'-GGG TAA GGA CGA AAA GAC TCA-3' (forward) and 5'-CGG CAG CGA TAA TCA GGA TAG-3' (reverse) for *tefA* as previously described (Oberhofer and Leuchtmann, 2012), and sequenced using BigDye Terminator v. 3.1 on a 3130 Genetic Analyzer (Applied Biosystems, Foster City, California). Forward and reverse strands were assembled in Geneious v. 6.1.8 (Drummond et al., 2013) and consensus sequences were aligned using the default plugin of the software. The obtained nucleotide sequences have been submitted to GenBank under accessions KP064360 – KP064391.

**Data Analysis of Natural Populations**

Number of alleles, allelic richness at each locus, gene diversity as well as genotypic linkage disequilibrium based on 10,500 permutations were determined using the program FSTAT v. 2.9.3 (Goudet, 1995). In the analyses of linkage disequilibrium, the data were coded as diploid data, where single-allelic loci were considered to be homozygous. The genetic differentiation among species was assessed in POPGENE v. 1.31 (Yeh et al., 1999) using Fisher exact tests and pairwise *F* sub_{ST} statistics (Weir and Cockerham, 1984). The software program GENALEX v. 6.2 (Peakall and Smouse, 2006) was used to analyze the molecular variance (AMOVA) among and within species of *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii*, and *E. typhina* subsp. *poae* with 9,999 permutations. Sequential Bonferroni
corrections were performed for simultaneous multiple tests (Rice, 1989).

To infer population structure of *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*, PCO-MC analysis was used, a method that combines principal coordinate analysis (PCO) with testing of significant clusters for population structure (Reeves and Richards, 2009). Following the authors’ recommendations (http://lamar.colostate.edu/~reevesp/PCOMC/PCOMC.html), stability of clusters was first assessed with a cut-off set to 15 and significance was then tested at the 0.05 level. Furthermore, the Bayesian model-clustering algorithm implemented in the program STRUCTURE v. 2.3.2 (Falush et al., 2003; Pritchard et al., 2000) was also employed. Iteration parameters were set to 950,000 Monte Carlo Markov Chain (MCMC) iterations preceded by a burn-in period of 50,000 iterations. Ten independent simulations were performed to test for the consistency of the results for *K* ranging from 1 to 5. To infer the number of clusters, i.e. *K*, that best fit the data the method described by Evanno et al. (2005) as implemented in STRUCTURE HARVESTER was used (Earl and vonHoldt, 2012). Membership coefficients estimated by STRUCTURE were visualized using the software programme DISTRACT v. 1.1 (Rosenberg, 2004).

**Results**

*Primer Design*

Blasting of the total 75,432 primer pairs generated in this study against *E. festucae* revealed that 10,406 primer pairs covering 2,022 microsatellites with 100 % perfectness are potentially suitable also for application on other *Epichloë* species. By restricting the repeat number of the microsatellite motifs to 5-20, 885 primer pairs covering 180 microsatellites remained. The full list of these candidate loci with respective descriptions and primer combinations is available as supplementary material (Tab. S1). Within the full set of microsatellites, the motif of hexanucleotide repeats (67 %) was the most abundant class, followed by trinucleotide (15 %), pentanucleotide (12 %), tetranucleotide (5 %) and dinucleotide repeats (1 %; Fig. 1a). In contrast, when filtering for 5-20 motif repeats, trinucleotide repeats (79 %) were the most abundant, followed by dinucleotide (15 %), tetranucleotide (3 %), pentanucleotide (2 %) and hexanucleotide repeats (1 %; Fig. 1b). Among the 15 fluorescent labelled primers developed on *E. typhina* subsp. *poae*, seven were recovered also in the genome of *E. festucae* and yielded successful PCR products when tested *in vitro*. Of the remaining eight microsatellites, three yielded successful PCR products in *E. festucae*, despite the fact that they were not found *in silico* in this species. This indicates that microsatellites can be missed if a draft version of the genome is used, and that unsuccessful blasting does not necessarily translate into negative *in vitro* results.
Genetic Diversity and Structure of Sympatric Epichloë Species

The allelic diversity of 99 isolates of *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii*, and *E. typhina* subsp. *poae* was analyzed within each species at 15 polymorphic loci. The number of alleles per locus ranged from two to 12 for *E. typhina* subsp. *typhina*, from one to 14 for *E. typhina* subsp. *clarkii* and from one to four for *E. typhina* subsp. *poae*. The allelic richness corrected for uneven sample size ranged from two to 11.87 per locus for *E. typhina* subsp. *typhina*, from one to 12.86 for *E. typhina* subsp. *clarkii* and from one to four for *E. typhina* subsp. *poae* (Tab. 3). The mean allelic richness across the three species was 9.58, of which 6.07 alleles were observed across isolates from *E. typhina* subsp. *typhina*, 6.55 alleles in *E. typhina* subsp. *clarkii*, and 2.33 alleles were observed among isolates from *E. typhina* subsp. *poae*. All loci from *E. typhina* subsp. *typhina* were polymorphic, whereas *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae* had 93 % and 80 % polymorphic loci, respectively. Overall, 91 % of the loci from the three species were polymorphic. Significant linkage disequilibrium ($\rho_{\text{adj}} < 0.0005$) was detected in one, two and four of the 105 possible pairwise combinations of loci in *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*, respectively. The mean total gene diversity ($H$) calculated using the allele frequencies across all isolates was 0.53. This was similar to the values obtained for gene diversity estimates, namely 0.664 and 0.628, in isolates from *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*, respectively, and 0.345 observed in isolates from *E. typhina* subsp. *poae* (Tab. 3).
Table 3: Number of alleles, allelic richness and gene diversity per locus at 15 microsatellite loci for *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*.

<table>
<thead>
<tr>
<th>Locus</th>
<th><em>E. typhina</em> subsp. <em>typhina</em></th>
<th><em>E. typhina</em> subsp. <em>clarkii</em></th>
<th><em>E. typhina</em> subsp. <em>poae</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Na</td>
<td>R</td>
<td>H</td>
</tr>
<tr>
<td>E8</td>
<td>5</td>
<td>4.937</td>
<td>0.709</td>
</tr>
<tr>
<td>E39</td>
<td>2</td>
<td>2.000</td>
<td>0.262</td>
</tr>
<tr>
<td>E29</td>
<td>10</td>
<td>9.872</td>
<td>0.875</td>
</tr>
<tr>
<td>E50</td>
<td>9</td>
<td>8.904</td>
<td>0.855</td>
</tr>
<tr>
<td>E13</td>
<td>12</td>
<td>11.871</td>
<td>0.909</td>
</tr>
<tr>
<td>E47</td>
<td>6</td>
<td>5.968</td>
<td>0.672</td>
</tr>
<tr>
<td>E32</td>
<td>5</td>
<td>4.999</td>
<td>0.658</td>
</tr>
<tr>
<td>E45</td>
<td>10</td>
<td>9.873</td>
<td>0.892</td>
</tr>
<tr>
<td>E4</td>
<td>3</td>
<td>2.969</td>
<td>0.544</td>
</tr>
<tr>
<td>E33</td>
<td>2</td>
<td>2.000</td>
<td>0.359</td>
</tr>
<tr>
<td>E36</td>
<td>6</td>
<td>5.937</td>
<td>0.744</td>
</tr>
<tr>
<td>E22</td>
<td>2</td>
<td>2.000</td>
<td>0.484</td>
</tr>
<tr>
<td>E11</td>
<td>5</td>
<td>4.936</td>
<td>0.567</td>
</tr>
<tr>
<td>E27</td>
<td>12</td>
<td>11.840</td>
<td>0.892</td>
</tr>
<tr>
<td>E46</td>
<td>3</td>
<td>2.969</td>
<td>0.544</td>
</tr>
<tr>
<td>Mean</td>
<td>6</td>
<td>6.072</td>
<td>0.664</td>
</tr>
</tbody>
</table>

*Abbreviations: Na number of alleles per locus, R corrected allelic richness, H gene diversity

The observed levels of genetic differentiation (Fischer exact test; *p* < 0.001) among *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae* reflect a clear structuring of the three species. These results were supported by high pairwise *F*ST values of 0.446 between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *poae*, 0.453 between *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*, and 0.280 between the more closely related *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* (Tab. 4). The AMOVA confirmed these results by showing that 62% of the molecular variance is found within species and 38% among species (Tab. 5).

Table 4: Pairwise *F*ST values among *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*. An asterisk indicated the *F*ST value is highly significantly different from 0 (*p* < 0.001). *P* values were adjusted using the Bonferroni procedure.

<table>
<thead>
<tr>
<th></th>
<th><em>E. typhina</em> subsp. <em>typhina</em></th>
<th><em>E. typhina</em> subsp. <em>clarkii</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td>0.280*</td>
<td></td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>poae</em></td>
<td>0.446*</td>
<td>0.453*</td>
</tr>
</tbody>
</table>
Table 5: AMOVA results testing the genetic structure among *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae* at 15 microsatellite loci (9999 permutations).

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among species</td>
<td>2</td>
<td>172.57</td>
<td>2.67</td>
<td>38</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Within species</td>
<td>96</td>
<td>414.53</td>
<td>4.32</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>98</td>
<td>587.10</td>
<td>6.98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PCO-MC analyses revealed three stable and statistically significant clusters (*p* < 0.05) that corresponded to the three host-specific species (Fig. 2). In this analysis, 19% of the variation was explained by the first axis, 11.6% by the second axis, and 4.3% by the third axes. No samples were found in between the clusters indicating that all samples are assigned to the three genotypically separate species. The STRUCTURE analysis yielded results consistent with those obtained from PCO-MC. The distribution of the maximum likelihood was the highest for *K*=3. The first cluster included all genotypes of isolates representing *E. typhina* subsp. *typhina* while the second cluster consisted of genotypes of *E. typhina* subsp. *clarkii* and the third cluster of *E. typhina* subsp. *poae*. The assignment of the individuals in the plot showed very reduced levels of admixture.

![Figure 2](image-url)

**Figure 2:** A) PCO-MC analysis based on 15 microsatellite loci for *E. typhina* subsp. *typhina* (circles), *E. typhina* subsp. *clarkii* (squares) and *E. typhina* subsp. *poae* (diamonds). Solid lines indicate stable and statistically significant (*p* < 0.05) clusters, whereas dashed line indicates stable cluster that do not pass the significance threshold. B) Bayesian clustering of *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae* at 15 microsatellite loci as inferred from STRUCTURE under the assumption of *K*=3 populations.

The alignment of the sequences of the nuclear gene *tefA* was ~620 bp in length and showed a considerable number of species-specific fixed single nucleotide polymorphisms (SNPs) for all three species of the study. In particular, four, three, and eight SNPs were found
to be fixed among representatives of *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poeae*, respectively. Hence, these SNPs have diagnostic value for the identification of the three species (Fig. S1).

**Discussion**

*Genetic Diversity and Structure of Sympatric Epichloë Species*

Host-specificity is hypothesised to promote reproductive isolation between sympatric fungal species when mating is constrained to occur within the host (Kohn, 2005). Evidence of the role of host-specificity in maintaining species boundaries has emerged from a number of systems including *Ascochyta* species, *Mycosphaerella graminicola* and *Microbotryum* species, hence this mechanisms is hypothesised to play an important role in the speciation of these fungi (Büker et al., 2013; Peever, 2007; Stukenbrock et al., 2011). To date, no in-depth study has addressed the validity of species boundaries in host-specific endophytes, and we contribute to fill this gap by presenting evidence from sympatric populations of *Epichloë*.

Our study provides clear evidence of genetic discontinuity between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*, indicating that reproductive isolation mechanisms exist to keep sympatric populations both morphologically and genetically distinct. Moreover, although the third species (*E. typhina* subsp. *poeae*) was represented by only three samples from Aubonne, these had clearly distinct genotypes that clustered with additional samples of this species from two different geographical regions. This offers preliminary evidence that also *E. typhina* subsp. *poeae* is maintained isolated from *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*. In fact, both PCO-MC and Bayesian analyses placed the isolates representing the three host-specific species in separate clusters. High pairwise *F_{ST}* values and the evidence provided by sequencing data from the nuclear gene *tefA*, where several species-specific fixed SNPs for each species were recovered, further supports these findings. Consistently, *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poeae* represent separate evolutionary lineages as previously shown in phylogenetic trees reconstructed from *actA*, *tubB* and *tefA* sequences (Craven et al., 2001). Overall, our results indicate high genetic differentiation among the three species possibly maintained by very limited gene flow.

To date, studies that have assessed the genetic structure of *Epichloë*/*Neotyphodium* populations have focused on asexual species. Because asexual endophytes do not form wind-dispersed ascospores and are vertically transmitted through the infection of developing seeds of their host grasses, gene flow may only be mediated by the process of plant seed dispersal. As this is expected to have a small effect on geographically distant populations (Arroyo García et al., 2002; Sullivan and Faeth, 2004), high levels of *F_{ST}* are expected among endophyte populations. However, results of *F_{ST}* analyses indicate that asexual reproduction is hardly a good predictor of levels of *F_{ST}* among geographically isolated...
populations. Low levels of gene flow were found among populations of asexually reproducing *Neotyphodium sibiricum* \((F_{ST} = 0.6799)\) and *N. gansuense* \((F_{ST} = 0.3490)\), consistent with the observed dominance of clonal reproduction (Zhang et al., 2010). On the other hand, gene flow was inferred to be high in asexual *E. festucae* infecting *Festuca rubra* \((F_{ST} = 0.197\) in Garcia et al. (2002) and \(F_{ST} = 0.0814\) in Wäli et al. (2007)), where genetic exchange may be attributed to the presence of migrants. Despite this contrasting evidence, results from a study on the endophytes of *Brachypodium sylvaticum* has shown that sexual stromata-forming and asexual asymptomatic isolates form genetically clearly differentiated subpopulations (Bucheli and Leuchtmann, 1996). While evidence is accumulating on the genetic structure of populations from asexually reproducing *Epichloë* species, to our knowledge no study has addressed the genotypic delimitation of sympatric sexually reproducing taxa.

This study showed high genetic differentiation among populations with pairwise \(F_{ST}\) values of 0.446 between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *poae*, 0.453 between *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*, and 0.280 between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*. The higher \(F_{ST}\) value of *E. typhina* subsp. *poae* in pairwise comparisons with the other two species indicates that *E. typhina* subsp. *poae* is a more genetically isolated taxon, supporting previous results (Craven et al., 2001). While this result can be caused by the geographic isolation of a great proportion of the *E. typhina* subsp. *poae* samples external to the Aubonne site, comparisons of the population statistics among the three species reveal peculiar characteristics for this taxon. We found similar levels of allelic richness for *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*, but these values were more than two times higher than for *E. typhina* subsp. *poae* (Tab. 3). Furthermore, *E. typhina* subsp. *poae* showed lower percentage of polymorphic loci and lowest gene diversity (Tab. 3), despite the fact that it was represented by samples from three different locations. These characteristics of *E. typhina* subsp. *poae* may be explained by its different transmission mode. In contrast to *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* that are only horizontally transmitted by sexual ascospores, *E. typhina* subsp. *poae* can transmit both horizontally and vertically by seed. This may lead more frequently to clonal endophytes in plants growing from seeds. Accordingly, only 16 genotypes out of 22 were unique in *E. typhina* subsp. *poae*, whereas in *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* all genotypes were found to be unique.

**Reproductive Isolation of Epichloë Species**

Hypotheses on the origin of reproductive barriers should include both intrinsic and extrinsic mechanisms, and our results provide preliminary insights to evaluate alternative scenarios. An obvious mechanism of extrinsic prezygotic isolation includes the preferential transfer of spermata by *Botanophila* flies that actively contribute to exchange of gametes. Yet, given that *Botanophila* flies are largely unspecific in their visiting behavior, spermata from
different hosts are collected and actively transported in mixtures among the stromata (Bultman et al., 2011), thus a preference by fly vectors is unlikely to constitute an efficient barrier. Alternatively, gene flow between *Epichloë* species could be restricted via prezygotic assortative mating on the stroma, if individuals were able to discriminate between conspecific and heterospecific spores. Preliminary data in *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* showed that the outcome of matings depends on the concentrations of conspecific and heterospecific spores applied in mixtures on the stroma (Schirrmann, unpublished data). For example, if conspecific and heterospecific spores were applied in equal concentrations, only progeny of intraspecific matings were found, thus indicating a preference for spores from the same species. These results are consistent with previous evidence for effective interference between matings of different *Epichloë* species on different parts of the stroma surface, depending upon the fungal strains involved (Chung and Schardl, 1997). The observed interferences suggested that signals generated after positive mating could be transported throughout the stroma and thus prevent or abort other mating interactions on the same stroma.

In contrast, intrinsic postzygotic isolation is often associated with hybrid sterility or inviability as shown by *in vitro* crosses that are initiated and then later aborted (Giraud et al., 2008). In *Neurospora* species interspecific compared to intraspecific crosses resulted in fewer viable ascospores, and *Microbotryum* species produced fewer viable mycelia in interspecific crosses (Dettman et al., 2003; Le Gac et al., 2007). Furthermore, in *Saccharomyces* species, the incompatibility between genes of the nuclear and mitochondrial genomes causes hybrid sterility, which might have evolved as a by-product of ecological adaptation to different nutrient sources (Lee et al., 2008). In our system, the viability of experimentally generated hybrids between species of the *E. typhina* complex was not reduced and their radial growth rates *in vitro* were not significantly different, suggesting a lack of obvious intrinsic postzygotic isolation barriers between the parental strains (Schirrmann, unpublished data). Therefore, postzygotic barriers may depend on extrinsic factors as for example a decreased fitness of the hybrids due to a likely maladaptation to host grasses, preventing infection or persistence of hybrids. Host-specificity was shown in a study by Chung et al. (1997) where the seedling infection frequency and stability of *E. typhina* isolates and their progeny were tested. A reduced infectivity of interspecific progeny in parental hosts was demonstrated previously also for *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* (Leuchtmann and Steinebrunner, 2012) and was confirmed by reciprocal inoculation experiments for these species and *E. typhina* subsp. *poae* (Schirrmann, unpublished data).

Overall, the evidence collected so far indicates that natural populations of sympatric *Epichloë* species may be genetically differentiated due to assortative mating on the stroma and host-specificity of parental strains. In the future, further investigations will hopefully test these hypotheses thoroughly.
Primer Design and Testing

In this study, we developed new microsatellite markers to study population structure at the interspecific level within the *E. typhina* complex. Of the 38 successfully amplifying microsatellites tested on the investigated lab strains, i.e., *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii*, *E. typhina* subsp. *poae*, *E. bromicola* and the four progeny from inter- and intraspecific crosses, 82 % to 100 % yielded a PCR product. This suggests that our set of loci has a high probability of cross-species transferability within *Epichloë* and that hybrids could be detected if they would be present in nature (Tab. 1). Additionally, the large set of markers identified in *E. typhina* subsp. *poae* and successfully *in silico* verified in *E. festucae* provides a valuable resource of candidate primers applicable to a broad taxonomic range of *Epichloë*. However, careful testing should always be performed to optimise amplification protocols and multiplex design, as well as to rule out unspecific amplification products. Overall, the loci detected and tested here constitute a substantial genetic resource that complement previous microsatellite markers developed for *Epichloë/Neotyphodium* species (Groppe et al., 1995; Groppe and Boller, 1997; Moon et al., 1999; van Zijl de Jong et al., 2003).

Conclusions

Using a new set of 15 microsatellites, we assessed the population structure of sympatric species of the *E. typhina* complex, i.e. *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*, at a natural field site in Aubonne, Switzerland. Our results showed that high genetic differentiation and very limited gene flow exist among these three species, indicating that mechanisms of reproductive isolation keep natural *Epichloë* populations genetically distinct. Preliminary evidence suggests that host-specificity and maladaptation of hybrids to host grasses may act as reproductive isolation barriers in *Epichloë* and therefore promote their speciation. However, further studies are needed to disentangle the complex interplay of isolation barriers that may be responsible to maintain species boundaries in natural populations.

Author’s contributions

MKS performed experiments, analyzed data and wrote the paper. SZ contributed analytical tools. SF analyzed data and wrote the paper. AL designed experiments and wrote the paper.

Acknowledgements

This study was funded by the Swiss National Science Foundation (31003A_138479). Microsatellite data were generated in the GDC of ETH Zurich. C. L. Schardl gave advice and provided access to unpublished genome data. We thank B. Blattmann for laboratory assistance, and M. C. Fischer and N. Zemp for providing methodological details.

References


Peever, T.L., 2007. Role of host specificity in the speciation of Ascochyta pathogens of cool


of microsatellite markers in the genome of the plant pathogen Ceratocystis fimbriata.
Fungal Biol 117, 545–555. doi:10.1016/j.funbio.2013.06.004
Figure S1: Species-specific fixed and diagnostic single nucleotide polymorphisms for *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae* in sequences of the nuclear gene *tefA*

Only informative portions of the alignments are shown. Numbers on top of the alignment indicate original nucleotide positions prior to removal of uninformative characters (dots).

Table S1: published only online (http://dx.doi.org/10.1007/s00248-014-0556-5).
CHAPTER 2

THE ROLE OF HOST-SPECIFICITY IN THE REPRODUCTIVE ISOLATION OF *EPICHLOË* ENDOPHYTES REVEALED BY RECIPROCAL INFECTIONS

Published as

¹ETH Zurich, Institute of Integrative Biology, Plant Ecological Genetics, Zürich, Switzerland
Abstract
Speciation in sexually reproducing organisms hinges on reproductive barriers that reduce gene flow between species or preclude the formation of hybrids. Here, we studied potential reproductive barriers in four members of the *Epichloë typhina* (Ascomycota, Clavicipitaceae) complex, i.e. *E. typhina* subsp. *typhina*\(^4\) infecting *Dactylis glomerata*, *E. typhina* subsp. *clarkii* infecting *Holcus lanatus*, *E. typhina* subsp. *paeae* infecting *Poa nemoralis* and *E. typhina* subsp. *typhina*\(^5\) infecting *P. trivialis*. Reciprocal inoculation tests showed that these endophytes are host-specific. This suggests that reproductive isolation among *Epichloë* strains may be the result of specialization to one host, on which mating between different individuals occurs. Furthermore, significantly lower infection frequencies of F1 progeny from crosses between host-strains compared to parental strains and within host-strain progeny suggest that host-dependent effects upon hybrid fitness exist, which would conform to an extrinsic postzygotic isolation barrier. Our results may explain, why members of the *E. typhina* complex remain genetically differentiated in natural populations.

Introduction
Speciation, the process by which biodiversity is generated, is characterized by the evolution of barriers to genetic exchange between previously interbreeding populations (Rieseberg and Willis, 2007). To define species, different species concepts are used depending on the properties upon which they are based, e.g. ecology, morphology or phylogeny (De Queiroz, 2007). A criterion that is commonly used to define biological species is the lack of interbreeding between species (Mayr, 1969). Understanding the ecological and genetic basis of isolation barriers is therefore of fundamental interest in evolutionary biology. Especially mating between sympatric species, which provides the potential for hybridization and may be a regular occurrence in natural populations, should be prevented (Butlin et al., 2008). Hybridization can be limited by a complex interplay of pre- and postzygotic reproductive barriers that are expected to arise as a result of divergence of populations or incipient species (Dobzhansky, 1937; Mallet, 2005). Prezygotic isolation mechanisms prevent the formation of hybrid zygotes, because mating does not occur due to differences in the time of reproduction, morphology, habitat, or by specialization of a biotic vector (Bultman and Leuchtmann, 2003). In contrast, intrinsic postzygotic barriers may be associated with inviability or reduced fitness of hybrid strains (Giraud et al., 2008). However, hybrids can be viable and fertile under *in vitro* conditions but unfit in a natural environment due to extrinsic postzygotic barriers that are linked to ecological factors. This can be the case if hybrids display intermediate traits between parental phenotypes and are poor competitors in either parental environment.

\(^4\) Name in the publication: *E. typhina* infecting *D. glomerata*
\(^5\) Name in the publication: *E. typhina* infecting *P. trivialis*
Reproductive isolation has been widely studied in animals and plants (Coyne and Orr, 2004; Seehausen et al., 2014; Widmer et al., 2008), but much fewer cases have been investigated in fungi (Butlin et al., 2012; Giraud et al., 2008; Kohn, 2005; Schardl and Craven, 2003). The sexual species of the genus *Epichloë* (Ascomycota, Clavicipitaceae) are an interesting fungal model system for studying mechanisms of speciation, because both pre- and postzygotic reproductive barriers may potentially be involved. *Epichloë* species are endophytes of pooid grasses (Pooidaeae), growing in plant apoplasts primarily in shoot apical meristems and leaf sheaths (Schardl et al., 2004) with little or no negative effects on the vegetative tissues of their hosts. The fungi may infect host embryos and seeds either vertically via the infected mother plant or horizontally via ascospores (Leuchtmann and Clay, 1997). During their sexual life cycle they may form external reproductive structures, i.e., stromata, on flowering shoots, which arrest the development of the affected host inflorescence by surrounding the host flag-leaf sheath and immature inflorescence. Sexual reproduction is therefore directly linked to the fitness of the host and is known as ‘choke disease’ (Western and Cavett, 1959). *Epichloë* species are heterothallic with a bipolar mating system, and receptive hyphae (male gametes) and spermatia are both formed on the same stroma (White and Bultman, 1987). In natural ecosystems, *Botanophila* flies (Anthomyiidae) transfer spermatia between stromata of opposite mating types (mat-1 and mat-2) in a process similar to pollination (Bultman et al., 1998). After successful mating, young asci develop within perithecia and undergo karyogamy and meiosis. Asci contain eight haploid ascospores that are ejected, wind-dispersed and mediate horizontal transmission through infections of host florets and then seeds (Chung and Schardl, 1997a).

Previously, *Epichloë* species have been classified under two different genera depending on differences in their mode of reproduction. Asexual species were referred to *Neotyphodium* (Glenn et al., 1996), whereas sexual species were placed in the genus *Epichloë*. Following recent changes in nomenclatural rules (McNeill et al., 2012), these fungi were realigned under one genus by Leuchtmann et al. (2014), accepting 10 teleomorph-typified species and 24 anamorph-typified species of *Epichloë* (including three subspecies and six varieties). Sexual species have been described based on morphological characters that include size and disarticulation patterns of ascospores, size and density of perithecia, and host species preferences (Leuchtmann, 2003). Besides their morphological characters, most species can further be distinguished by apparent sexual incompatibility among them established by experimental mating tests and have therefore been referred to mating populations (Chung and Schardl, 1997b; Leuchtmann and Schardl, 1998; Leuchtmann et al., 1994; Schardl et al., 1997).

The recognition of species as mating populations and characterization of their host ranges (Leuchtmann, 2003) allowed us to study the genetics of host compatibility and incompatibility. Most mating populations show rather narrow host ranges restricted to grass species of one genus or closely related host genera with the exception of *E. typhina* that has
been assumed to be able to colonize multiple hosts (Craven et al., 2001), but is in fact represented by a complex of different taxa, including the subspecies \textit{E. typhina} subsp. \textit{clarkii} and \textit{E. typhina} subsp. \textit{poae} (Leuchtmann et al., 2014). However, although members of the \textit{E. typhina} complex are usually interfertile in experimental mating tests, many of the host-associated strains appear to be genetically distinct even in sympatry (Schardl et al., 2007; Schirrmann et al., 2014), suggesting that they remain reproductively isolated in nature.

Here, we studied \textit{in vitro} and \textit{in vivo} growth patterns of four members of the \textit{E. typhina} complex, i.e., \textit{E. typhina} subsp. \textit{typhina} infecting \textit{Dactylis glomerata} (Dg), \textit{E. typhina} subsp. \textit{clarkii} infecting \textit{Holcus lanatus} (Hl), \textit{E. typhina} subsp. \textit{poae} infecting \textit{Poa nemoralis} (Pn) and \textit{E. typhina} subsp. \textit{typhina} infecting \textit{Poa trivialis} (Pt), and F1 progeny from reciprocal crosses (if available) to gain insight into mechanisms of their reproductive isolation. Radial growth was measured \textit{in vitro} as a proxy for competitiveness to infect grass hosts, and to assess the effect of temperature on fungal growth. Reduced growth of hybrids compared to parental strains would indicate a postzygotic isolation barrier. By inoculation of wounded seedlings (Latch and Christensen, 1985), we infected new host grasses with parental and hybrid strains, as it may occur in the sexual life cycle, and assessed their host-specificity. Furthermore, we investigated the long-term stability of experimental infections, because infections may be lost, if genotypes are maladapted (Saikkonen et al., 2010). Finally, by counting the number of stromata and symptomless tillers, i.e. shoots that sprout from the base of a grass, we assessed the reproductive success of the inoculated strains.

\textbf{Material and methods}
\textit{Epichloë isolates and crosses}

All fungal isolates were cultured from systemically infected \textit{Holcus lanatus}, \textit{Dactylis glomerata}, \textit{Poa nemoralis} or \textit{P. trivialis} collected in Switzerland or France. These isolates were obtained from surface sterilized stems and leaf sheaths and subsequently grown on potato dextrose agar (PDA; BD Company, Sparks, Maryland) as previously described (Leuchtmann, 1994). In total, four isolates (representing two different clonal strains) of \textit{E. typhina} subsp. \textit{clarkii} (Hl), five of \textit{E. typhina} subsp. \textit{typhina} (Dg), two of \textit{E. typhina} subsp. \textit{poae} (Pn), and two of \textit{E. typhina} subsp. \textit{typhina} (Pt) from their native host grasses were used in the study. These will be referred to “Hl strains”, “Dg strains”, “Pn strains” and “Pt strains”, respectively. Identifiers of individual parental strains can be found in Tab. 1.
Table 1: *Epichloë* isolates with taxonomic identity, host, sampling location in Switzerland or France (F) and mating type used in this study.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>No.</th>
<th>Fungal taxon</th>
<th>Host grass</th>
<th>Location</th>
<th>Mating type</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI.1</td>
<td>0807*</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>Aubonne</td>
<td>mat-1</td>
</tr>
<tr>
<td>HI.2</td>
<td>0808*</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>La Rippe</td>
<td>mat-2</td>
</tr>
<tr>
<td>HI.3</td>
<td>1205*</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>Aubonne</td>
<td>mat-1</td>
</tr>
<tr>
<td>HI.4</td>
<td>1206*</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>La Rippe</td>
<td>mat-2</td>
</tr>
<tr>
<td>DG.1</td>
<td>1203</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>D. glomerata</em></td>
<td>Vesancy (F)</td>
<td>mat-1</td>
</tr>
<tr>
<td>DG.2</td>
<td>1204</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>D. glomerata</em></td>
<td>Vesancy (F)</td>
<td>mat-2</td>
</tr>
<tr>
<td>DG.3</td>
<td>1217</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>D. glomerata</em></td>
<td>Merishausen</td>
<td>mat-2</td>
</tr>
<tr>
<td>DG.4</td>
<td>1218</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>D. glomerata</em></td>
<td>Zumikon</td>
<td>mat-1</td>
</tr>
<tr>
<td>DG.5</td>
<td>1305</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>D. glomerata</em></td>
<td>Changins</td>
<td>mat-2</td>
</tr>
<tr>
<td>PN.1</td>
<td>1207</td>
<td><em>E. typhina</em> subsp. <em>poae</em></td>
<td><em>P. nemoralis</em></td>
<td>Aubonne</td>
<td>mat-1</td>
</tr>
<tr>
<td>PN.2</td>
<td>1208</td>
<td><em>E. typhina</em> subsp. <em>poae</em></td>
<td><em>P. nemoralis</em></td>
<td>Aubonne</td>
<td>mat-2</td>
</tr>
<tr>
<td>PT.1</td>
<td>1201</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>Poa trivialis</em></td>
<td>Küsnacht</td>
<td>mat-1</td>
</tr>
<tr>
<td>PT.2</td>
<td>1202</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>P. trivialis</em></td>
<td>Küsnacht</td>
<td>mat-2</td>
</tr>
</tbody>
</table>

* These four isolates are representing two strains only (0807/1205 and 0808/1206)

Crosses between the host-associated strains were conducted in the greenhouse using different isolates and freshly emerging stromata on the grasses from which parental strains were obtained. To initiate fertilization, stromata of opposite mating type, which has been determined previously, were rubbed against each other. Beginning three weeks after mating, perithecia were squashed in aniline blue stain (Clark et al., 1983) and examined by light microscopy for ascospores. Once mature ascospores were evident, F1 progeny was established in single-ascospore cultures (Leuchtmann et al., 1994). In this context, each ascospore represents one progeny.

In addition to matings within host-strains, the following pairings were performed: (1) between *E. typhina* subsp. *typhina* (Dg) and *E. typhina* subsp. *clarkii*, (2) between *E. typhina* subsp. *typhina* (Dg) and *E. typhina* subsp. *poae*, (3) between *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae*, (4) between *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *typhina* (Pt) and (5) between *E. typhina* subsp. *typhina* (Dg) and *E. typhina* subsp. *typhina* (Pt). Identifiers of single spore progeny from successful crosses can be found in Table 2. Crosses between *E. typhina* subsp. *typhina* (Dg) and *E. typhina* subsp. *poae*, and between *E. typhina* subsp. *clarkii* and *E. typhina* subsp. *poae* were unsuccessful and are therefore not listed in the Table.
Table 2: Progeny from crosses within and between *Epichloë* host-strains (see Table 1) and number of progeny used in this study.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Cross no.</th>
<th>Parental host-strains</th>
<th>No. of progeny</th>
</tr>
</thead>
<tbody>
<tr>
<td>HL.4_Hl.3</td>
<td>1101</td>
<td>HL.4 (E. typhina subsp. clarkii) × HL.3 (E. typhina subsp. clarkii)</td>
<td>4</td>
</tr>
<tr>
<td>HL.4_Dg.4</td>
<td>1103</td>
<td>HL.4 (E. typhina subsp. clarkii) × Dg.4 (E. typhina subsp. typhina)</td>
<td>3</td>
</tr>
<tr>
<td>HL.3_Hl.4</td>
<td>1105</td>
<td>HL.3 (E. typhina subsp. clarkii) × HL.4 (E. typhina subsp. clarkii)</td>
<td>3</td>
</tr>
<tr>
<td>Dg.3_Hl.3</td>
<td>1106</td>
<td>Dg.3 (E. typhina subsp. typhina) × HL.3 (E. typhina subsp. clarkii)</td>
<td>2</td>
</tr>
<tr>
<td>Dg.1_Dg.2</td>
<td>1109</td>
<td>Dg.1 (E. typhina subsp. typhina) × Dg.2 (E. typhina subsp. typhina)</td>
<td>4</td>
</tr>
<tr>
<td>Dg.1_Hl.4</td>
<td>1110</td>
<td>Dg.1 (E. typhina subsp. typhina) × HL.4 (E. typhina subsp. clarkii)</td>
<td>3</td>
</tr>
<tr>
<td>Dg.5_Pt.1</td>
<td>1211</td>
<td>Dg.5 (E. typhina subsp. typhina) × Pt.1 (E. typhina subsp. typhina)</td>
<td>4</td>
</tr>
<tr>
<td>HL.4_Pt.1</td>
<td>1213</td>
<td>HL.4 (E. typhina subsp. clarkii) × Pt.1 (E. typhina subsp. typhina)</td>
<td>4</td>
</tr>
</tbody>
</table>

Genotyping of progeny

To confirm parentages of the F1 progeny, DNA was extracted from the parental strains as well as from the progeny using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer’s protocol. Extracted DNA was subjected to electrophoresis to confirm the presence of template for PCR and quantified using a Nanodrop spectrophotometer (Thermo Scientific, Inc., Waltham, Massachusetts). For genotyping of the fungal strains, 15 newly developed microsatellite markers were amplified as previously described (Schirrrmann et al., 2014). All PCR products were run with GeneScan-500 LIZ as size standard on a 3130 Genetic Analyzer (Applied Biosystems, Foster City, California), and electropherograms were analyzed using Geneious 6.1.7 (Drummond et al., 2013).

In vitro growth of fungal isolates

In successful matings radial growth of hybrids and parental strains was measured and the effects of temperature on growth examined. Eleven parental isolates and 27 F1 progeny from eight crosses were selected. The source material for the growth experiment was a small block of agar (1mm²) removed from the margin of a 14 days old colony on PDA. Six replicate petri dishes were inoculated for each of the 38 selected fungal isolates adding up to a total of 228 dishes. Half of the replicates of each isolate were subjected to their assumed optimal temperature (24°C) and the other half to restrictive temperature (18°C). All colonies were grown in the dark on PDA for four weeks. Three measurements of diameter per colony, located 120° apart, were taken and then averaged for each colony.
Measurements were made after 14, 21 and 28 days, but only those after 14 days were reported, because relative growth rates were identical. Significant differences of radial growth were tested with a binominal analysis of variance (ANOVA). All statistical analyses were conducted with R (version 2.13.0; R Development Core Team 2011).

Germination of grass seeds
For the inoculation experiments, seeds of commercial cultivars of D. glomerata (cv. ‘Prato’, Agroscope Reckenholz-Tänikon), H. lanatus (3624/K01/0.1, Otto Hauenstein seeds), P. nemoralis (cv. ‘Enhary’, Otto Hauenstein seeds), and P. trivialis (37185/N01/0.2, Otto Hauenstein seeds) were used. Seeds were surface sterilized (Leuchtmann and Clay, 1990), placed on water agar (1.5 %) in petri dishes and incubated for germination in a climate cabinet set to 12h day at 24°C and 12h night at 18°C (Sanyo Versatile Environmental Test Chamber, Sanyo Commercial Solutions, Bensenville, USA). Illumination devices were Osram L 40W 20SA and Osram L 40W 25 SA (Munich, Germany).

Inoculation of seedlings with parental isolates and F1 progeny
To assess compatibility of parental isolates and F1 progeny with host plants, seedlings of each host species were inoculated by the procedure described in Latch and Christensen (Latch and Christensen, 1985) with small modifications. For the inoculation treatment, fungal cultures were grown on agar medium containing 1 % malt extract, 1 % glucose, 0.25 % bacto peptone, and 0.25 % yeast extract (Leuchtmann, 1994) for a maximum of two weeks. Care was taken that cultured fungal strains used as inoculum produce conidia, because degraded morphotypes produce no aerial hypha or conidia and would have reduced infectivity (Christensen, 1995).

Seedlings were inoculated 5-7 days after germination when the first leaf was 3-12 mm long and still partly enclosed by the coleoptile. The shoot meristem was visible by a light microscope as a slight swelling at the junction of the mesocotyl and the radicle. Fungal mycelium was introduced into seedlings through a stem slit near the meristems using a sterile hypodermic needle. This was done with the help of a stereomicroscope within a laminar-flow bench. To exclude contamination from another source than the inoculum, control seedlings were treated the same, except using sterile water instead of an inoculum. Each dish contained only one inoculum/seedling combination and for each combination a separate hypodermic needle was used. At least ten seedlings were inoculated with each parental strain or F1 progeny selected. Inoculated seedlings were kept for another 14 days in the climate cabinet before being transplanted into commercial potting substrate (Ökohum Universalerde, Herbertingen, Germany). Although some seedlings died before transplantation, at least seven of the inoculated seedlings survived in each case.
**Analysis of infection frequencies**

After about two months, plants were checked for infection using two leaf sheaths per plant. Samples were stained with aniline blue and examined at 400x magnification under the microscope for the presence of fungal hyphae. A haphazard subset of fungal strains in successfully infected host plants was genotyped with 15 microsatellites as described above to confirm the identity of the strains. Infection frequency was determined as the number of infected plants over the total number of inoculated plants tested for infection. Significant differences between infection frequencies of progeny from crosses within compared to between host-strains were tested with a binominal analysis of variance (ANOVA). One cross included several single-ascospore progeny. Furthermore, significance of infection frequencies of hybrid compared to parental strains was tested using a Dunnett generalized linear mixed model (GLM). All statistical analyses were conducted with R (version 2.13.0; R Development Core Team 2011).

**Infection stability of fungal isolates**

To test for long-term stability of infected *D. glomerata, H. lanatus* and *P. nemoralis*, plants were cloned and kept for another 12 months in the greenhouse set to 22°C 10 h day (30 kLux) and 18°C 14 h night. Because *E. tychina* subsp. *typhina* infecting *P. trivialis* was inoculated at a later stage of the experiment, this species was not screened for infection stability. For vernalization of the grasses, temperatures were set to 18°C 10 h day (30 kLux) and 15°C 14 h night for 5 days, 14°C 8 h day (10 kLux) and 10°C 16 h night for 21 days, 4°C 8 h day (10 kLux) and 4°C 16 h night for 2 months, 14°C 8 h day (15 kLux) and 11°C 16 h night for 21 days, 18°C 10 h day (30 kLux) and 15°C 14 h night for 5 days, and finally 22°C 10 h day (30 kLux) and 18°C 14 h night for 5 days. At the end of the growing period, total DNA was extracted from lyophilized ground leaf sheaths of approximately 400 mg fresh weight, following the same method as for the genotyping of the progeny. If all plants infected with an isolate maintained their infection after this time period, the isolate-host species association was considered highly stable; loss of infections from some, but not all previously infected plants was considered moderately stable; and if all infections with an isolate were lost, the association was considered unstable. Furthermore, we counted the number of symptomless tillers and stromata of *E. tychina* subsp. *clarkii* infecting *H. lanatus*. This was not possible for the other inoculated host-associated strains, because they did not produce stromata at the end of the growing period although they were vernalized the same way.
Results

Genotyping of progeny

We analysed all available parental isolates at 15 microsatellite loci and compared them to the progeny. Because alleles in the F1 progeny segregate in Mendelian fashion, we could confirm the parents of five crosses (Hl.4_Dg.4, Dg.3_Hl.3, Dg.1_Hl.4, Dg.5_Pt.1 and Hl.4_Pt.1; Tab. 2). However, for three of the crosses involving Pn strains the expected paternal parent could not be confirmed, instead fertilization occurred within host-strains (Hl.4_Hl.3, Hl.3_Hl.4 and Dg.1_Dg.2; Tab. 2). These crosses presumably occurred unintentionally with spores from conspecific stroma-forming plants present in the greenhouse, because E. typhina subsp. poae obviously cannot hybridize with E. typhina subsp. clarkii or E. typhina subsp. typhina (Dg). In the following experiments, we used progeny from within and between host-strain crosses for inoculation of D. glomerata, H. lanatus and P. nemoralis. By comparing infection frequencies of both types of progeny we could rule out any artefact of the inoculation treatment, and subsequently used only progeny from between host-strain crosses for the inoculations of P. trivialis.

In vitro growth of fungal isolates

Measurements of radial growth of the parental isolates and their F1 progeny in vitro showed a clear effect of temperature. As expected, colonies were growing significantly faster at their near optimal temperature of 24°C compared to 18°C (ANOVA; p < 0.001; Fig. 1). Relative growth rates were the same after 14, 21 and 28 days and therefore only results after 14 days are shown. However, the growth rate could not be related to the origin of the endophyte (ANOVA; p > 0.1). Different progeny from one cross showed differences in their growing behavior compared to each other and to their parental isolates, depending on the genotype involved (Fig. 1). The parental isolates showed similar growth patterns, regardless of whether they belonged to the same or to a different host-associated strain. Furthermore, we could not find a correlation between in vitro growth and in vivo infection frequencies (data not shown). Some strains showed a high in vitro growth rate but low in vivo infection frequencies and vice versa.

Infection frequencies of parental isolates

All 13 parental isolates of Dg strains, Hl strains, Pn strains and Pt strains and isolates of within host-strain progeny used as inoculum showed host-specificity (Table S1, Table S2). None of these failed to infect their natural host grass, but did not infect any other host with one exception. The strain Hl.3 from H. lanatus infected one seedling of P. nemoralis (infection frequency of 11 %). Overall, infection frequencies differed between host-strains and were variable among the progeny used for inoculation. They were highest for Hl strains ranging from 40-60 %, followed by Dg strains with 39-52 %, Pn strains with 33-42 % and Pt strains with 10-35 %.
**Figure 1**: Boxplot diagrams of colony diameters of the parental strains *E. typhina* subsp. *typhina* (Dg), *E. typhina* subsp. *clarkii* (Hl), *E. typhina* subsp. *poae* (Pn) and *E. typhina* subsp. *typhina* (Pt) and their F1 progeny after 14 days of growth at 18°C (grey) and 24°C (white) on PDA. For isolate identity see Table 1.

**Infection frequencies of F1 progeny**

Inoculations of control seedlings were always negative, excluding the possibility of seed infections or contamination with other *Epichloë* strains during the experiment. When comparing infection frequencies of different types of progeny in *D. glomerata* and *H. lanatus*, between host-strain progeny performed significantly worse than within host-strain progeny (ANOVA; *p* < 0.001). Progeny from within host-strain crosses were all able to infect their respective host grass with infection frequencies being sometimes even higher than for their parental strains (Fig. 2). However, not all progeny from crosses between host-strains were capable of infecting the parental host grasses and infection frequencies were often significantly lower compared to the parental strains and the within host-strain crosses, depending on the genotype (Fig. 2).

All progeny from the within host-strain cross Dg.1_Dg.2 were able to infect *D. glomerata* and were highly variable in infection frequencies, ranging from 5 to 50 %. Furthermore, all progeny from Hl.4_Hl.3 and Hl.3_Hl.4 were able to infect *H. lanatus* with infection frequencies between 55 to 77 % and 13 to 83 %, respectively. Surprisingly, some progeny of these two crosses could also infect seedlings of *P. nemoralis* with infection frequencies ranging from 8 to 10 % and 18 to 31 %, respectively. Infection frequencies of within host-strain progeny did not differ significantly from that of parental strains except for progeny Hl.4_Hl.3 inoculated into *P. nemoralis*. This progeny showed a significantly lower infection rate compared to the parental strains (Fig. 2).
Figure 2: Infection frequencies of host-associated parental strains and their progeny from crosses within and between host-strains inoculated into Dactylis glomerata (A), Holcus lanatus (B), Poa nemoralis (C) and P. trivialis (D). Asterisks indicate significant differences between parental strains and their progeny (**p < 0.01, ***p < 0.001). For isolate identity see Table 1.

Regarding between host-strain cosses involving Dg strains and Hl strains, not all progeny from a particular cross were able to infect the host grasses. In D. glomerata, only isolate Hl.4_Dg.4_3 out of three progeny was successful with an infection frequency of 8 %. This is significantly lower compared to the parental strains (GLM Dunnett; p < 0.01; Table S2). However, all progeny from this cross were able to infect P. nemoralis with infection frequencies ranging from 6 to 46 %, whereas none of the parental strains showed positive infections in this grass species. Progeny of the within host-strain cross Dg.3_Hl.3 were able to infect both host grass species with infection frequencies of 8 to 43 % in D. glomerata and 38 % in H. lanatus. Infection frequencies of progeny from this cross were significantly lower compared to the parental strains only in H. lanatus (GLM Dunnett; p < 0.01). Positive infections in both host grass species of the parental strains were also detected for two
progeny of the between host-strain cross Dg.1_Hl.4 with an infection frequency of 4 % in *D. glomerata* and 14 to 55 % in *H. lanatus*. The third progeny of this cross was not able to infect neither *D. glomerata* nor *H. lanatus*. However, infection frequencies of this strain were significantly lower than the parental strains in both host species (GLM Dunnett; \( p < 0.001 \) and \( p < 0.001 \) for *D. glomerata* and *H. lanatus*, respectively).

In comparison, progeny of crosses between Pt strains and Dg or Hl strains were only able to infect one of their parental hosts (Fig. 2). Only one progeny out of four from cross Dg.5_Pt.1 was able to infect *D. glomerata* with an infection frequency of 10 %. This infection frequency is significantly lower than that of the parental strains (GLM Dunnett; \( p < 0.01 \)). From cross Hl.4_Pt.1 only one progeny out of four was able to infect *P. trivialis* with an infection frequency of 40 %. However, none of the progeny from crosses with Pt strains were able to infect *H. lanatus*.

**Infection stability of fungal isolates**  
The long-term stability of infected *D. glomerata*, *H. lanatus* and *P. nemoralis* was tested using microsatellite genotyping after 12 months in the greenhouse (Tab. 3). None of the initially infected *D. glomerata* and *H. lanatus* seedlings lost the infection status. In comparison to *D. glomerata* that included plants infected either with a parental strain or progeny from within or between host-strain crosses, *H. lanatus* had only plants infected with parental strains or progeny of Hl crosses. In *P. nemoralis*, infections with the parental strain were highly stable. However, the within host-strain progeny Hl.3_Hl.4_2 were only moderately stable, and the between host-strain progeny Hl.4_Dg.4_3 were unstable with all infections eventually being lost.

All infected *H. lanatus* plants formed stromata during the growing period of the experiment, whereas none of the *D. glomerata* and *P. nemoralis* plants formed stromata. The ratio of tillers with stromata and symptomless tillers differed considerably among plants, depending on the inoculated endophyte genotype (Table S3). The ratio was the smallest for the parental strain Hl.4 and the highest for the within host-strain progeny Hl.4_Hl.3_2.
Table 3: Number of stable infections of different Epichloë isolates inoculated into Dactylis glomerata, Holcus lanatus and Poa nemoralis after 12 months in the greenhouse.

<table>
<thead>
<tr>
<th>Host grass</th>
<th>Fungal isolate</th>
<th>No. of infections¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. glomerata</td>
<td>Dg.4</td>
<td>10 (10)</td>
</tr>
<tr>
<td></td>
<td>Dg.3_Hl.3_1</td>
<td>10 (10)</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Dg.2_2</td>
<td>10 (10)</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Dg.2_4</td>
<td>10 (10)</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Dg.2_5</td>
<td>8 (10)</td>
</tr>
<tr>
<td>H. lanatus</td>
<td>Hl.2</td>
<td>10 (10)</td>
</tr>
<tr>
<td></td>
<td>Hl.4</td>
<td>10 (10)</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Hl.3_2</td>
<td>10 (10)</td>
</tr>
<tr>
<td></td>
<td>Hl.3_Hl.4_4</td>
<td>4 (4)</td>
</tr>
<tr>
<td></td>
<td>Hl.3_Hl.4_1</td>
<td>9 (9)</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Hl.3_1</td>
<td>9 (9)</td>
</tr>
<tr>
<td>P. nemoralis</td>
<td>Pn.2</td>
<td>9 (9)</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Dg.4_3</td>
<td>0 (8)</td>
</tr>
<tr>
<td></td>
<td>Hl.3_Hl.4_2</td>
<td>2 (8)</td>
</tr>
</tbody>
</table>

¹Total number of screened plants is given in parentheses

Discussion

Host-specificity is hypothesised to constitute an efficient reproductive barrier in fungi, which may reduce gene flow between species even in sympathy (Giraud et al., 2008). However, only few studies have assessed the role of host-specificity for reproductive isolation in plant-associated fungi (Büker et al., 2013; Peever, 2007; Stukenbrock et al., 2011), whereas none has focused on endophytes. By means of reciprocal inoculation experiments, we observed a clear host-specificity of four members of the E. typhina complex and a reduced infectivity of experimental hybrids in the parental host grasses. These results provide evidence that reproductive isolation observed among natural Epichloë taxa may be promoted through host-specificity.

In vitro growth of fungal isolates

As expected, all strains showed a significantly reduced radial growth on PDA plates at 18°C compared to 24°C, indicating that a difference of 6°C is enough to slow down growth. The fungal strains seem to be responsive to higher temperatures, which is necessary for fast growth during emergence of stromata in spring and thus for their sexual reproduction. Furthermore, growth patterns varied among progeny from the same cross with some growing either slower or faster than their parental strains, suggesting that growth is genotype dependent. However, none of the F1 progeny tested in this study exhibited significantly reduced growth in vitro, which has been suggested as a potential intrinsic postzygotic isolation mechanism (Giraud et al., 2008). Indeed, parental strains and strains from within and between host-strain crosses did not differ significantly in their growth rate.
This conforms to a previous study that found no significant difference in radial growth between parental and hybrid isolates from several crosses (Leuchtmann and Steinebrunner, 2012). Our results therefore suggest that growth performance does not present an obvious intrinsic postzygotic isolation barrier for hybrids.

**Infection frequencies of parental isolates**

Previously, it has been assumed that *E. typhina* is able to colonize multiple hosts (Craven et al., 2001), but this study has shown that four host-associated strains of the *E. typhina* complex are clearly host-specific. Isolates could only infect the host grass from which they were isolated. An exception was one Hl strain (*E. typhina* subsp. *clarkii*) that could infect one seedling of *P. nemoralis*. In comparison to the other three host grass species that typically grow on meadows, *P. nemoralis* grows in forest habitats and on rocks, which could render cross-strain infections more difficult in nature. Furthermore, endophyte infections of *P. nemoralis* are often asymptomatic and seed-transmitted, i.e. no stromata are formed. Because of the absence of contagious ascospores allowing for horizontal transmission, a closer adaptation to a particular host may be favoured (Leuchtmann and Clay, 1997). In contrast, stroma-forming *Epichloë* spp. should have wider host ranges due to ascospores that can potentially infect new host species. This could explain why Hl strains, which are stroma-forming and not seed-transmitted, can infect *P. nemoralis* but not vice versa. A similar case was found for isolates from *Elymus villosus*, *Elymus virginicus* and *Hystrix patula* that could infect seedlings of each of these hosts plus *Brachyelytrum erectum*, whereas isolates from *B. erectum* were not capable of infecting the other grass species (Leuchtmann and Clay, 1993). However, we could not confirm the long-term stability of the Hl strain in *P. nemoralis*, because the seedling died prior to the screening. It is possible that infection would have become lost, which could explain why such infections are not found in natural habitats.

Results from previous studies using seedling inoculations indicate that there is considerable variation in the host range among different endophytes. A rather broad compatibility among hosts was shown for *Epichloë elymi* and *Epichloë coenophiala* (Christensen, 1995; Leuchtmann and Clay, 1993) and for *Epichloë bromicola* (Brem and Leuchtmann, 2003). In contrast, host-specificity was evident for *E. typhina* subsp. *typhina* infecting *D. glomerata* or *Epichloë lolii* infecting *Lolium perenne* that could only establish and maintain infections in their original host (Chung et al., 1997). Our study supports the view that narrow host ranges or even specificity to a single host may not be uncommon among *Epichloë* endophytes.

Infection frequencies of Dg strains (39-52 %) and Hl strains (40-60 %) found in this study were lower compared to a previous study (100 % and 58-93 %, respectively) (Leuchtmann and Steinebrunner, 2012). Because the seeds of the host grasses used were of the same cultivar and the fungal strains were isolated from the same host grasses, other
factors than the genotype of the grasses and fungi seem to influence the outcome of the infection process. Such factors may include the physiological stage of the cultures used or a different performance of the seedling inoculation technique. Therefore, values of infection frequencies from different studies should be compared with caution.

Infection frequencies of F1 progeny

Compared to parental strains and within host-strain progeny, infection frequencies of between host-strain progeny were significantly lower and genotype dependent (Table S1, Table S2). This is in agreement with a previous study using the same system, where the infectivity of the progeny of crosses between Dg and Hl strains was also reduced compared to that of natural strains from their original hosts (Leuchtmann and Steinebrunner, 2012). Similarly, inoculations of progeny from crosses between E. bromicola from Bromus erectus and E. bromicola from B. benekenii showed reduced compatibility with either host compared to parental strains (Brem and Leuchtmann, 2003). Overall, these findings suggest that multiple genetic determinants are involved in the compatibility of endophytes and host grasses, and indicate that a complex interplay of fungal and host grass genotypes may be responsible for host-compatibility and infection success (Brosi et al., 2011; Saikkonen et al., 2004; Tadych et al., 2014).

In our study, several Epichloë isolates originating from crosses involving D. glomerata and H. lanatus host-strains were able to infect P. nemoralis seedlings (e.g. a Hl progeny and a Hl_Dg progeny). This was surprising, because P. nemoralis has never been reported to be infected by other than E. typhina subsp. poae. However, this finding conforms to the observation that an Hl parental strain was able to infect P. nemoralis, as discussed above.

Infection stability of fungal isolates

All successful infections in D. glomerata and H. lanatus of our experiment, including parental strains and progeny from within and between host-strain crosses, were highly stable after one year. Although it remains to be seen whether stability will last for a more extended period of time and under natural conditions, it appears that compatible hybrid genotypes can form stable associations once they have established infection. However, because the inoculation procedure used was artificial involving forceful wounding of young seedlings, it is unknown whether the progeny could infect host grasses in reasonable numbers in natural habitats. By comparison, in P. nemoralis only native E. typhina subsp. poae strains were stable, whereas progeny from crosses involving other host-associated strains became lost after initial infection. Thus, it appears that hybrid progeny with no E. typhina subsp. poae genotypes in their background are not compatible with P. nemoralis in the long-term.

In the study by Chung et al. (1997), the stability of the inoculated hybrid progeny differed after eight months in the greenhouse, suggesting that infection frequency and stability are oligo- or polygenically controlled with some likely epistatic interactions.
However, we could not confirm this hypothesis on stability, because in our two host grasses the hybrid progeny were stable after 12 months. This is in agreement with another study involving endophytes of *Lolium* and *Festuca*, where for most associations stable associations between hosts and non-native endophytes could be created (Naffaa et al., 1999). However, there have also been reports where initially successful infections of seedlings inoculated with a non-native *Neotyphodium* endophyte failed to persist for more than six to eight months (Clay, 1993; Leuchtmann, 1992).

The large variation in the stroma/tiller ratio found among *H. lanatus* plants inoculated with different strains (Table S3) suggests that host- and endophyte-haplotype combinations have an impact on the reproductive success of the fungus. However, because the number of symptomless tillers and stromata were only counted for the first year and may be different in the following years, these data should be considered preliminary.

Reproductive isolation barriers

Results from a previous study showed clear genetic discontinuities among natural populations of the three host-associated endophytes of *D. glomerata*, *H. lanatus* and *P. nemoralis*, providing evidence for their reproductive isolation (Schirrmann et al., 2014). The present study indicates that reproductive barriers between these endophytes forming the *E. typhina* complex may be promoted by mechanisms of host-specificity. Because of the difficulty to separate between development and mating, mutations providing adaptation to a new habitat can affect both the fitness and the ability to mate in this habitat (Giraud et al., 2008). Host specialization can thus be sufficient to restrict gene flow, if mating occurs within each host without necessarily requiring active assortative mating (Giraud, 2006; Kohn, 2005). In this context, host-specificity has been suggested to act as an isolation barrier and to play an important role in the speciation of *Ascochyta* plant pathogens and *Mycosphaerella graminicola* (Peever, 2007; Stukenbrock et al., 2011). Thus, the evolution of host-specificity appears to be often the initial step in the speciation of fungi.

A further aspect to consider regarding potential reproductive barriers is the fitness of hybrids. While crosses involving *E. typhina* subsp. *poae* were unsuccessful due to likely strong intrinsic barriers, the other three members of the *E. typhina* complex could produce viable offspring. However, hybrids may be viable under *in vitro* conditions but unfit in the natural environment of the host, if they display intermediate traits between parental phenotypes. Our results show that the potential of between host-strain progeny to infect host grasses is generally lower than that of the host-specific parental strains due to a likely maladaptation to host plants (Giraud et al., 2008). Similar findings were made in inoculation experiments with *Microbotryum*, which showed that hybrids are less successful at infecting host plants than the progeny of intraspecific crosses (Büker et al., 2013). The inferior performance of hybrids under natural conditions could therefore act as an extrinsic postzygotic isolation barrier, preventing host grasses from becoming infected with hybrids.
However, as some of the hybrid progeny appears to be compatible with either host, we cannot rule out other possible isolation mechanisms. These may include a preference for particular host grasses of the gamete transferring Botanophila flies, assortative mating on the stromata or a low survival rate of hybrid ascospores that are sometimes formed. While no or only preliminary data are available regarding the last two hypotheses, Botanophila flies show some preference and thus may have an effect on the level of between host-strain matings (Bultman et al., 2011). Furthermore, direct competition between hybrids and non-hybrids growing simultaneously within host grasses could result in the displacement of hybrid strains, if they show inferior fitness. This hypothesis should be tested in future experiments.

Conclusions
Reciprocal inoculation experiments involving four members of the E. typhina complex, i.e. E. typhina subsp. typhina from D. glomerata, E. typhina subsp. clarkii from H. lanatus, E. typhina subsp. poae from P. nemoralis and E. typhina subsp. typhina from P. trivialis, revealed distinct patterns of host-specificity, which can be sufficient to restrict or even prevent gene flow between taxa. Furthermore, the inferior performance of experimentally generated hybrids compared to parental strains and progeny of within host-strain crosses found in planta could act as an extrinsic postzygotic isolation barrier. Our results suggest that natural populations of closely related Epichloë fungi may be genetically differentiated because of host-specificity and a maladaptation of hybrids to host grasses, which is maintaining boundaries among host-associated strains and may represent the initial step in the speciation of these fungi.

Author's contributions
MKS performed experiments, analyzed data and wrote the paper. AL designed experiments and wrote the paper.

Acknowledgements
This study was funded by the Swiss National Science Foundation (31003A_138479). Microsatellite data were generated in the Genetic Diversity Centre of ETH Zurich (GDC). We thank B. Blattmann for laboratory assistance, M. Frei for taking care of the grasses and N. Zemp for support with statistics. We thank N. Borst for helping with the growth assay, inoculation tests and preliminary analysis. Furthermore, we are grateful to M. Roumet and R. Berndt for helpful comments on the manuscript. We also thank F. Baena-Diaz, M. Scharmann and D. Brödlin for their help in the greenhouse.
References


Clark, E.M., White, J.F., Jr, Patterson, R.M., 1983. Improved histochemical techniques for the detection of *Acremonium coenophialum* in tall fescue and methods of in vitro culture of


Supplementary material

Table S1: Infection rates of progeny from crosses within and between *E. typhina* subsp. *clarkii* (Hl), *E. typhina* subsp. *typhina* (Dg) and the parental strains in *Holcus lanatus*, *Dactylis glomerata* and *Poa nemoralis* plants two months after inoculation.

<table>
<thead>
<tr>
<th>H/C/P</th>
<th>Isolate</th>
<th><em>Holcus lanatus</em></th>
<th><em>Dactylis glomerata</em></th>
<th><em>Poa nemoralis</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of plants</td>
<td>Infection in %</td>
<td>Number of plants</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Infected</td>
<td>Total</td>
</tr>
<tr>
<td>H</td>
<td>Hl.4_Dg.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hl.4_Dg.4_2</td>
<td>20</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Dg.4_3</td>
<td>11</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Dg.4_5</td>
<td>12</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>H</td>
<td>Dg.3_Hl.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dg.3_Hl.3_1</td>
<td>13</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Dg.3_Hl.3_2</td>
<td>15</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>H</td>
<td>Dg.1_Hl.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dg.1_Hl.4_1</td>
<td>21</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Hl.4_2</td>
<td>11</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Hl.4_3</td>
<td>12</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>C</td>
<td>Hl.4_Hl.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hl.4_Hl.3_1</td>
<td>24</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Hl.3_2</td>
<td>12</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Hl.3_3</td>
<td>11</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Hl.3_4</td>
<td>13</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>C</td>
<td>Hl.3_Hl.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hl.3_Hl.4_1</td>
<td>12</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Hl.3_Hl.4_2</td>
<td>19</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Hl.3_Hl.4_4</td>
<td>12</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>H/C/P</td>
<td>Isolate</td>
<td>Holcus lanatus</td>
<td></td>
<td>Dactylis glomerata</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>---------------</td>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of plants</td>
<td>Infection in %</td>
<td>Number of plants</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Infected</td>
<td>Total</td>
</tr>
<tr>
<td>C</td>
<td>Dg.1_Dg.2</td>
<td>24</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Dg.2_1</td>
<td>12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Dg.2_2</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Dg.2_4</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dg.1_Dg.2_5</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>Hl.1</td>
<td>25</td>
<td>15</td>
<td>60</td>
</tr>
<tr>
<td>P</td>
<td>Hl.2</td>
<td>58</td>
<td>35</td>
<td>60</td>
</tr>
<tr>
<td>P</td>
<td>Hl.3</td>
<td>8</td>
<td>4</td>
<td>50</td>
</tr>
<tr>
<td>P</td>
<td>Hl.4</td>
<td>11</td>
<td>6</td>
<td>55</td>
</tr>
<tr>
<td>P</td>
<td>Dg.1</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>Dg.2</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>Dg.3</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>Dg.4</td>
<td>31</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>Pn.1</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>Pn.2</td>
<td>12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>76</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

1 Abbreviations are used to indicate parental (P) or hybrid (H) strain, or intraspecific cross (C).
2 The following abbreviations are used to indicate the taxonomic identity of fungi: Hl = E. typhina subsp. clarkii infecting Holcus lanatus; Dg = E. typhina subsp. typhina infecting Dactylis glomerata; Pn = E. typhina subsp. poae infecting Poa nemoralis.
3 The total number of inoculated plants that have survived and were transferred to multi-pot trays three weeks after inoculation. Infection rates were determined two months after inoculation.
Table S2: Infection rates of progeny from crosses within and between *E. typhina* subsp. *typhina* (Pt) and *E. typhina* subsp. *clarkii* (Hl) or *E. typhina* subsp. *typhina* (Dg) and the parental strains in *Holcus lanatus*, *Dactylis glomerata* and *Poa trivialis* plants two months after inoculation.

<table>
<thead>
<tr>
<th>P/H</th>
<th>Isolate</th>
<th><em>Holcus lanatus</em></th>
<th><em>Dactylis glomerata</em></th>
<th><em>Poa trivialis</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of plants</td>
<td>Infection in %</td>
<td>Number of plants</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Infected</td>
<td>Total</td>
</tr>
<tr>
<td>H</td>
<td>Dg.5_Pt.1</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dg.5_Pt.1_2</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dg.5_Pt.1_3</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dg.5_Pt.1_4</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dg.5_Pt.1_4</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>H</td>
<td>Hl.4_Pt.1</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Pt.1_1</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Pt.1_2</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Pt.1_3</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Hl.4_Pt.1_4</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>Pt.1</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pt.2</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>Hl.4</td>
<td>10</td>
<td>4</td>
<td>40</td>
</tr>
<tr>
<td>P</td>
<td>Dg.5</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

1 Abbreviations are used to indicate parental (P) or hybrid (H) strain.

2 The following abbreviations are used to indicate taxonomic identity of fungi: Hl = *E. typhina* subsp. *clarkii* infecting *Holcus lanatus*; Dg = *E. typhina* subsp. *typhina* infecting *Dactylis glomerata*; Pt = *E. typhina* subsp. *typhina* infecting *Poa trivialis*.

3 The total number of inoculated plants that have survived and were transferred to multi-pot trays 3 three weeks after inoculation. Infection rates were determined two months after inoculation.
**Table S3:** Average number of stromata, symptomless tillers and stroma/tiller ratio of *Holcus lanatus* infected with different *Epichloë* parental isolates and progeny from crosses within host-strains after 12 months in the greenhouse.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>No. of stromata</th>
<th>No. of tillers</th>
<th>Ratio (stroma/tiller)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hl.2</td>
<td>31.875</td>
<td>59.5</td>
<td>0.574</td>
</tr>
<tr>
<td>Hl.4</td>
<td>0.875</td>
<td>76.25</td>
<td>0.013</td>
</tr>
<tr>
<td>Hl.4_Hl.3_2</td>
<td>38</td>
<td>60.5</td>
<td>0.617</td>
</tr>
<tr>
<td>Hl.4_Hl.3_3</td>
<td>12.25</td>
<td>70.875</td>
<td>0.171</td>
</tr>
<tr>
<td>Hl.3_Hl.4_1</td>
<td>19.375</td>
<td>58.5</td>
<td>0.371</td>
</tr>
<tr>
<td>Hl.3_Hl.4_4</td>
<td>24.375</td>
<td>49.125</td>
<td>0.502</td>
</tr>
<tr>
<td>Control</td>
<td>0</td>
<td>97.875</td>
<td>0.000</td>
</tr>
</tbody>
</table>
CHAPTER 3

REPRODUCTIVE ISOLATION BETWEEN GRASS-ASSOCIATED ENDOPHYTES – NEW INSIGHTS FROM MATING TESTS AND GENOME SCANS

Unpublished manuscript

ETH Zurich, Institute of Integrative Biology, Plant Ecological Genetics, Zürich, Switzerland
ETH Zurich, Genetic Diversity Centre (GDC), Zürich, Switzerland
ETH Zurich, Institute of Integrative Biology, Plant Pathology, Zürich, Switzerland
**Abstract**

Adaptation to different hosts is common among closely related taxa of plant-associated fungi, and is often interpreted as an underlying force driving the divergence of lineages, particularly in pathogenic fungi. To date, little is known about the genetics of host-specificity in endophytes, which are generally considered to establish mutualistic interactions with the infected host. *Epichloë* taxa constitute an ideal system to investigate the underlying mechanisms of host-adaptation in endophytes, as they form cryptic species complexes including several host-specific taxa. Here, we addressed the nature of reproductive barriers that maintain divergence between host-specific taxa in two sympatric subspecies of the *E. typhina* species complex (i.e. *E. typhina* subsp. *typhina* infecting *Dactylis glomerata* and *E. typhina* subsp. *clarkii* infecting *Holcus lanatus*) by combining evidence from a mating experiment with genomic analyses to detect candidate loci under selection. Our results indicate that strong prezygotic barriers hinder the formation of hybrids between the two subspecies, and assortative mating may contributes to lineage divergence. Moreover, genome scans revealed that genes showing the strongest signal of selection encode for secreted proteins that may be involved in the infection phase and stroma formation of the endophytes. These pathogenic stages of the endophyte require a close adaptation to the host in order to avoid defense reactions. The secreted proteins may therefore be involved in host adaptation and reproductive isolation observed between host-specific *Epichloë* taxa.

**Introduction**

Endophytic fungi and their host plants may establish mutualistic interactions, in which endophytes provide direct defense against natural enemies of the host, e.g. pathogenic organisms or herbivores (Caroll, 1992; Clay, 1988; Faeth, 2002). In return, the endophytes profit by receiving photosynthetically derived carbohydrates from the host. With successful infection of the host plant a close relationship mediated by multiple signalling systems is initiated (Shaw, 2014), including the interaction with the defense mechanisms of the host. This relationship may arise through the coevolution of reciprocal life history traits driven by the response to selective forces that make the endophyte an obligate participant of the plant life cycle. Moreover, in outcrossing plant systems the endophyte is confronted continuously with new biochemical compounds originating from genetic recombination associated with sexual reproduction of the host. The survival of the fungus therefore may depend on life strategies that allow to increase its genetic diversity and adaptive potential, e.g. via the establishment of a phase of sexual reproduction (Peters and Lively, 1999; Saikkonen et al., 2004).

In the most common evolutionary model describing plant-fungi interactions, recurrent changes in the biochemical structure of the interacting organisms are expected to trigger bouts of selection of genetic variants. These genetic variants enable reciprocal adaptation to the biochemical environment, what is defined as an ‘arms race’ between the two organisms.
infecting *Epichloë typhina* genetically distinct even in sympatry. (Leuchtmann and Schardl, 1998; Schardl et al., 2007; White, 1993) demonstrated host stromata from differential mating types (mat) in opposite mating types. The flies are largely unspecific in their visiting behaviour and therefore spermata from different host grasses are collected and actively transported in mixtures among the stromata (Bultman and Leuchtmann, 2009), potentially allowing for hybridization between host-associated taxa. However, although taxa within the same species complex have been demonstrated to be able to hybridize in artificial experiments of cross-fertilization (Leuchtmann and Scharld, 1998; Scharld et al., 2007; White, 1993), many appear to be genetically distinct even in sympatry. One particular example includes the host-associated *Epichloë typhina* subsp. *typhina* infecting *Dactylis glomerata* and *E. typhina* subsp. *clarkii* infecting *Holcus lanatus*. Because of apparent sexual compatibility, the two taxa have been

Sexually reproducing *Epichloë* endophytes of grasses (Poaceae) (Leuchtmann, 1992) provide an ideal system to study adaptation to different hosts, because they include species complexes of taxa forming host-specific associations, which mediates reproductive isolation. During the asexual life cycle, *Epichloë* fungi grow asymptomatically in plant apoplasts primarily in shoot apical meristems and leaf sheaths (Schardl et al., 2004) with little or no negative effects on the vegetative tissues of their hosts. However, during the sexual life cycle they form external reproductive structures, i.e. stromata, which emerge on flowering shoots (Clay and Schardl, 2002). Because stromata prevent the development of host inflorescence and seed by surrounding the flag-leaf sheath, this antagonistic stage is called ‘choke disease’ (Western and Cavett, 1959). Haploid spermata (male gametes) and female receptive hyphae are simultaneously formed on the young emerging stromata (White and Bultman, 1987). In natural ecosystems, *Botanophilia* flies (Anthomyiidae) transfer spermata between stromata of opposite mating types (mat-1 and mat-2) in a process similar to pollination (Bultman et al., 1998). The flies are largely unspecific in their visiting behaviour and therefore spermata from different host grasses are collected and actively transported in mixtures among the stromata (Bultman and Leuchtmann, 2009), potentially allowing for hybridization between host-associated taxa. However, although taxa within the same species complex have been demonstrated to be able to hybridize in artificial experiments of cross-fertilization (Leuchtmann and Schardl, 1998; Schardl et al., 2007; White, 1993), many appear to be genetically distinct even in sympatry. One particular example includes the host-associated *Epichloë typhina* subsp. *typhina* infecting *Dactylis glomerata* and *E. typhina* subsp. *clarkii* infecting *Holcus lanatus*. Because of apparent sexual compatibility, the two taxa have been
referred to the same species within the *E. typhina* complex (Craven et al., 2001; Leuchtmann et al., 2014). Although the two taxa can be distinguished based on morphological characters, e.g. the size and disarticulation patterns of the ascospores, and different hosts (White, 1993), phylograms based on *β*-tubulin, translation elongation factor 1-α (*tefA*) and actin (*act1*) show that *E. typhina* subsp. *clarkii* is nested within *E. typhina* subsp. *typhina*, indicating a very close relationship between the taxa (Craven et al., 2001; Leuchtmann et al., 2014). The two host-associated taxa occur sympatrically in Switzerland at the site of Aubonne. Previous investigations focusing on the populations at this site provided evidence of established reproductive barriers involved in the maintenance of subspecies divergence. Hybrid ascospores were found in small proportions of less than 10 % in field surveys (Bultman et al., 2011), and molecular analyses based on 15 microsatellite markers did not detect genotypes of recent hybrid origin in host plants (Schirrmann et al., 2015). The rarity of hybrid ascospores may underlie genetic incompatibilities that cause intrinsic reproductive isolation during sexual reproduction (Seehausen et al., 2014). Furthermore, reciprocal infections with several host-associated strains, including *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*, showed distinct patterns of host-specificity and lower infection frequencies of hybrid progeny compared to the parental strains due to a likely maladaptation to host plants (Schirrmann and Leuchtmann, 2015). This evidence suggests that mechanisms of reproductive isolation keep natural populations of *E. typhina* in different hosts genetically distinct.

In this study, we investigated mechanisms underlying host-specific adaptation in *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* to better understand the components involved in reproductive barriers between the two subspecies. As *Botanophila* flies usually show no preferences for one or the other host-associated taxa and carry a mixture of spermata, tests for spermata competition on the stroma were performed to investigate potential intrinsic barriers, which may hinder the formation of hybrids and therefore contribute to the maintenance of subspecies divergence. Furthermore, genome scans were performed on a population sampling of *Epichloë* from *D. glomerata* and *H. lanatus* occurring in sympatry at the Aubonne site to identify candidate genes responsible for host-specificity driving the divergence of the two systems. Our results showed that strongly diverged genes under positive selection are mainly associated with functions of endophyte-grass interactions, and secreted proteins are hypothesised to play a major role in the establishment of reproductive barriers.

**Material and methods**

**Mating experiment**

Fungal isolates were obtained from *Dactylis glomerata* infected with *E. typhina* subsp. *typhina* and *Holcus lanatus* infected with *E. typhina* subsp. *clarkii* collected in Switzerland as previously described (Schirrmann and Leuchtmann, 2015) (Table S1). Cultures were
established from a small agar stripe (0.5 mm x 2 mm) removed from the margin of a 14 days old colony on potato dextrose agar (PDA; BD Company, Sparks, Maryland) (Leuchtmann, 1994). These were homogenized with a pestle in an Eppendorf tube and evenly spread on a new PDA plate. Conidia (homologous to spermatia from stromata) were harvested from seven days old colonies in 2 ml distilled water by scraping the colony surface with a sterilized spreader rod. After counting the number of conidia of each fungal strain in a Neubauer counting chamber, the suspensions of *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* were adjusted to 48 conidia per microliter and mixed in equal concentrations, in 1:10 ratios and in 1:100 ratios. Finally, 5 μl of the conidia suspensions were applied with a pipette in the middle of each susceptible, freshly emerged stroma on *D. glomerata*. Conidia of only host-compatible *E. typhina* subsp. *typhina*, or only host-incompatible *E. typhina* subsp. *clarkii* served as controls. Each conidia suspension in different ratios was replicated 20 times, including six different genotype combinations. However, because only a limited number of stromata were available and less parental strains of *E. typhina* subsp. *typhina* than *E. typhina* subsp. *clarkii* were used, control matings with conidia from *E. typhina* subsp. *clarkii* were replicated only 10 times. Furthermore, three *D. glomerata* plant genotypes each infected with a different genotype of *E. typhina* subsp. *typhina* were used for the experiment and cloned five times, resulting in 15 experimental plants in total. After fertilization of the stromata, the host grasses were grown in growth-chambers at 21°C 16 h day (30 klux) and 16°C 8 h night for five weeks. To exclude contamination from another source than the applied suspensions, stromata were surrounded by plastic tubes. The reproductive success of each stroma was assessed five weeks later when perithecia with ascospores developed. Stroma fragments from fertile matings were placed on the lids of overturned petri dishes containing PDA, on which ascospores were ejected. An agar block of about 1 mm² carrying ejected ascospores was then transferred to new PDA plates for further growth. Eventually, DNA was extracted from each fungal colony originating from a bulk of germinated ascospores and genotyped. Protocols followed the methodology previously published in Schirrmann et al. (2015). Briefly, DNA was extracted using the cetyl-trimethyl ammonium bromide (CTAB) standard protocol (Doyle and Doyle, 1987) and genotyped with 15 previously developed microsatellite markers (Schirrmann et al., 2015). All PCR products were run with GeneScan-500 LIZ as size standard on a 3130 Genetic Analyzer (Applied Biosystems, Foster City, California), and electropherograms were analyzed using Geneious 6.1.7 (Drummond et al., 2013).

*Genome-wide re-sequencing of host-associated Epichloë*

*Illumina library preparation for Pool sequencing*

Twenty stromata of *E. typhina* subsp. *typhina* from *D. glomerata* plants and 20 stromata of *E. typhina* subsp. *clarkii* from *H. lanatus* plants were sampled in spring 2013 from the sympatric populations at the natural grassland site in Aubonne, Canton Vaud, Switzerland.
Genomic DNA of each individual was extracted with the DNeasy Plant Kit (Quiagen) from each stroma using dried fungal material from the interior part of the stroma split open under sterile conditions. DNA quality was checked on 1.5 % agarose gels stained with GelRed on a UV-Vis Spectrometer and DNA quantity was measured with a Qubit fluorometer (dsDNA BR). Whole-genome resequencing data were obtained using a pool sequencing (Pool-Seq) approach (Fischer et al., 2013; Schlötterer et al., 2014). Equal amounts of high-quality DNA from 20 stromata from each of the two host-associated taxa were pooled, thus producing two samples of 5 μg RNA-free genomic DNA representing the fungal populations. Illumina libraries of ~300 bp insert size were generated following the instructions of the Illumina Paired-End Sample Preparation Kit. Sequencing was performed on an Illumina MiSeq lane using 150 bp paired-end reads.

Read mapping and SNP calling
Adapters of forward and reverse reads were quality trimmed using Trimmomatic with a sliding window size of 4 bp and a required quality of 15 (Bolger et al., 2014). Bases at the start and end of a read were cut off below a threshold of 5, whereas the read was dropped below a length of 50 bp. The available genome assembly of *E. typhina* subsp. *poeae* (E5819) (Schardl et al., 2013) was used as reference to map the reads produced in this study. *Epichloë typhina* subsp. *poeae* is a close relative of the studied taxa within the same species complex (Leuchtmann and Schardl, 1998). The *E. typhina* subsp. *poeae* reference genome included 34 Mb assembled in 2072 contigs, with N50 of 36475 bp (Schardl et al., 2013). In total, 15.2 Mb of the genome consisted of genic regions, whereas 10.5 Mb (25.1 %) were coding sequences.

Reads were mapped with BWA-MEM using the default settings (H. Li and Durbin, 2009). Alignments were filtered for a minimum mapping quality of 20 before the remaining high quality reads were sorted and indexed with Samtools v. 0.1.18 (http://samtools.sourceforge.net/). Single nucleotide polymorphisms (SNPs) in the two host-associated taxa were called with Samtools (mpileup; Li et al., 2009), using the default settings and population statistic measures were computed using software specifically developed for Pool-Seq data.

Population genomic analyses
All population genomic analyses were done gene-wise, because reads of both host-associated taxa mainly mapped to coding regions. Genetic differentiation was characterized by estimating the pairwise fixation index $F_{ST}$ according to (Hartl and Clark, 2007) to identify genes with elevated genetic differentiation between host-associated taxa using the program Popoolation2. Synchronisation and filtering of the mpileup-file were performed using the java script mpileup2sync.jar (Kofler et al., 2011b). Subsequently, the structural annotation of *E. typhina* subsp. *poeae* E5819 (Schardl et al., 2013) was transferred onto the synchronized
file using create-gene-wise-sync.pl (Kofler et al., 2011b) to generate a gene-based dataset, which enables the calculation of population statistics for whole genes. Genes with elevated genetic differentiation were identified using fst-sliding.pl. The minimum count of the minor allele was set to 2 to account for sequencing errors, and minimum coverage of 6 and maximum coverage of 60 per population were used as thresholds to minimise the risk of calling variants within repeated sequences. Pool size per population was set to 20 and the window size to 40,000 bp, this value being longer than the length of any gene as recommended by the authors of the program (Kofler et al., 2011b). Within each gene, at least 50 % of SNPs had to fulfil the minimum coverage of 6 and maximum coverage of 60 per taxon. Because the minimum covered fraction in PoPoolation2 is calculated by using the size of the window, it is heavily biased towards very low values. This parameter was therefore adjusted by only considering the fraction of the gene that fulfils the coverage thresholds and dividing it through the total length of the gene, giving more precise estimates about the minimum covered fraction of a gene. The set of genes containing highly differentiated SNPs above the 95 % $F_{ST}$ quantile were considered significant outliers.

The absolute differentiation statistics $D_{XY}$ was calculated as the DNA sequence divergence for each SNP based on allele frequencies, summed across all SNPs in each gene, respectively, and divided by the total base pairs in the gene (Smith and Kronforst, 2013). SNP allele frequencies were estimated with snp-frequency-diff.pl filtering for a minimum quality of 20. Calculations of $D_{XY}$ were only performed for bi-allelic loci (~99.2 % of the total allele estimates).

Gene-wise genetic diversity $\pi$ (Nei and Li, 1979) and Tajima’s $D$ (Tajima, 1989) within host-associated taxa were estimated using variance-at-position.pl in PoPoolation (Kofler et al., 2011a). Minimum requirements for coverage and allele count used in SNP calling were set as described above for the $F_{ST}$ calculations. All statistical tests mentioned below were computed in R version 2.13.0 (R Development Core Team 2011).

We compared ratios of non-synonymous (N) to synonymous (S) changes ($dN/dS$) of genes showing a consistent signal of selection in $F_{ST}$ and $D_{XY}$ analyses ($F_{ST}$ - $D_{XY}$ outliers) to genes estimated to be outliers only in the $F_{ST}$ analyses. Given the high measures of $F_{ST}$ for both sets of genes, the allelic differences can be considered to represent fixation events and therefore the $dN/dS$ test can be applied as originally conceived for independently diverging lineages (Kryazhimskiy and Plotkin, 2008). For each host-associated taxon, we computed the 90 % majority sequence and estimated $dN/dS$ using the Nei and Gojobori method (Nei and Gojobori, 1986) as implemented in the PAML software package (Yang, 2007). We furthermore computed the McDonald-Kreitman test to compare polymorphic synonymous and non-synonymous within a population with fixed synonymous and non-synonymous substitutions between populations (Egea et al., 2008). Input data for this test were obtained by computing the 25 % minority and 90 % majority sequences for E. typhina subsp. typhina and E. typhina subsp. clarkii.
Gene functions
Annotations were obtained with Blast2Go by blasting genes with a homology search against a local installation of the NCBI nr database (Conesa et al., 2005). In addition, protein domains detection was performed with a local installation of interproscan version 5RC7 (Jones et al., 2014). To explore in which biological processes genes within $F_{ST}$-$D_{ST}$ outliers are involved, we performed Gene Ontology (GO) enrichment analyses using Blast2Go (Conesa and Götz, 2008). Significance for each individual GO category was computed with Fisher’s exact test and a significance threshold of 1 %. The E5819 E. typhina subsp. poae genes that fulfilled the coverage thresholds in the $F_{ST}$ analyses were used as a background reference. Furthermore, signal peptides within both sets of outliers were predicted using the software SignalP (Petersen et al., 2011).

Results
Assortative mating among closely related endophytes
We mimicked the fertilization of stromata through Botanophila flies by using conidia suspensions of E. typhina subsp. typhina and E. typhina subsp. clarkii in different ratios to test for the potential of forming hybrids on the stromata (Table S2). All stromata that were fertilized with only conidia from E. typhina subsp. typhina produced viable ascospore progeny (100 %, Fig. 1). Genotyping of the progeny recovered the expected parental alleles, but also other alleles of possibly contaminating strains growing on conspecific grasses present in the growth-chamber in various amounts, ranging from 20 to 80 % (Fig. 1). In contrast, much fewer positive matings were found among stromata that were fertilized with only conidia of E. typhina subsp. clarkii or different ratios of conidia from E. typhina subsp. typhina and E. typhina subsp. clarkii, ranging from 40 to 70 % (Fig. 1). In negative matings either the stromata dried out shortly after fertilization or did not produce any viable ascospores. When conidia from only E. typhina subsp. clarkii were applied, in about 20 % of the positive matings (hybrids between host-associated strains) the expected paternal parent could be confirmed, whereas in 80 % we found offspring between matings of E. typhina subsp. typhina (Fig. 1). The amount of hybrids is significantly lower compared to matings within E. typhina subsp. typhina (ANOVA; $p < 0.001$). The same pattern could be observed when equal concentrations or 1:10 ratios of E. typhina subsp. typhina and E. typhina subsp. clarkii were applied. In both cases, significantly fewer positive matings, 7 % and 10 % respectively, were identified as hybrids, whereas 93 % and 90 % were crosses between different E. typhina subsp. typhina genotypes (ANOVA; $p < 0.001$; Fig. 1). In contrast, when conidia were applied in 1:100 ratios, we could detect only crosses between E. typhina subsp. typhina genotypes and no hybrids (Fig. 1).
Fig. 1: Percentage of Epichloë stromata growing on D. glomerata that produced viable ascospores after fertilization. Stromata were fertilized with conidia of only E. typhina subsp. typhina, with conidia in 1:1 ratios, 1:10 ratios or 1:100 of E. typhina subsp. typhina and E. typhina subsp. clarkii, and with conidia of only E. typhina subsp. clarkii. The dark grey portions of the bar show the amount of crosses between E. typhina subsp. typhina and E. typhina subsp. clarkii (i.e. hybrids), the light grey portions show the amount of contamination with E. typhina subsp. typhina, and the white portions show the amount of progeny having the genetic background of the applied E. typhina subsp. typhina. Asterisks indicate significant differences between hybrids and crosses within host-associated strains (**p < 0.001). See Table S2 for the actual numbers and percentages.

**Genome re-sequencing and read mapping**

Sequencing of the two population pools resulted in 69 million paired-end reads, corresponding to 8.5 Gb of sequence data. After adapter and quality trimming, we retained 67 million paired-end reads (97.8 %) with a median Phred-score of 38 (mean = 33.7 million paired-end reads per pool; range: 30–37). Approximately 17 % and 21 % of all reads of E. typhina subsp. typhina and E. typhina subsp. clarkii mapped to the reference genome of E. typhina subsp. poae, and respectively to 1911 and 1875 contigs. A large proportion of the remaining unmapped reads is likely constituted by plant genomic material. Mapping against the genome of the grass species Brachypodium distachyon strain Bd21-1 downloaded from the Brachypodium Genome Database (http://www. brachypodium.org) revealed that 35 % of the reads could be aligned. Another proportion of unmapped reads likely included fungal genomic elements that are too divergent from the reference genome. In fact, we observed an increased proportion of coverage in coding regions in comparison to more variable portions of the genome. Actual median coverage for annotated genes was 41x for E. typhina subsp. typhina and 23x E. typhina subsp. clarkii, whereas genome-wide coverage was 32x and 20x.
Population genomic analyses

Between the two populations studied, we identified 647,099 polymorphic SNPs at the genome-wide level. Considering only SNPs within the 8,206 mapped genes, we identified 177,421 SNPs in total that fulfilled the coverage thresholds, with an average of 21.9 SNPs per gene. Pairwise $F_{ST}$ among host-associated strains based on SNPs within genes was very heterogeneous across the genome with an average of 0.53. The shape of the frequency of gene-wise $F_{ST}$ values shows a bimodal distribution with a low peak at an $F_{ST}$ value of about 0.1 and a high peak at an $F_{ST}$ value of about 0.6 (Fig. 2). In genes above the 95% quantile of $F_{ST}$, corresponding to an $F_{ST}$ threshold of 0.843, we identified 5,344 strongly differentiated SNPs located within 410 genes (with an average number of 13.1 SNPs per gene). This conforms to 0.83% of all SNPs in the data set.

![Figure 2: The shape of the frequency distributions of gene-wise $F_{ST}$ values between E. typhina subsp. typhina and E. typhina subsp. clarkii. The threshold for the positive $F_{ST}$ outliers (0.843) is shown with a dashed vertical line.](image)

We further explored the genomic dataset to search for complementary evidence of selection as inferred by methods alternative to $F_{ST}$ scans to pinpoint major candidate genes that show a consistent signal in multiple statistics. For this reason, we took the overlap between $F_{ST}$ and $D_{XY}$ outlier genes to more reliably identify regions of locally reduced gene flow mediated by selection (Cruickshank and Hahn, 2014). The mean of the absolute divergence measure $D_{XY}$ was significantly higher in $F_{ST}$ outlier genes compared to the rest of the genome, i.e. 0.0119 compared to 0.0074 (two-tailed Wilcoxon rank-sum test; $p < 0.001$), and the set of genes above the 95% quantile of $D_{XY}$, corresponding to a threshold of 0.01996, were considered significant outliers. Fifty-seven genes are included among the overlap of $F_{ST}$ and $D_{XY}$ outlier genes ($F_{ST}$-$D_{XY}$ outliers), and they are considered as forming a set of loci showing the strongest signal of selection (Fig. 3).
**Figure 3:** $F_{ST}$ values plotted against $D_{XY}$ values of all mapped genes in the genome. The horizontal line represents the threshold for the 5% quantile $F_{ST}$ outliers and the vertical line the threshold for the 5% quantile $D_{XY}$ outliers. The overlap between $F_{ST}$ and $D_{XY}$ outliers is shown in red in the rectangle in the upper right.

The mean of the nucleotide diversity $\pi$ was small in both populations with 0.0038 in *E. typhina* subsp. *typhina* and 0.0017 in *E. typhina* subsp. *clarkii*, but about 40% higher in *E. typhina* subsp. *typhina* (Table 1). This supports the observation that the assumed population size of *E. typhina* subsp. *typhina* at the field site in Aubonne is higher than the population size of *E. typhina* subsp. *clarkii*, enhancing a higher intraspecific diversity (Nielsen and Slatkin, 2013). Furthermore, in *E. typhina* subsp. *clarkii* $\pi$ was significantly reduced in genes within $F_{ST}$-$D_{XY}$ outliers compared to the rest of the genome (two-tailed Wilcoxon rank-sum test; $p < 0.05$), whereas in *E. typhina* subsp. *typhina* this difference was not significant (Table 1). Mean values of Tajima’s $D$ (-0.514 and -0.306 in *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*, respectively) were slightly negative in both host-associated taxa, but only in *E. typhina* subsp. *typhina* Tajima’s $D$ was significantly reduced in genes within $F_{ST}$-$D_{XY}$ outliers compared to the rest of the genome (two-tailed Wilcoxon rank-sum test; $p < 0.01$; Table 1).

<table>
<thead>
<tr>
<th>Population</th>
<th>$F_{ST}$-$D_{XY}$ outliers</th>
<th>Non-outliers</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$\pi$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td>0.0034</td>
<td>0.0038</td>
<td>0.502</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td>0.0009</td>
<td>0.0017</td>
<td>0.022*</td>
</tr>
<tr>
<td><strong>Tajima’s $D$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td>-0.865</td>
<td>-0.514</td>
<td>0.006**</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td>-0.236</td>
<td>-0.306</td>
<td>0.367</td>
</tr>
</tbody>
</table>

* $p < 0.01$; * $p < 0.5$
**Gene Ontology enrichment analyses**

Using the set of $F_{ST-D_{XY}}$ outlier genes, we conducted Gene Ontology (GO) enrichment analyses. We classified the aligned genes using Blast2Go (Conesa and Götz, 2008). GO categories could be assigned to 5669 (71 %) of all genes aligned between *E. typhina* subsp. *typhina*, *E. typhina* subsp. *clarkii* and the reference genome of *E. typhina* subsp. *poae*. Among the 57 $F_{ST-D_{XY}}$ outliers, GO categories could be assigned to 22 genes (39 %) with 19 categories being significantly overrepresented ($p < 0.001$; Table 2). These 19 categories could be assigned to three main biological processes, i.e. ‘modification of the cell wall’ (GO:0042545), ‘secretion of proteins’ (GO:0009306) and ‘catabolic processes of xylan’ (GO:0045493), which is a group of hemicelluloses found in plant cell walls and grasses (Bastawde, 1992). Using a false discovery rate (FDR) of 5 %, these GO terms were however not significant, possibly because of a too low sample size of $F_{ST-D_{XY}}$ outlier genes (Pawitan et al., 2005).

**Table 2:** The 19 enriched GO terms that were significantly overrepresented ($p < 0.01$) among $F_{ST-D_{XY}}$ outliers in *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*.

<table>
<thead>
<tr>
<th>GO ID</th>
<th>GO term</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0071554</td>
<td>cell wall organization or biogenesis</td>
<td>1.80E-04</td>
</tr>
<tr>
<td>GO:0044036</td>
<td>cell wall macromolecule metabolic process</td>
<td>8.92E-04</td>
</tr>
<tr>
<td>GO:0005618</td>
<td>cell wall</td>
<td>1.82E-03</td>
</tr>
<tr>
<td>GO:0030312</td>
<td>external encapsulating structure</td>
<td>2.04E-03</td>
</tr>
<tr>
<td>GO:0004713</td>
<td>protein tyrosine kinase activity</td>
<td>2.78E-03</td>
</tr>
<tr>
<td>GO:0031176</td>
<td>endo-1,4-beta-xylanase activity</td>
<td>3.81E-03</td>
</tr>
<tr>
<td>GO:0019028</td>
<td>viral capsid</td>
<td>3.81E-03</td>
</tr>
<tr>
<td>GO:0019013</td>
<td>viral nucleocapsid</td>
<td>3.81E-03</td>
</tr>
<tr>
<td>GO:0030599</td>
<td>pectinesterase activity</td>
<td>3.81E-03</td>
</tr>
<tr>
<td>GO:0042545</td>
<td>cell wall modification</td>
<td>3.81E-03</td>
</tr>
<tr>
<td>GO:0006807</td>
<td>nitrogen compound metabolic process</td>
<td>3.89E-03</td>
</tr>
<tr>
<td>GO:0005976</td>
<td>polysaccharide metabolic process</td>
<td>4.92E-03</td>
</tr>
<tr>
<td>GO:1901e360</td>
<td>organic cyclic compound metabolic process</td>
<td>7.22E-03</td>
</tr>
<tr>
<td>GO:0045493</td>
<td>xylan catabolic process</td>
<td>7.61E-03</td>
</tr>
<tr>
<td>GO:0045491</td>
<td>xylan metabolic process</td>
<td>7.61E-03</td>
</tr>
<tr>
<td>GO:0009306</td>
<td>protein secretion</td>
<td>7.61E-03</td>
</tr>
<tr>
<td>GO:0004114</td>
<td>3’,5’-cyclic-nucleotide phosphodiesterase activity</td>
<td>7.61E-03</td>
</tr>
<tr>
<td>GO:0004112</td>
<td>cyclic-nucleotide phosphodiesterase activity</td>
<td>7.61E-03</td>
</tr>
<tr>
<td>GO:0010410</td>
<td>hemicellulose metabolic process</td>
<td>7.61E-03</td>
</tr>
</tbody>
</table>
Putative secreted proteins

The GO-term enrichment analyses indicated that genes involved in the secretion of proteins formed a significant category showing a consistent signal of selection between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*. Differences in secreted proteins have been shown to be key to determine adaptation to different hosts, because they can suppress plant defense responses and modulate plant physiology to accommodate fungal invaders (Presti et al., 2015; Stukenbrock et al., 2010). Therefore, we further explored the genes in our dataset related to these functions.

In total, 358 genes out of 8206 mapped genes were predicted to contain N-terminal signal peptides. Gene functions of secreted proteins within both outlier sets, i.e. $F_{ST}$ outliers and $F_{ST-D_{XY}}$ outliers, are listed in Table S3. Within $F_{ST-D_{XY}}$ outliers five genes were encoding for putative secreted proteins, of which two had no functional annotations. Genes with known function were homologous to genes essential for phytopathogenicity in other fungal species. Two proteins showed similarity to enzymes involved in the degradation of cell walls (Eaton et al., 2015), i.e. ‘carbohydrate esterase family 8 protein’ and ‘endo-1,4-beta xylanase’. One gene encoded for a ‘chloroperoxidase’, an enzyme that was shown to have key functions in the suppression of host defense (Rudd et al., 2015). Genes encoding for the secreted proteins had high $dN/dS$ ratios ranging from 1.337 to 2.173 with an average of 1.663, which might be indicative for positive selection (Table 3; detailed information in Table S3). In comparison, genes occurring only in $F_{ST}$ outliers had $dN/dS$ ratios ranging from 0.047 to 0.711 with an average of 0.354 (Table 3; detailed information in Table S3).

**Table 3**: Average $dN/dS$ ratio and results of the multi-locus McDonald-Kreitman test between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* for genes encoding for putative secreted proteins within the $F_{ST}$ and $F_{ST-D_{XY}}$ outliers.

<table>
<thead>
<tr>
<th>Outlier</th>
<th>$dN/dS$</th>
<th>$\alpha$</th>
<th>$\omega_{MH}$</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{ST-D_{XY}}$</td>
<td>1.663</td>
<td>0.614</td>
<td>0.35</td>
<td>0.006**</td>
</tr>
<tr>
<td>$F_{ST}$</td>
<td>0.354</td>
<td>-0.488</td>
<td>1.86</td>
<td>0.22</td>
</tr>
</tbody>
</table>

$\alpha$ – mean proportion of adaptive substitutions; $\omega_{MH}$ – Mantel-Haenszel estimator (equivalent to Neutrality Index); **$p < 0.01$

To confirm the signal of positive selection within $F_{ST-D_{XY}}$ outlier genes we applied the standard McDonald-Kreitman test. Using *E. typhina* subsp. *clarkii* as outgroup, the test showed an excess of fixed non-synonymous substitutions compared to polymorphic substitutions, but only one gene (i.e. endo-1,4-beta xylanase) was significant (Chi-square test; $p < 0.01$; Table S3). The other genes might have missed significance, because of too low within-population polymorphism. Therefore, the multi-locus McDonald-Kreitman test was used (Rand and Kann, 1996), which showed a significant excess of fixed non-synonymous substitutions compared to polymorphic substitutions across genes overlapping between $F_{ST}$ and $D_{XY}$ outliers (Chi-square test; $p < 0.01$; Table 3). In contrast, the combination of genes
occurring only in the \( F_{ST} \) outliers showed a non-significant excess of polymorphic non-synonymous substitutions, which may be a signal of negative selection (Egea et al., 2008). When using \( E. typhina \) subsp. \( typhina \) as outgroup, we failed to detect a signal of selection in the standard as well as in the multi-locus McDonald-Kreitman test (data not shown). Overall, these results indicate positive selection acting on genes within \( F_{ST} \)-\( D_{xy} \) outliers, which are likely involved in the pathogenic behaviour of \( Epichloë \) endophytes, with selection acting stronger within \( E. typhina \) subsp. \( typhina \).

**Discussion**

Although modern approaches relying on next generation sequencing provide a powerful resource to detect loci that may be involved in adaptation processes, genomic studies were done on only a few fungal genera, e.g. \( Saccharomyces \), the human pathogen \( Coccidioides \), \( Neurospora \), \( Mycosphaerella \), and the mycorrhizal fungus \( Suillus \) (Branco et al., 2015; Ellison et al., 2011; Liti et al., 2009; Neafsey et al., 2010; Stukenbrock et al., 2011). Moreover, none of the studies focused on the genomic architecture underlying adaptation to different host species. In this study, we undertook the first genome-wide analysis of genetic differentiation between two closely related endophytes. Our complementary approach including multiple inferences for signature of selection from between- and within-population estimates gave insight into the genes underlying differentiation between \( E. typhina \) subsp. \( typhina \) and \( E. typhina \) subsp. \( clarkii \). This information should provide grounds to generate hypotheses on the genetic architecture of reproductive isolation among natural populations. To better understand potential isolation mechanisms, we furthermore tested in mating experiments the conditions for forming hybrids on stromata.

Indeed, hybrid genotypes could be retrieved from ascospores in small proportions (7-10 %) when applying conidia from \( E. typhina \) subsp. \( typhina \) and \( E. typhina \) subsp. \( clarkii \) in different ratios on stromata of \( E. typhina \) subsp. \( typhina \). However, at applied \( E. typhina \) subsp. \( typhina \) conidia concentrations 100 times lower than that of \( E. typhina \) subsp. \( clarkii \) no hybrid genotypes were found. This may be because the proportion of hybrids was too small for microsatellite genotyping. In control experiments we could reliably detect proportions only down to only 5 %, suggesting that the amount of hybrids formed in the mating experiment may be underestimated. Furthermore, the high number of unwanted fertilizations with contaminating genotypes of other \( E. typhina \) subsp. \( typhina \) present in the greenhouse may have obscured the actual number of hybrids formed. However, under natural conditions mixtures of spermatia of different genotypes from neighboring infected grasses are also transferred, thus allowing for cross-fertilization. Our results therefore confirm that small proportions of hybrids may be formed on stromata, which is in agreement with a previous study monitoring hybrid ascospores (~9 %) in natural populations (Bultman et al., 2011). Nevertheless, genetic conflicts between spermatia of \( E. typhina \) subsp. \( typhina \) and \( E. typhina \) subsp. \( clarkii \) resulting in assortative mating could promote
matings within host-associated strains. Such a prezygotic reproductive barrier has also been suggested for sister species of *Microbotryum* based on specific recognition at variations of the mating type locus (Büker et al., 2013).

Given that hybridization between host-associated strains may occur, it is surprising that grasses infected with hybrids have never been found in natural populations (Bultman et al., 2011). Recent research based on experimental infections have suggested that this may be due to a maladaptation of hybrids to either parental host plant (Schirrmann et al., 2015). In fact, the inferior performance of experimentally generated hybrids compared to parental strains and progeny of within host-strain crosses found in planta could act as a postzygotic isolation barrier. The evidence for pre- and postzygotic barriers found between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* suggests that experimental interfertility tests may not be sufficient to delimit species boundaries in *Epichloë*. For full interfertility both pre- and postzygotic barriers should to be absent (Kohn, 2005). Because reproductive barriers were not considered for defining species of *Epichloë* in previous studies, interfertile host-associated strains were treated within a species complex consisting of subspecies or cryptic species (Craven et al., 2001; Leuchtmann et al., 2014). However, the surprisingly high genetic differentiation found between the two host-associated taxa of the *E. typhina* complex in this study indicates that despite apparent interfertility strong reproductive barriers exist and that the taxa may be seen as two separate species.

To test for signatures of selection within the two host-associated taxa, we combined multiple inferences from both within- and between-population measures of diversity computed from genome scans. Despite their ability to cross under laboratory conditions, genome-wide data using whole genome resequencing of pooled individuals representing the two host-associated taxa revealed strong genetic differentiation between host-associated strains with a mean gene-wise *F*$_{ST}$ of 0.53, consistent with previous results (Schirrmann et al., 2015). The overall high levels of *F*$_{ST}$ are indicative of a long period of isolation of genotypes, which could either result from a phase of geographic isolation or from a persistent process of sympatric divergence between the two taxa (Giraud, 2006). Teasing apart these two scenarios is a challenging task of evolutionary biology research (Bierne et al., 2013; Seehausen et al., 2014) and beyond the goals of the present work. In a conservative manner, we interpret the scenario at the Aubonne site as a case of secondary contact between populations that developed strong differentiation and host-specific adaptation in an allopatric setting, without reaching complete reproductive barriers (Feder et al., 2012; Strasburg et al., 2012). Following secondary contact, alleles could newly recombine between populations except in regions under strong selection regimes that remain protected from gene flow (Delmore et al., 2015). The mating experiments performed in the current study indicate that a small proportion of hybrids can originate from controlled crosses of between taxa, providing evidence to hypothesise that populations at Aubonne maintain the potential for recombination of genetic material between genotypes. This is supported by the genomic
pattern of divergence described by the bimodal $F_{ST}$ distribution, in which the lower peak can result from the portion of the genome that is able to recombine between the two species. In this perspective, genes underlying host-specific adaptation maintaining the divergence between the two species are hypothesised to be resistant to gene flow between populations in comparison to the homogenising remainder of the genome (Cruickshank and Hahn, 2014; Stukenbrock et al., 2011), and therefore they are expected to maintain the greatest proportion of new mutations accumulated after species divergence. Hence, in our approach we choose to consider the set of genes that appear as outliers in both $F_{ST}$ and $D_{XY}$ as the most likely candidates under selection, which underlie processes involved in the maintenance of species barriers between the two studied taxa.

Using this approach, we identified five candidate genes encoding for putative secreted proteins among the $F_{ST}$-$D_{XY}$ outliers, which are involved during plant resistance responses or in the degradation of cell walls (Eaton et al., 2015; Rudd et al., 2015), suggesting a role in the pathogenic stage of the endophytes, e.g. infection phase or stroma formation. These genes showed high $dN/dS$ ratios, indicating that they exhibit a significantly higher amount of adaptive evolution. The $dN/dS$ ratio is a powerful test of the neutral model of evolution and less subject to false positives, because it requires fewer assumptions than many others in population genetics (Kryazhimskiy and Plotkin, 2008; Yang and Bielawski, 2000). It therefore provides additional confidence in the role of selection in driving the differentiation of genes showing a significant signal. To further explore within which of the host-associated endophytes selection is acting, we used the McDonald-Kreitman test. However, the standard McDonald-Kreitman test failed to detect a significant signature of positive selection for single candidate genes, because of the lack of polymorphism within our two host-associated taxa (Bierne and Eyre-Walker, 2004; Fay, 2011). We therefore applied the test to multiple loci to correct for too low within-population polymorphism. The multi-locus McDonald-Kreitman test indicated significant positive selection on genes within $F_{ST}$-$D_{XY}$ outliers of E. typhina subsp. typhina, whereas it failed to detect a signal in E. typhina subsp. clarkii. This is in agreement with measures of within-population diversity, showing that Tajima’s $D$ was significantly reduced within $F_{ST}$-$D_{XY}$ outlier genes only in E. typhina subsp. typhina (Table 1). These results are indicative of a more distinguishable signature of positive selection within E. typhina subsp. typhina for genes that confer a local advantage over the ancestral genotypes (Ellison et al., 2011).

It has been shown that natural selection in host-associated fungi in particular affects genes encoding secreted proteins (Stukenbrock et al., 2010; Terauchi and Yoshida, 2010). Secreted proteins have been hypothesised to play a key role between Epichloë species and their respective hosts by mediating communication with the host for the maintenance of mutualism or suppression of host defense, similar to effector proteins of pathogenic fungi (Eaton et al., 2015; Rep, 2005). Secreted proteins are therefore primary targets of host-specific coevolution subjected to heterogeneous selection pressures and may evolve more
rapidly than other genes during host specialization (Rafiqi et al., 2012). The potential role of our candidate genes in the infection phase or stroma formation requires a close interaction with the host plant. Therefore, the endophytes need to evolve mechanisms that allow them to escape detection by the host, which may impose a strong evolutionary pressure on these genes leading to accelerated evolution (Giraud et al., 2006; Presti et al., 2015). Specialization to one host may therefore be beneficial for the endophyte to keep pace in the arms race, as it allows a more specific reaction towards host molecules. A link between the diversification of secreted proteins and host adaptation was established for the oomycete *Phytophthora*, in which a member of the cysteine-like effector family changed specificity towards its associated cysteine protease from the host plant (Dong et al., 2014). Secreted proteins may therefore mediate host specialization and eventually reproductive isolation among grass-associated *Epichloë* endophytes.

**Author’s contributions**

MKS performed experiments, analyzed data and wrote the paper. SZ contributed analytical tools and analyzed data. DC helped with interpretation of data. SF designed experiments and wrote the paper. AL designed experiments and wrote the paper.

**Acknowledgments**

This study was funded by the Swiss National Science Foundation (31003A_138479). Microsatellite and genomic data were generated in the Genetic Diversity Centre of ETH Zurich (GDC). C. L. Schardl provided access to annotations of the reference genome, and A. Widmer gave advice. We thank B. Blattmann and C. Michel for laboratory assistance, and M. Frei for taking care of the grasses. We also thank M. C. Fischer, M. Scharmann, and N. Zemp for providing methodological details.
References


Cruickshank, T.E., Hahn, M.W., 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Mol Ecol.


Rohmer, L., Guttmann, D.S., Dangl, J.L., 2004. Diverse evolutionary mechanisms shape the
type III effector virulence factor repertoire in the plant pathogen *Pseudomonas syringae*. Genetics 167, 1341–1360. doi:10.1534/genetics.103.019638


Supplementary material

Table S1: Individuals of *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii* used for matings with identity of species, host, sampling location in Switzerland and mating type.

<table>
<thead>
<tr>
<th>Isolate no.</th>
<th>Species</th>
<th>Host</th>
<th>Location</th>
<th>Mating type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1217</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>D. glomerata</em></td>
<td>Merishausen</td>
<td>mat-2</td>
</tr>
<tr>
<td>1206</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>La Rippe</td>
<td>mat-2</td>
</tr>
<tr>
<td>1403</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>Aubonne</td>
<td>mat-2</td>
</tr>
<tr>
<td>0702/15</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>Aubonne</td>
<td>mat-2</td>
</tr>
<tr>
<td>1218</td>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td><em>D. glomerata</em></td>
<td>Zumikon</td>
<td>mat-1</td>
</tr>
<tr>
<td>1402</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>Aubonne</td>
<td>mat-1</td>
</tr>
<tr>
<td>1401</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>Aubonne</td>
<td>mat-1</td>
</tr>
<tr>
<td>1205</td>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td><em>H. lanatus</em></td>
<td>Aubonne</td>
<td>mat-1</td>
</tr>
</tbody>
</table>

Table S2: Number and percentage of positive matings after fertilization. For positive matings, the number and percentage of progeny having the genetic background of the applied endophyte strains, crosses between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*, i.e. hybrids, and contamination with *E. typhina* subsp. *typhina* are listed.

<table>
<thead>
<tr>
<th>Fertilization</th>
<th>Positive matings</th>
<th>Applied <em>E. typhina</em> subsp. <em>typhina</em> genotype</th>
<th>Hybrids</th>
<th>Contamination with <em>E. typhina</em> subsp. <em>typhina</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>typhina</em></td>
<td>10</td>
<td>100</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td><em>E. typhina</em> subsp. <em>clarkii</em></td>
<td>11</td>
<td>55</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Equal</td>
<td>14</td>
<td>70</td>
<td>10</td>
<td>73</td>
</tr>
<tr>
<td>1:10</td>
<td>10</td>
<td>50</td>
<td>6</td>
<td>60</td>
</tr>
<tr>
<td>1:100</td>
<td>8</td>
<td>40</td>
<td>3</td>
<td>40</td>
</tr>
</tbody>
</table>
Table S3: Functions of genes encoding for putative secreted proteins within the $F_{ST}$ and $F_{ST-D_{xy}}$ outliers, including the $dN/dS$ ratio and the standard McDonald-Kreitman test between *E. typhina* subsp. *typhina* and *E. typhina* subsp. *clarkii*.

<table>
<thead>
<tr>
<th>Outlier</th>
<th>Gene name</th>
<th>Gene function</th>
<th>$dN/dS$</th>
<th>FS</th>
<th>FNS</th>
<th>PS</th>
<th>PNS</th>
<th>$\alpha$</th>
<th>NI</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{ST}$</td>
<td>maker-contig00008-augustus-gene-0.258</td>
<td>abc transporter</td>
<td>0.047</td>
<td>19</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.692</td>
</tr>
<tr>
<td></td>
<td>genemark-contig00033-processed-gene-0.25</td>
<td>Ca influx-promoting protein ehs1</td>
<td>0.313</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>0.349</td>
</tr>
<tr>
<td></td>
<td>maker-contig00183-fgenesh-gene-0.60</td>
<td>guanyl-specific ribonuclease f1</td>
<td>-1.000</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>maker-contig00095-fgenesh-gene-0.80</td>
<td>oligosaccharyl transferase subunit</td>
<td>0.166</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>-6.5</td>
<td>7.5</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>maker-contig00038-augustus-gene-0.162</td>
<td>NA</td>
<td>0.479</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0.6</td>
<td>0.4</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>maker-contig00042-fgenesh-gene-0.91</td>
<td>NA</td>
<td>0.341</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>-1</td>
<td>2</td>
<td>0.611</td>
</tr>
<tr>
<td></td>
<td>maker-contig00394-fgenesh-gene-0.45</td>
<td>NA</td>
<td>0.307</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>maker-contig00439-fgenesh-gene-0.31</td>
<td>NA</td>
<td>0.711</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>7</td>
<td>-4.6</td>
<td>5.6</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>maker-contig00258-snap-gene-0.61</td>
<td>cem domain-containing protein</td>
<td>0.564</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>maker-contig00259-snap-gene-0.67</td>
<td>cell surface protein</td>
<td>0.338</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0.333</td>
<td>0.7</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>maker-contig00712-snap-gene-0.33</td>
<td>killer kp4 smk-core</td>
<td>-1.000</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>snap-masked-contig00153-processed-gene-0.36</td>
<td>hypothetical protein</td>
<td>0.542</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>snap-masked-contig00356-processed-gene-0.28</td>
<td>transcription factor</td>
<td>0.083</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.396</td>
</tr>
<tr>
<td>$F_{ST-D_{xy}}$</td>
<td>maker-contig00477-fgenesh-gene-0.41</td>
<td>carbohydrate esterase family 8</td>
<td>1.420</td>
<td>8</td>
<td>29</td>
<td>0</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>0.299</td>
</tr>
<tr>
<td></td>
<td>maker-contig01280-augustus-gene-0.17</td>
<td>chloroperoxidase</td>
<td>1.337</td>
<td>5</td>
<td>17</td>
<td>2</td>
<td>2</td>
<td>0.705</td>
<td>0.29</td>
<td>0.257</td>
</tr>
<tr>
<td></td>
<td>maker-contig00175-fgenesh-gene-0.57</td>
<td>endo-1,4-beta xylanase</td>
<td>1.678</td>
<td>9</td>
<td>53</td>
<td>8</td>
<td>11</td>
<td>0.766</td>
<td>0.23</td>
<td>0.009**</td>
</tr>
<tr>
<td></td>
<td>maker-contig00477-augustus-gene-0.55</td>
<td>NA</td>
<td>2.173</td>
<td>4</td>
<td>33</td>
<td>5</td>
<td>11</td>
<td>0.733</td>
<td>0.27</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>snap_masked-contig00572-processed-gene-0.15</td>
<td>NA</td>
<td>1.707</td>
<td>2</td>
<td>17</td>
<td>3</td>
<td>5</td>
<td>0.803</td>
<td>0.20</td>
<td>0.099</td>
</tr>
</tbody>
</table>

FS – fixed synonymous substitutions; FNS – fixed non-synonymous substitutions; PS - polymorphic synonymous substitutions; PNS – polymorphic non-synonymous substitutions; NI – Neutrality Index; **p < 0.01; NA – not available
GENERAL DISCUSSION

In this PhD thesis, I aimed to identify the mechanisms underlying reproductive isolation in *Epichloë* endophytes, using different ecological and genetic approaches. Host-associated taxa within the *Epichloë typhina* complex are ideal to study pre- and postzygotic barriers, because they are still interfertile. However, genetic tests to examine population structure of three taxa of the *E. typhina* complex described in this thesis revealed that no hybrid strains could be detected in any of the sympatric host grasses, suggesting that these taxa are reproductively isolated (Chapter I). I therefore tested for reproductive barriers among four taxa of the *E. typhina* complex using reciprocal infections with parental and hybrid strains, which revealed a strict host-specificity for the parental strains (Chapter II). To get a more detailed idea about reproductive barriers during sexual reproduction, I assessed the formation of hybrids on stromata using mixtures of compatible and incompatible spores. Finally, to investigate the genetic basis of these barriers, I sequenced whole genomes of sympatric populations of two host-associated taxa to find genes under selection that might be involved in host-specificity and therefore reproductive isolation.

Strong differentiation of host-associated taxa within the *E. typhina* complex

The development of new microsatellite markers using the genome of *E. typhina* subsp. *poae* infecting *Poa nemoralis* allowed us to study the population structure of different *Epichloë* taxa in more detail. The results showed high genetic differentiation among interfertile taxa of the *E. typhina* complex, i.e. *E. typhina* infecting *Dactylis glomerata*, *E. typhina* subsp. *clarkii* infecting *Holcus lanatus* and *E. typhina* subsp. *poae* infecting *P. nemoralis*, at the field site in Aubonne, Switzerland, suggesting that reproductive isolation keeps natural populations genetically distinct (Chapter I). We furthermore found evidence that another sympatrically growing taxa of the *E. typhina* complex, i.e. *E. typhina* infecting *Poa trivialis* may also be reproductively isolated, because these samples form a clearly separated cluster from the other taxa in a PCA analysis. However, to get a more complete picture about the genetic differentiation of host-associated *Epichloë*, more taxa from the same species complex as well as from other *Epichloë* species should be included in future studies of population genetic structure. Preliminary results indicated that high genetic differentiation not only exists between samples from the field site in Aubonne, but also between samples from different locations. Therefore, it would be interesting to include more locations to be able to test whether reproductive isolation is consistently observed across geographic regions.
Host-specificity of parental strains and inferior performance of hybrids in planta

To understand the mechanisms that reproductively isolate *Epichloë* taxa, I tested for host-specificity and the performance of hybrid strains of *E. typhina* subsp. *typhina* (*D. glomerata*), *E. typhina* subsp. *clarkii*, *E. typhina* subsp. *poae* and *E. typhina* subsp. *typhina* (*P. trivialis*) (Chapter II). In addition to the inferior infection frequencies of hybrids *in planta*, which would conform to a postzygotic barrier, I also found strong host-specificity for these taxa. Taken together with the results from the population study, these results indicate that each host grass in this study harbours an independent *Epichloë* lineage, which is reproductively isolated from those of other host grasses. Although host grasses could be infected with hybrid strains using artificial inoculations, plants infected with hybrids could not be detected in natural populations. A complex interplay of different factors seems to prevent the infection of host grasses with hybrids in the field, although hybrid ascospores were found in small proportions on the stroma in a previous study (Bultman et al., 2011). A possible explanation could be that hybrid ascospores are outnumbered by parental spores and therefore have lower chances of infecting new host grasses. Furthermore, to confirm whether hybrids that may form initial infections are in competition with parental strains when they infect new host grasses, it would be useful to conduct double inoculation experiments on host plants.

Spore competition on the stroma and whole-genome re-sequencing

Another explanation for hybrids being only rarely formed on the stromata and not detected *in planta* in natural populations could be intrinsic reproductive barriers during sexual reproduction. After applying different ratios of conspecific and heterospecific conidia from two different host-associated strains, hybrid progeny was found only in small proportions among positive matings, indicating that matings between strains from the same host grass were more successful. This may be because of a maladaptation of heterospecific conidia to the different host grass. Heterospecific conidia could therefore have a growth disadvantage on stromata of a different host grass species, because germination and proliferation may take longer in the maladapted environment, allowing for a higher success rate of positive matings between only conspecific conidia. Positive matings within taxa may then interfere with other mating events on different parts of the stroma surface, possibly aborting matings between different taxa (Chung and Schardl, 1997). However, interference on the stroma was not tested in this study and would be the topic of further research. Another interesting aspect would be to quantify the actual amount of hybrids in each mating, using for example pyrosequencing (Ronaghi et al., 2007). Pyrosequencing would help estimate the ratio of matings between and within different taxa found on the same stroma. Furthermore, it needs to be seen whether the hypothesis of assortative mating holds true for other host-associated taxa than *E. typhina* infecting *D. glomerata* and *E. typhina* infecting *H. lanatus*.
The re-sequencing of whole genomes of *E. typhina* subsp. *typhina* (from *D. glomerata*) and *E. typhina* subsp. *clarkii* gave important insights into the genomic differentiation between interfertile taxa. The results of the measures of between species diversity, e.g. *F*$_{ST}$, indicate that the two taxa are two clearly separated species. However, based on previous gene trees that showed a nested position within *E. typhina* from other hosts, they are treated as subspecies, belonging to the same species complex (Craven et al., 2001; Leuchtmann et al., 2014). Our genomic results therefore suggest that it is likely that the actual species diversity of *Epichloë* is underestimated. By using genome-wide genetic markers, e.g. Restriction Associated DNA Sequencing (RAD-Seq), phylogenetic trees would more likely reflect the true relationships of different *Epichloë* taxa in comparison to phylogenies based on only a few genes (Pante et al., 2015).

By examining genome-wide patterns of nucleotide polymorphism, I furthermore aimed to identify genes under positive selection, which might be involved in host adaptation and eventually in reproductive isolation. It is hypothesized that secreted proteins interact with host molecules, and are therefore subjected to heterogeneous selection pressures imposed by the host (Brunner et al., 2013; Stukenbrock et al., 2011). Interestingly, genes encoding for putative secreted proteins, which were identified by a combination of *F*$_{ST}$, *D*$_{XY}$ and *dN/dS* ratios measures, were mainly involved in the infection phase of hosts (Eaton et al., 2015; Rudd et al., 2015). Some genes had functions related to the degradation of cell walls, which was rather unexpected for endophytes that are suspected to not cause damage to the host grasses. However, we do not know in which processes and in which stage of the endophyte life cycle these genes are involved. Besides the formation of stromata and the infection of new hosts (Eaton et al., 2015), it may be possible that these genes encode for proteins necessary for the uptake of nutrients in the apoplast by degrading sugars or other metabolites. Another explanation could be that these proteins were secreted as a side product by the endophytes without specific function and are therefore under positive selection to avoid further recognition by the host. Future research needs to test which processes these genes are involved in and how they affect the host-specificity of *Epichloë* endophytes. This would include more populations of *E. typhina* subsp. *typhina* (*D. glomerata*) and *E. typhina* subsp. *clarkii*, as well as more taxa from different species complexes to test whether the same genes/pathways are involved in host adaptation in different populations and species complexes. Furthermore, gene expression studies during different stages of the fungal development within the host grass may help to decipher in which processes the pathogen-related genes are actually involved.
References


ACKNOWLEDGEMENTS

First of all I would like to thank my thesis supervisor Adrian Leuchtmann for his support and ideas during my PhD. He opened the fascinating world of working with endophytes to me.

I would also like to thank Otmar Holdenrieder and Florian Schiestl for being co-examiners and reviewing my thesis.

Many thanks go to the members of the Plant Ecological Genetics group. I would like to thank Alex Widmer for always having an ear when needed. Special thanks goes to Simone Fior for being always very supportive and for sharing good as well as bad moments. Thanks go also to Dominik Brödlin, James Buckley, Martin Fischer, Marie Roumet, Mathias Scharmann, and Niklaus Zemp for all the (scientific) support and the laughs we had together.

Special thanks goes also to Maja Frei for the help in the greenhouse and keeping the grasses clean from ‘bad’ fungi. I would like to thank Bea Blattmann and Claudia Michel for their help and support with laboratory work. Without these persons I could not have finished my PhD thesis in time.

I would also like to thank Stefan Zoller, Silvia Kobel and Aria Minder from the Genetic Diversity Center (GDC) for their help and work during the PhD thesis.

Many thanks go to Kathrin Rentsch and Franziska Berger for help with any practical aspect, and Karsten Rohweder for helping me with sometimes challenging computer problems.

Last but not least I would like to thank my family and friends for being patient and supportive during my thesis.