Received: 8 February 2023 | Revised: 23 November 2023 | Accepted: 30 November 2023

DOI: 10.1111/mec.17242

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Two-speed genomes of *Epichloe* fungal pathogens show contrasting signatures of selection between species and across populations

Artemis D. Treindl[1,2] | Jessica Stapley[3] | Daniel Croll[4] | Adrian Leuchtmann[1]

[1]Plant Ecological Genetics Group, Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

[2]Biodiversity and Conservation Biology, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

[3]Plant Pathology Group, Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

[4]Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland

Correspondence
Adrian Leuchtmann, Plant Ecological Genetics Group, Institute of Integrative Biology, ETH Zurich, Universitästrasse 16, CH- 8092 Zurich, Switzerland.
Email: adrian.leuchtmann@retired.ethz.ch

Funding information
Swiss National Science Foundation, Grant/Award Number: 31003A_169269

Handling Editor: Sara Branco

Abstract

Antagonistic selection between pathogens and their hosts can drive rapid evolutionary change and leave distinct molecular footprints of past and ongoing selection in the genomes of the interacting species. Despite an increasing availability of tools able to identify signatures of selection, the genetic mechanisms underlying coevolutionary interactions and the specific genes involved are still poorly understood, especially in heterogeneous natural environments. We searched the genomes of two species of *Epichloe* plant pathogen for evidence of recent selection. The *Epichloe* genus includes highly host-specific species that can sterilize their grass hosts. We performed selection scans using genome-wide SNP data from seven natural populations of two co-occurring *Epichloe* sibling species specialized on different hosts. We found evidence of recent (and ongoing) selective sweeps across the genome in both species. However, selective sweeps were more abundant in the species with a larger effective population size. Sweep regions often overlapped with highly polymorphic AT-rich regions supporting the role of these genome compartments in adaptive evolution. Although most loci under selection were specific to individual populations, we could also identify several candidate genes targeted by selection in sweep regions shared among populations. The genes encoded small secreted proteins typical of fungal effectors and cell wall-degrading enzymes. By investigating the genomic signatures of selection across multiple populations and species, this study contributes to our understanding of complex adaptive processes in natural plant pathogen systems.

KEYWORDS
effector genes, fungal pathogen, genome compartmentalization, genome scan, population genomics, selective sweep

1 | INTRODUCTION

Identifying adaptive traits, elucidating their genetic basis and understanding the forces driving divergence of populations and species lie at the core of evolutionary biology research. Host–pathogen interactions are fascinating biological models to study adaptation over short evolutionary timescales because pathogens are entangled in tight biotic interactions with their hosts driving rapid coevolution (Kurtz et al., 2016; Paterson et al., 2010). Pathogens have negative effects on host fitness and exert strong selection pressure on traits

that increase resistance. In turn, pathogens must adapt to overcome host defences and establish infection as they depend on colonizing their host for survival or to complete their reproductive cycle. This interaction manifests in ongoing cycles of adaptation and counteradaptation driven by strong antagonistic selection and leading to rapid evolutionary change (Papkou et al., 2019; Tellier et al., 2014).

Antagonistic selection between the host and pathogen was formalized as two main types of coevolutionary dynamics: the *arms race* model (also called *Escalatory Red Queen*) and the *trench warfare* model (or *Fluctuating Red Queen*, Brockhurst et al., 2014, Ebert & Fields, 2020; Woolhouse et al., 2002). Arms race dynamics are characterized by recurrent selective sweeps continuously driving new beneficial variants to fixation within a population. Because variants that are linked to the beneficial allele will also reach high frequency or become fixed (genetic hitchhiking), selection produces localized depletion of genetic variation around the selected locus, creating the signature of a selective sweep. Selective sweeps initially result in strong linkage disequilibrium (LD) and an excess of low-frequency variants across the affected regions (Tajima, 1989). In contrast, trench warfare dynamics involve the continuous rise and fall in allele frequencies driven by negative frequency-dependent selection resulting in the maintenance of polymorphism at the coevolving locus. Because none of the different alleles are expected to become fixed, this dynamic is detectable as an excess of intermediate frequency variants around the genomic region under selection matching the genomic signature of balancing selection (Charlesworth, 2006).

The arms race and trench warfare coevolutionary models represent extreme scenarios of what is in reality a continuum of allele frequency dynamics observable at the genomic level (Brown & Tellier, 2011; Ebert & Fields, 2020; Tellier et al., 2014; Woolhouse et al., 2002). Importantly, selection pressure also varies across space in heterogeneous environments creating spatially structured populations where coevolution with local host genotypes can lead to asynchrony in evolutionary dynamics across populations and local adaptation (Thompson, 2005; Woolhouse et al., 2002). Theory suggests that the way antagonistic selection ultimately plays out in terms of evolutionary dynamics largely depends on population structure and levels of gene flow among host and pathogen populations (Gandon & Michalakis, 2002; Gandon & Nuismer, 2009). Additionally, the relative strength of selection versus stochastic processes such as genetic drift largely depends on the N_e of a given population (Charlesworth, 2009). To improve our understanding of antagonistic coevolutionary dynamics in spatially and genetically complex biological systems, analyses should cover multiple populations and closely related species to assess demographic processes simultaneously (Amandine et al., 2022).

Fungal plant pathogens are appropriate systems to study such coevolutionary dynamics, because they often have small tractable genomes and large population sizes facilitating the investigation of host-driven adaptation (Gladieux et al., 2014). Fungi specialized on a single host species with the ability to kill or sterilize provide particularly attractive systems (Ashby & Gupta, 2014), as strong antagonistic selection is expected to foster arms race dynamics via selective sweeps (Persoons et al., 2017; Tellier et al., 2014). Research on fungal crop pathogens has provided broad insights into the key genes involved in host infection (McDonald & Stukenbrock, 2016). To suppress or evade host defence responses, plant pathogens produce a set of small secreted proteins. Identified as effectors, these cysteine-rich proteins are characterized by high expression upon contact with the host (reviewed in Lo Presti et al., 2015, van der Does & Rep, 2017). Many effector genes lack conserved domains and homologues in closely related species (Franceschetti et al., 2017) making their prediction challenging. A second group of proteins important for plant infection are cell wall-degrading enzymes (CWDEs) helping to access nutrients following infection (Kubicek et al., 2014). As key players in host–pathogen interactions, effectors and CWDEs are common targets of selective sweeps and the underlying genes are among the most rapidly evolving in pathogen genomes (Aguileta et al., 2009; Plissonneau et al., 2017).

Several fungal plant pathogens possess compartmentalized genomes with a conserved gene-rich compartment and a dynamic repeat-rich compartment largely devoid of genes (Croll & McDonald, 2012). This bipartite genome organization has been referred to as the 'two-speed genome' because the compartments are thought to experience different evolutionary rates (Dong et al., 2015; Raffaele & Kamoun, 2012). Repeat-rich regions are hot spots for duplication, deletion and recombination as a result of transposable element (TE) activity rendering these regions highly polymorphic. Furthermore, these regions may be targeted by DNA defence mechanisms such as repeat-induced point mutation (RIP) leading to rapid mutation accumulation (Möller & Stukenbrock, 2017; Seidl & Thomma, 2017; Stukenbrock & Croll, 2014). Effector genes are often localized in or near these dynamic genome compartments where they may experience faster rates of sequence evolution (Dong et al., 2015; Faino et al., 2016; Grandaubert et al., 2014; Schirawski et al., 2010). Together, these observations have highlighted the role of such dynamic compartments in the adaptive evolution of fungal pathogens.

Most research on fungal plant pathogen adaptation has focussed on agricultural systems; however, coevolution of plant–fungal interactions in natural ecosystems remains poorly explored. The pathogenic members of the ascomycete genus *Epichloe* provide an ideal system as they infect different grasses in natural and semi-natural ecosystems, reproduce sexually and severely affect host fitness through sterilization (Chung & Schardl, 1997; Clay & Schardl, 2002; White Jr. & Bultman, 1987). Furthermore, rapid adaptation and divergence at both the population and species level may be fuelled by the two-speed genome organization with conserved gene-rich and dynamic repeat-rich RIP-targeted compartments, hereafter referred to as 'AT-rich regions' (Treindl, Stapley, Winter, et al., 2021). Genome-wide analyses of SNPs of seven natural sympatric populations of two *Epichloe* sibling species, *E. typhina* (Pers.) Brockm. and *E. clarkii* J. F. White, revealed substantial genetic variation and high recombination rates (Treindl et al., 2023). Populations showed clear evidence of genetic structure that may reflect selection acting at the population level, driving divergence and generating locally adapted pathogen

populations (see Croll & McDonald, 2016). Although the two species share nearly the same ecological niche, they differ substantially in population demography possibly related to different dispersal abilities. *E. clarkii* populations have lower genetic diversity suggesting smaller N_e and have lower levels of gene flow among populations compared with *E. typhina* (Treindl et al., 2023).

In this study, we aim to investigate these two closely related species *E. typhina* infecting *Dactylis glomerata* and *E. clarkii* infecting *Holcus lanatus* across seven populations each to identify signatures of recent selection. We analysed genomic regions showing signatures of selection to assess shared and divergent selective sweeps among populations and species. By combining information on the localization in the genome, the predicted function of genes in selective sweep regions and differential expression analyses, we identified the possible drivers of coevolution between *Epichloe* species and their hosts.

2 | MATERIALS AND METHODS

2.1 | Study system and population sampling

We analysed genome sequencing data for the two systemic endophytes *E. typhina* and *E. clarkii* specialized on different host grasses. Both focal species reproduce sexually by forming external fruiting bodies, so-called stromata, which enclose undeveloped inflorescences of the host plant and inhibit flowering and seed production, a phenomenon referred to as choke disease (Leuchtmann, 2003). Whole plants are usually infected by a single haploid genotype (Leuchtmann & Clay, 1997), and obligate outcrossing is mediated by a symbiotic fly of genus *Botanophila* which transfers gametes between stromata (bipolar heterothallic mating system; Bultman & Leuchtmann, 2008). After successful fertilization with gametes of the opposite mating type, ascospores are produced that are then wind-dispersed and mediate horizontal transmission to new hosts (Chung & Schardl, 1997; for a complete life cycle, see Schirrmann et al., 2018). Taxonomically *E. typhina* and *E. clarkii* are sibling species and members of the *E. typhina* species complex, an aggregate of genetically differentiated lineages infecting distinct host grasses (Treindl, Stapley, & Leuchtmann, 2021). Throughout this study, we use the name *E. typhina* to refer solely to the lineage specialized on *D. glomerata*.

The previously generated dataset of whole-genome SNP data from haploid isolates included seven pairs of natural sympatric populations across Europe sampled between 2016 and 2018 (Figure 1; Tables S1 and S2). Detailed sampling and SNP genotyping methods were described previously (Treindl et al., 2023). In brief, fungal strains were isolated from fruiting bodies encasing infected flowering tillers of host grasses. Each strain represents a distinct genotype from an individual plant. Genomic DNA was extracted from mycelium of pure liquid cultures and Illumina paired-end sequencing of 2×150-bp read length, and an insert size of 330 bp was performed

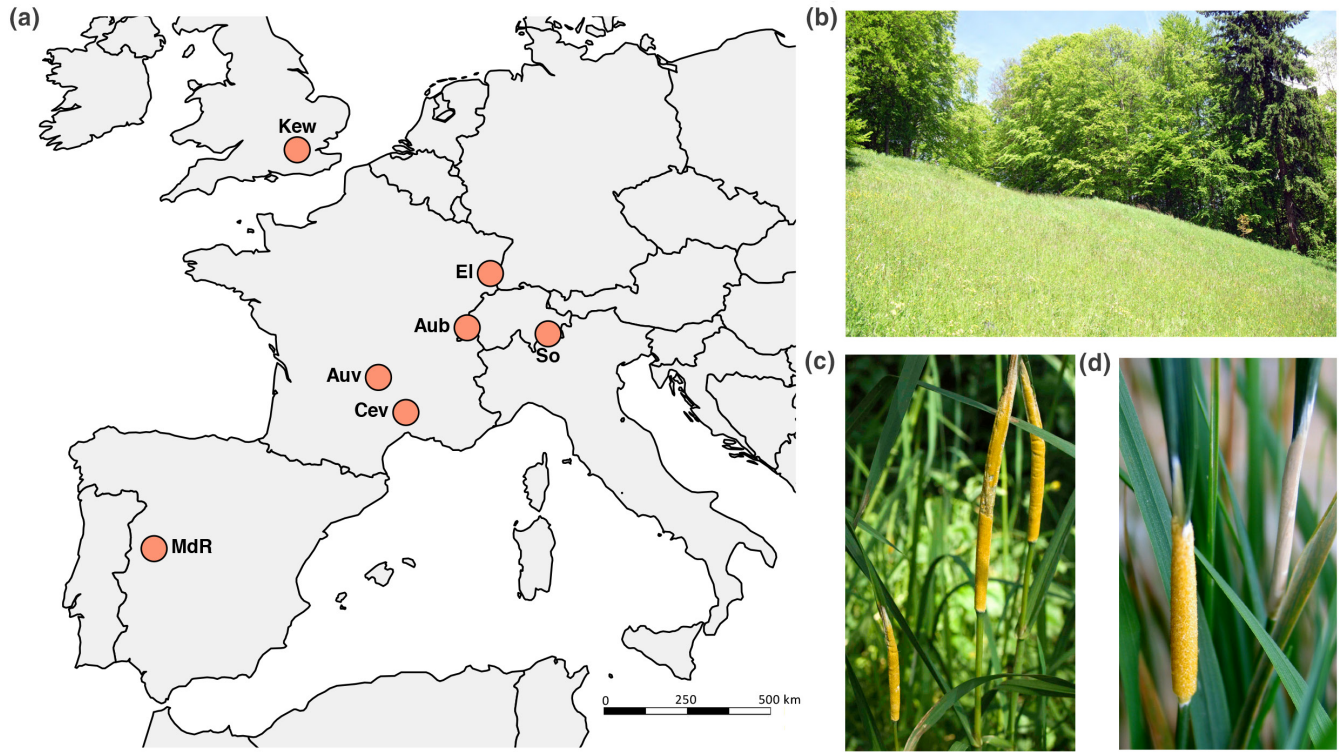


FIGURE 1 Sampling locations and disease symptoms of sympatric *E. typhina* and *E. clarkii* populations. (a) Map of Europe with collection points. Abbreviations refer to places of origin of the sampled populations: Aub = Aubonne, Auv = Auvergne, Cev = Cevennes, El = Alsace, Kew = Kew Gardens London, MdR = Montemayor del Rio, So = Soglio. For more details, see Tables S1 and S2. (b) Sampling site Aubonne, Switzerland. (c) Ripe fungal fruiting bodies (stromata) of *E. typhina* infecting *D. glomerata*. (d) Stromata of *E. clarkii* infecting *H. lanatus*.

yielding on average of 16× coverage per isolate. The sampling included isolates collected in 2005 from the population in Aubonne (Aub), Western Switzerland (16 *E. typhina* and 18 *E. clarkii*), and the population was resampled in 2018. The 2005 samples were genetically similar to the 2018 samples (Aub), but we have treated the data from these two sampling events as separate populations in analyses (2005 = AubX, 2018 = Aub). We only report results for the 2018 population unless indicated otherwise. In total, we analysed 191 *E. typhina* isolates ($n = 14–33$ per population) and 208 *E. clarkii* isolates ($n = 18–30$).

2.2 | Variant detection and data polarization

We used BWA-mem (Li & Durbin, 2009) to map reads against the chromosome-level assembly of the *E. typhina* reference genome (Ety_1756, BioProject ID PRJNA533210) and *E. clarkii* (Ecl_1605_22, BioProject ID PRJNA533212) and the Genome Analysis Toolkit (McKenna et al., 2010) to call variants. After filtering, we retained 658,021 biallelic SNPs in the *E. typhina* dataset and 400,033 SNPs in the *E. clarkii* dataset. To determine ancestral allele state in our focal species, we included high-coverage sequencing data from the closely related outgroup species *E. festucae* strain Fl1 (SRX5056454, 2 × 250 Illumina paired-end, 550 bp insert size, ~87× coverage; see Winter et al., 2018). Outgroup genotypes were extracted and subsequently used to annotate variant call files with the inferred ancestral allele information using bcftools (Li et al., 2009). We were able to assign ancestral states for 480,102 SNPs in *E. typhina* and 219,085 SNPs in *E. clarkii* (72% and 55% of the total number of SNPs in *E. typhina* and *E. clarkii*, respectively). We predicted the effect of SNPs on encoded proteins using SnpEff 4.3 m (Cingolani et al., 2012).

2.3 | Genome scans for signatures of selective sweeps within populations

To detect regions with signatures of positive selection in *E. typhina* and *E. clarkii* populations, we used two different selection scan approaches. Selection scans were performed separately for each of the seven populations using only SNPs with known ancestral states. We used a 99.9th percentile threshold of the test statistics distribution to identify the most likely selective sweep regions. As a first approach, we used extended haplotype homozygosity (EHH), a measure of linkage disequilibrium, which is expected to be increased across haplotypes that contain an allele under selection. We calculated the integrated haplotype scores (iHS) implemented in the R package REHH v3.0.1 (Gautier et al., 2017). Tests were performed for each chromosome separately and with SNPs filtered for a minor allele frequency > 5%. Candidate regions were determined by identifying clusters of outlier SNPs within sliding windows of 10-kb with 5-kb overlap for *E. typhina* and windows of 20-kb with 10-kb overlap for *E. clarkii*, respectively. Different window sizes were used to account for the different SNP densities in the two species and correspond to windows of 100 SNPs on average in both species. Windows with at least two markers displaying an absolute value of the statistic above the percentile threshold were classified as candidate sweep regions and neighbouring windows containing outlier SNPs were merged into one candidate region.

In a second approach, we analysed shifts in the allele frequency spectrum consistent with positive selection using the software SweeD v3.2.1 (Pavlidis et al., 2013). SweeD implements a composite likelihood ratio (CLR) test based on the SweepFinder algorithm (Nielsen et al., 2005). We excluded sites that were monomorphic (invariant) within each population and computed CLR on the unfolded SFS of each chromosome using a grid size of 5000 in *E. typhina* and 3000 in *E. clarkii*. The grid size specifies the number of positions in the sequence where the CLR will be computed and we chose different window sizes to account for the differences in SNP density between the two species resulting in approximate window sizes of 1 kb for *E. typhina* and 2 kb for *E. clarkii*. We repeated the analyses using smaller grid sizes with CLR tests computed every 100th SNP and every 200th SNP on average. This change decreased likelihood scores but did not increase the size of outlier sweep regions indicating that smaller windows were appropriate to detect sweep signatures. The 99.9th percentile of the distribution of the CLR statistic was used as an outlier threshold to identify SNPs under positive selection. We concatenated adjacent outlier positions into a single region if the distance was <5 kb. We extended candidate sweep regions based on genome-wide estimates of linkage disequilibrium (LD) decay. In *E. typhina*, LD decayed below $r^2 = 0.2$ within 1 kb in all populations. Hence, we added 1 kb at each end of a selective sweep region. In *E. clarkii*, we added 1 kb to each region for the populations Aub, AubX, Auv and So; 2 kb for populations El and Cev; and 5 kb for populations MdR and Kew to account for slower LD decay.

2.4 | Signatures of divergent selection between populations

In addition to within-population EHH analyses, we also performed cross-population EHH tests (XP-EHH), implemented in the R package REHH v3.0.1 (Gautier et al., 2017). We focussed on scans on a subset of the populations Aub, Auv and El, which had among the smallest geographic distances between them and showed relatively low levels of genetic differentiation based on F_{ST} (*E. typhina*: 0.06, 0.06 and 0.05; *E. clarkii*: 0.14, 0.13 and 0.19 for population pairs Aub-Auv, Aub-El and Auv-El, respectively; for comparison, medians across all population pairs excluding AubX were 0.1 and 0.23; Treindl et al., 2023). We considered these population pairs to be among the most likely to experience gene flow, thereby increasing power to detect loci under divergent selection. We used the same procedure as for the within-population iHS scans to determine outliers and define candidate regions.

TABLE 1 Number of selective sweeps identified by different methods (iHS or CLR) in populations of *E. typhina* and *E. clarkii* (for abbreviations and origin of populations, see Figure 1).

	Aub	Auv	Cev	El	Kew	MdR	So	All	Unique[a]
E. typhina									
iHS	30	39	32	38	19	38	32	228	163
CLR	22	21	24	10	13	19	15	124	85
Overlap[b]	1	4	3	2	5	2	4	21	15
E. clarkii									
iHS	13	12	15	12	9	12	17	90	62
CLR	16	6	17	15	3	12	18	87	66
Overlap[b]	0	0	1	0	0	0	1	2	2

Note: For detailed summary statistics of selective sweeps, see Figures S3 and S4.

[a] Number of sweep regions unique to a single population.

[b] Number of sweep regions that overlapped between methods.

2.5 | Association between selective sweeps and dynamic AT-rich regions

For each species, we combined iHS and CLR sweeps from all populations to compile 'species-wide' candidate region sets containing all loci under selection in any population. Shared sweeps that showed a signature of selection in multiple populations were only counted once and overlapping sweep regions from different populations were concatenated in larger regions. We tested whether sweeps and AT-rich regions overlapped more than expected by chance using permutation tests with the 'overlapPermTest' function in the R package *regionR* with 5000 iterations (Gel et al., 2016) with count.once = TRUE.

2.6 | Gene ontology enrichment analysis

We performed gene ontology (GO) enrichment analyses for genes located in selective sweep regions detected by the two different within-population selection scans (iHS and CLR). GO annotations for conserved protein domains were based on PFAM (Bateman et al., 2002) and PANTHER databases (Thomas et al., 2003). GO enrichment was assessed for all sweep regions using the R package 'topGO' and implementing a Fisher's exact test using the function 'runTest' (Alexa & Rahnenführer, 2009). We considered only GO terms with a minimum of five annotated genes. The analysis was performed separately for iHS sweeps and CLR sweeps. For each selection scan method, we performed the analysis for all sweeps independent of their location and additionally for the subsets of sweeps overlapping or not with AT-rich regions. We performed the same analysis for divergent sweep regions detected by XP-EHH among three populations.

2.7 | Differential gene expression analysis

RNAseq data were generated from isolates of reference strains grown *in culture* and *in planta* to identify differentially expressed genes (Treindl, Stapley, Winter, et al., 2021). Genes expressed *in culture* were identified from strains grown on supplemented malt-extract agar using three technical replicates per strain. For genes expressed *in planta* commercial cultivars of *D. glomerata* and *H. lanatus*, plants were artificially inoculated with the respective strains. Pseudostems of three plants (biological replicates) with two technical replicates per plant were then used for RNA extraction. We mapped RNAseq reads to the matching reference genomes using STAR v2.5.3a (Dobin et al., 2013) and counted the number of reads for each gene using the R package 'Rsubread' (Liao et al., 2019). Read counts from *in planta* technical replicates were merged. Then, we used the R package 'edgeR' (Robinson et al., 2010) to identify differentially expressed genes using normalized library size based on mapped reads.

2.8 | Data and code availability

Genome assemblies are available on GenBank (Accession nos. PRJNA533210 PRJNA533212). Genome sequencing data have been deposited in the European Nucleotide Archive ENA (Accession no. PRJEB59262). Additional figures of selection scans are available on Zenodo (10.5281/zenodo.7573676).

3 | RESULTS

3.1 | Evidence of recent selection in *Epichloe* populations

We analysed signatures of recent selection in seven sympatric population pairs of *E. typhina* and *E. clarkii* from Switzerland, France, Spain and the UK, covering the centre of distribution of both pathogens and host grasses (Figure 1). We scanned genome-wide SNP datasets with inferred ancestral states using linkage-based and SFS-based scans. Analyses were performed separately within each population and were based on approaches mitigating the confounding effects of demography (Nielsen et al., 2005; Vitti et al., 2013). Using a 99.9th percentile, we identified evidence for selective sweeps in all seven

chromosomes and populations from both species. *E. typhina* populations showed 228 sweeps using iHS (19–39 per population) and 124 using CLR scans (10–22 per population; Table 1). In *E. clarkii*, we detected 90 sweeps using iHS (9–17 per population) and 87 sweeps using CLR scans (3–18 per population). Regions identified by iHS and CLR scans showed weak concordance with only 21 sweeps (or 5.96%) in *E. typhina* and 2 sweeps (or 1.12%) in *E. clarkii* overlapping (Figure 2; Figure S1).

On average, sweeps detected using CLR tended to be longer than sweeps detected by iHS with a larger proportion of the genome being covered by CLR-based sweeps (*E. typhina* 2.94% vs. 1.39% and *E. clarkii* 1.87% vs. 0.85% averaged across populations; Tables S3 and S4). Inferred sweeps were narrower with shorter haplotypes in *E. typhina* consistent with the faster decay in LD compared with *E. clarkii* (Treindl et al., 2023). The difference is likely also influenced by the choice of larger window sizes during analysis due to lower polymorphism in *E. clarkii*. We detected the fewest and widest selective sweep regions in the Kew population of *E. clarkii* using CLR, which is consistent with the small sample size ($n=19$), low genetic diversity and high degrees of LD (Treindl et al., 2023).

We analysed whether the same loci were under selection in different populations based on overlaps in sweep regions. We performed this separately for each selection scan method and found most selective sweep regions were unique to a single population (*E. typhina*: 72.4% of iHS and 60% of CLR-based sweeps; *E. clarkii*: 62.9% of iHS and 75.8% of CLR-based sweeps; Table 1; Tables S5 and S6; Figure S2). Among the shared selective sweep regions, most were under selection in only two populations (*E. typhina*: 66.7% of iHS- and 70.6% of CLR-based sweeps; *E. clarkii*: 82.6% of iHS and 81.3% of CLR-based sweeps). In *E. typhina*, 15 genomic regions showed evidence of a selective sweep in three or more populations based on iHS scans and 10 regions based on CLR scans, whereas in *E. clarkii* only four (iHS) and three (CLR) regions were under selection in three or more populations. We selected population pairs which shared the most sweeps and tested if the number of sweeps overlapping by 50% of their length or more was higher than expected by chance. Population pairs sharing the most iHS-based sweeps in *E. typhina* were Cev-El ($n=7$, $Z=7.9697$, $p<0.001$) and El-MdR ($n=8$, $Z=7.8517$, $p<0.001$), and population pairs sharing the most CLR-based sweeps in *E. typhina* were Auv-El ($n=6$, $Z=8.497$, $p<0.001$), Auv-MdR ($n=6$, $Z=7.645$, $p<0.001$) and Auv-So $n=(6$, $p<0.001$). Population pairs sharing the most iHS-based sweeps in *E. clarkii* where Cev-So ($n=5$, $Z=6.7353$, $p<0.001$) and Aub–So ($n=6$, $Z=9.9234$, $p<0.001$), and population pairs sharing the most CLR-based sweeps in *E. clarkii* were Aub-El ($n=4$, $Z=5.857$, $p<0.003$) and Aub-So (4, $Z=10.684$, $p<0.001$). In our previous work, we found that both *E. typhina* and *E. clarkii* showed evidence for isolation-by-distance (IBD) with geographically more distant populations showing higher levels of pairwise F_{ST} (Treindl et al., 2023). We tested whether the number of shared sweeps between populations decreased with increasing genetic and geographic distance. For this, we calculated the fraction of overlapping sweeps for each population pair as the mean of the fraction of sweep intervals in one population that overlap with a sweep interval in the second population, that is (the fraction in population A + fraction in population B/2). We found no significant correlations between shared sweeps and geographic distance in either species (Figures S3 and S4). However, we found that the proportion of shared CLR-based sweeps was weakly negatively correlated with genetic distance in *E. clarkii* ($p=0.081$) consistent with the stronger isolation-by-distance in this species (Treindl et al., 2023).

We evaluated candidate regions showing signatures of divergent selection identified by cross-population XP-EHH scans between three population pairs: Aub, Auv and El. Overall, we detected almost twice as many divergent sweeps in *E. typhina* compared with *E. clarkii* (124 and 63 sweeps corresponding to 91 and 45 distinct regions, respectively). A similar proportion of divergent selective sweep regions also showed a signature of selection within at least one of the populations (23% and 22%, respectively). Most regions were under divergent selection between one population pair (67% and 69%, respectively). About a third of the divergent selective sweep regions were under divergent selection between one population compared with both other populations (31% and 31%, respectively). In *E. typhina*, two regions were under divergent selection between all three population pairs.

3.2 | Genomic context of selective sweep regions

We analysed whether specific genome compartments of *Epichloe* contributed disproportionally to recent adaptive evolution indicated by selective sweeps. Using permutation tests, we found that selective sweep regions identified by iHS overlapped significantly more often with AT-rich regions than expected by chance: 80.6% of regions overlapped in *E. typhina* (125 out of 155, $p=0.0002$) and 94.6% of regions overlapped in *E. clarkii* (53 out of 56, $p=0.0002$, Figure 3). We found a similar enrichment for CLR-based scans in both species: 86.8% of regions overlapped in *E. typhina* (66 out of 76, $p=0.0002$) and 92.2% of regions overlapped in *E. clarkii* (59 out of 64, $p=0.0004$).

3.3 | Gene content analyses of selective sweep regions

We analysed the gene content of different sweep regions to identify potential enrichment in specific gene functions. We grouped sweep regions by species by combining evidence from individual populations. In total there were 228 sweeps detected by iHS in *E. typhina* populations with 110 regions being shared between at least two populations resulting in a total of 163 unique sweep regions (Table 2; Figure S5). Many sweep regions were located fully in AT-rich regions; however, these regions contained only two predicted genes in either species. Neither of the two genes encoded conserved protein domains, however a blast search revealed that the genes share homology to genes

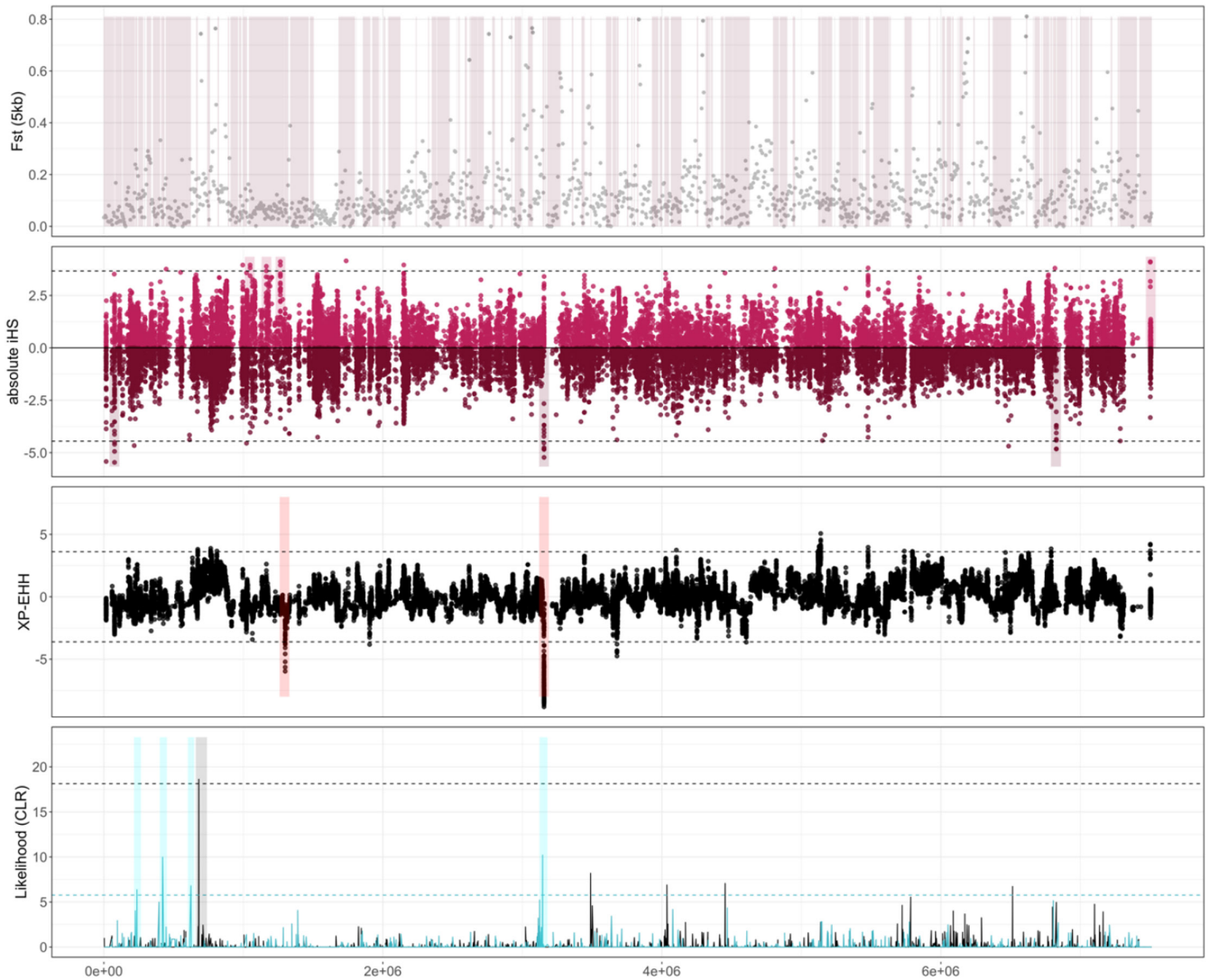


FIGURE 2 Signatures of selective sweeps on *E. clarkii* chromosome 4. Chromosome-wide SNP data and sweeps identified for the population pair Auv–So. The top panel shows pairwise F_{ST} values, averaged across 5-kb windows. Shaded rectangles represent the locations of AT-rich regions. The second panel shows the absolute values of the integrated haplotype score (iHS) calculated at each SNP locus for which the ancestral allele state was known. Scores for Auv are shown at the top and scores for So are negatively transformed and shown at the bottom. Horizontal dashed lines indicate the 99.9% percentile threshold which was used as a cut-off to identify outlier SNPs and inferred iHS sweeps are shown as shaded rectangles. The third panel shows the cross-population extended haplotype homozygosity (XP-EHH) scores calculated between the two populations. Dashed lines indicate 99.9% percentile threshold which was used as a cutoff to identify outlier SNPs and inferred divergent sweeps are shown as shaded rectangles. Positive and negative XP-EHH values refer to the direction of selection: Positive values indicate selection in Auv, and negative values indicate selection in So. In the bottom panel, composite likelihood ratio (CLR) scores are plotted for Auv (black) and So (blue), coloured dashed lines indicate respective 99.9% threshold and coloured rectangles highlight inferred CLR sweeps. Note the region at around 3 Mb in which a selective sweep was detected by both iHS and CLR scans. The same region is also under divergent selection between these two populations and overlaps with several relatively short AT-rich regions. Equivalent composite plots for each pairwise population combination and each of the seven chromosomes are available at https://doi.org/10.5281/zenodo.7573676 ($n = 294$).

or are members of gene clusters involved in alkaloid biosynthesis in *Epichloe* (Schardl, Young, et al., 2013): Ety_004235 showed homology to IolF (78% identity), Ety_007055 showed homology to the partial EAS gene cluster (71% identity); Ecl_003076 showed homology to the PER gene cluster (89% identity) and Ecl_004371 showed homology to IolF (87% identity).

We analysed the overall functions encoded by genes in selective sweep regions using gene ontology terms. We focussed on regions containing the highest numbers of genes. In *E. typhina*, protein functions related to oxidation–reduction processes (GO:0055114) and amide biosynthetic processes (GO:0043604) were enriched across all CLR-based sweep regions (Table 3). Both GO terms were also enriched in iHS-based sweep regions but less significantly. Proteins related to oxidation–reduction processes and intracellular signal transduction (GO:0035556) were also enriched in sweeps detected by both methods. In selective sweeps

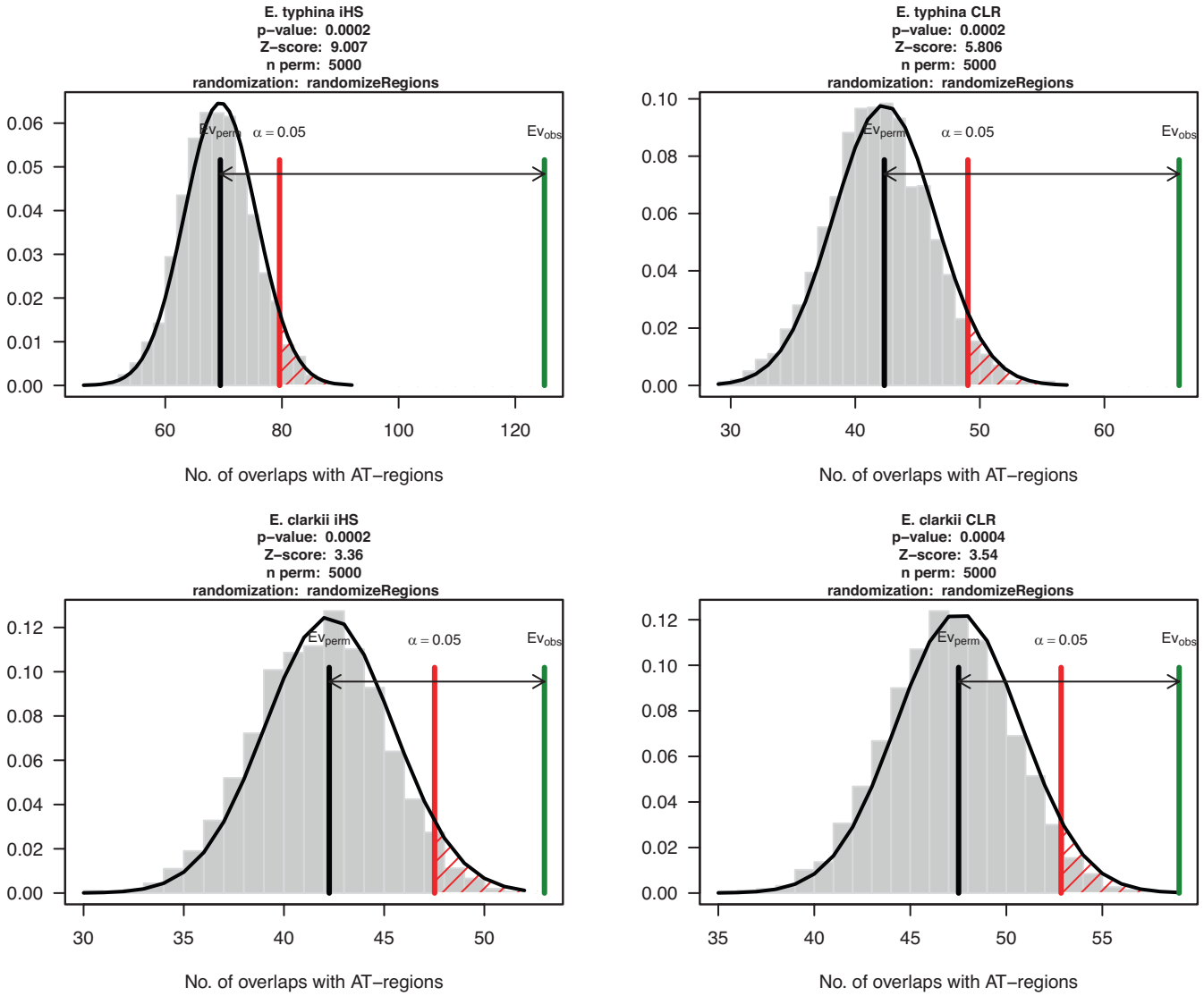


FIGURE 3 Association between regions under selection and AT-rich regions. Grey histogram represents the number of overlaps using a randomized region set with a fitted normal distribution Ev_{perm} (black line) showing the mean of the randomized counts and Ev_{obs} (green line) indicating the observed overlap count. Both *E. typhina* (top) and *E. clarkii* (bottom) and both sweep detection methods are shown (iHS left, CLR right).

overlapping with AT-rich regions, protein functions involved in cellular amide metabolic processes (GO:0043603) were enriched in iHS and CLR sweeps, but below our conservative cut-off (GO:0043603: $p = 0.012$). The strongest evidence of enrichment was found for biological functions related to protein phosphorylation (GO:0006468) across all CLR sweeps and the same GO term was enriched in AT-overlapping iHS and CLR sweeps ($p = 0.002$ and $p = 0.00012$, respectively).

In *E. clarkii*, the same GO term for protein phosphorylation (GO:0006468) showed the strongest evidence of enrichment across all CLR sweeps and for CLR sweeps overlapping with AT-rich regions. Proteins related to mycotoxin biosynthetic processes (GO:0043386) and establishment of cellular localization (GO:0051649) were also enriched across all sweeps and AT-overlapping sweeps detected by CLR-based scans and iHS-based scans, respectively. Among all iHS sweeps, functional domains involved in cellular-modified amino acid biosynthetic processes (GO:0042398) were enriched.

Among the 59 genes in *E. typhina* selective sweep regions with significant GO term annotations, 20 encoded protein kinases (GO:0006468), 7 encoded ribosomal proteins (GO:0043603 and GO:0043604) and 4 encoded cytochrome P450 monooxygenases (GO:0055114). Among the 32 genes in *E. clarkii* selective sweep regions with significant GO term annotations, 21 encoded protein kinases (GO:0006468) and 5 encoded proteins putatively involved in the production of toxic cyclic peptides (ustYa-like proteins, GO:0043386). Significantly enriched GO terms ($p < 0.01$) for all sweep regions and grouped by genomic compartment are available as Supplementary Material 10.5281/zenodo.7573776.

3.4 | Differential gene expression analysis

We compared expression levels of genes between *Epichloe* reference strains grown on culture plates (in vitro) and in their respective host plants (*in planta*). Overall, we found 1037 and 1598 genes with significantly higher expression levels *in planta* corresponding to 12.6% and 22.1% of genes in *E. typhina* and *E. clarkii*, respectively. A total of 1440 (17.5%) and 1458 (20.1%) genes had lower expression levels *in planta*. Differentially expressed genes in *E. typhina* and *E. clarkii* shared similar GO terms, including mycotoxin biosynthetic and oxidation–reduction processes being upregulated *in planta*. Genes involved in copper ion transmembrane transport had lower expression levels *in planta* relative to in vitro. Additionally, in *E. typhina*, genes related to pathogenesis and sterol metabolic processes were downregulated *in planta*. In *E. clarkii in planta* downregulated genes were enriched for functions related to gluconeogenesis, fungal cell wall organization and chitin biosynthesis. Enriched GO terms of differentially expressed genes between *in planta* and in vitro are listed in Table S7.

TABLE 2 Number of selective sweep regions and corresponding gene content for *E. typhina* and *E. clarkii* genomes subdivided by method (iHs or CLR) and compartment (GC- or AT-rich) with respective overlaps.

	E. typhina		*E. clarkii*	
	Sweep regions	Genes	Sweep regions	Genes
Method				
iHS	163	297	62	162
CLR	85	388	66	632
Overlap[a]	15	89	2	3
Compartment				
GC-rich iHS	31	172	3	47
GC-rich CLR	14	78	8	46
AT-overlap[b] iHS	66	124	36	114
AT-overlap CLR	48	323	42	605
AT-rich iHS	69	1	25	2
AT-rich CLR	27	1	19	0

[a]Number of sweep regions that overlapped between methods.
[b]Number of sweep regions overlapping but not fully located in AT-rich regions.

3.5 | Pathogenesis-related genes under selection

Regions under selection in the genomes of *E. typhina* and *E. clarkii* were not more likely to contain candidate effector genes than expected at random (Figure S6). In total, 19 out of 138 (*E. typhina*) and 23 out of 144 (*E. clarkii*) candidate effectors were in selective sweep regions identified by either iHS or CLR. Out of these, 9 and 10 genes encoded a conserved protein domain, in each species, respectively. One candidate effector gene (Ecl_000006) encoded a domain related to ribotoxins (fungal extracellular RNases) and was in a CLR-sweep region shared by all *E. clarkii* populations except MdR and the

TABLE 3 Biological functions overrepresented among selective sweeps in *E. typhina* and *E. clarkii* (Fisher's exact test: $p < 0.01$).

GO ID	GO term	Significant category	Genome-wide annotations	Significant annotations	Enrichment *p*-value
E. typhina					
GO:0006468	Protein phosphorylation	All sweeps CLR	148	18	6.5E-05
GO:0035556	Intracellular signal transduction	All sweeps iHS-CLR-overlap	30	2	5.4E-03
GO:0043604	Amide biosynthetic process	All sweeps CLR	127	8	8.0E-03
GO:0055114	Oxidation–reduction process	All sweeps iHS-CLR-overlap[a]	349	18	1.6E-05
GO:0006468	Protein phosphorylation	AT-overlap CLR and iHS	148	15/7	1.2E-04/2.0E-03
GO:0043603	Cellular amide metabolic process	AT-overlap iHS	143	6	4.3E-03
E. clarkii					
GO:0006468	Protein phosphorylation	All sweeps CLR	147	21	2.5E-04
GO:0043386	Mycotoxin biosynthetic process	All sweeps CLR	15	5	1.7E-03
GO:0042398	Cellular-modified amino acid biosynthetic process	All sweeps iHS	12	2	7.2E-03
GO:0051649	Establishment of localization in cell	All sweeps iHS	81	4	9.9E-03
GO:0006468	Protein phosphorylation	AT-overlap CLR	147	20	3.4E-04
GO:0043386	Mycotoxin biosynthetic process	AT-overlap CLR	15	5	1.4E-03
GO:0051649	Establishment of localization in cell	AT-overlap iHS	81	4	1.6E-03

Note: Significant category indicates the sweep subsets on which enrichment tests were performed; genome-wide annotations indicate the total number of genes annotated to the respective GO term and significant annotations indicate how many of those genes were located in selective sweeps.
[a]Also significant in CLR sweeps: 25 annotations ($p = 9.4E-03$) and iHS sweeps: 17 annotations ($p = 3.7E-02$.

orthologous locus in *E. typhina* (Ety_004220) also showed evidence of selection in one population (So). Interestingly, Ecl_000006 was significantly upregulated *in planta* (1.2 $\log_2$-fold-change [logFC]) whereas the ortholog Ety_004220 was significantly downregulated *in planta* (5.0 logFC). We investigated the haplotype structure of these loci in both species (Figure 4, Figure S7). In *E. clarkii* four populations in which the CLR scan identified a sweep were fixed for one haplotype in addition to some low-frequency variants. The MdR population was fixed for a different haplotype. An additional haplotype with little variation was present in low frequency in the Cev population and intermediate frequency in the So population. At the homologous locus, *E. typhina* carried two haplotypes at intermediate frequencies within populations. Additional candidate effectors are described in Text S1 and Figures S8 and S9.

3.6 | Genes with convergent signatures of selection among scan methods

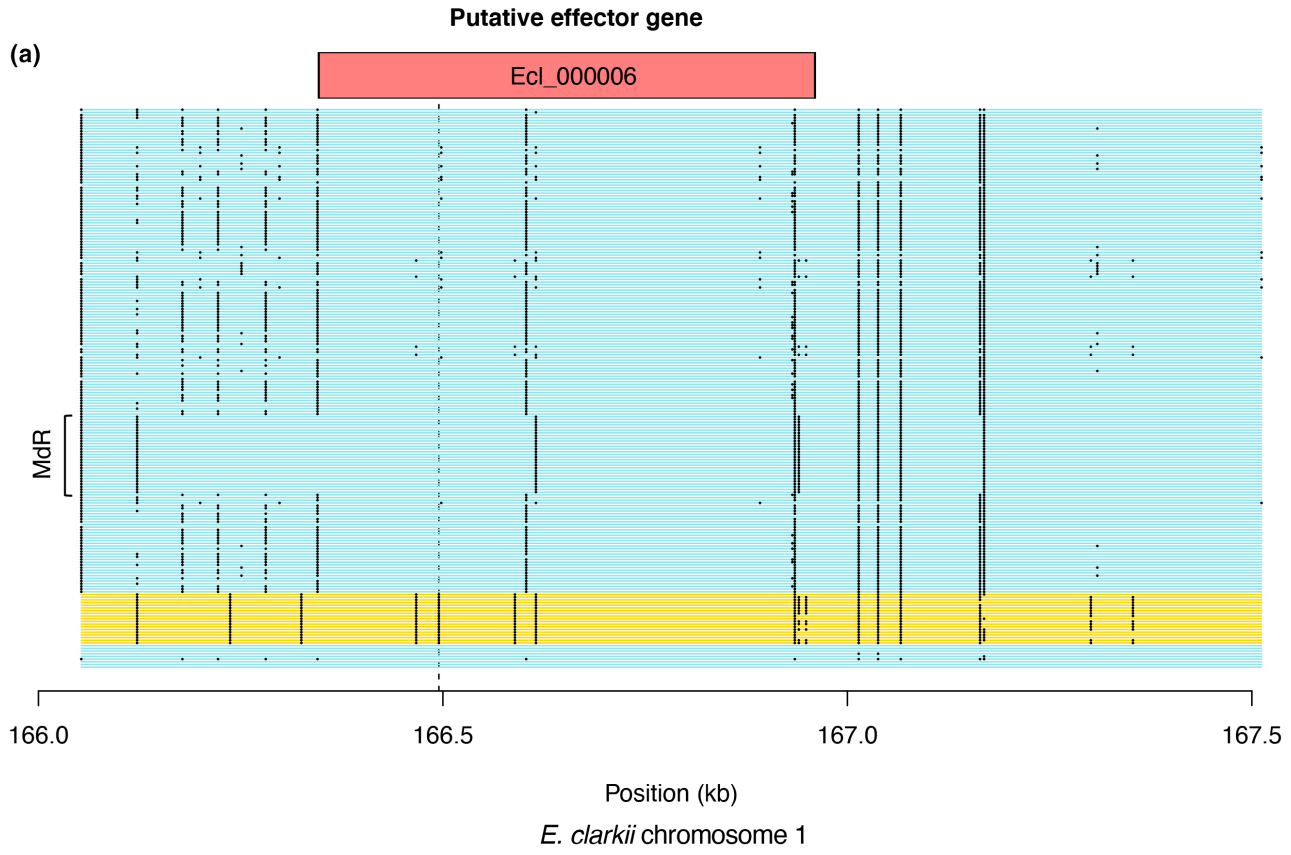
We found 21 regions with overlapping selective sweeps identified by both methods in *E. typhina* populations, which correspond to 15 distinct genomic regions given that some overlapping sweeps were shared among populations. Ten of these 15 regions encoded a total of 89 genes and 70 genes encoded conserved protein domain. No gene was predicted to be an effector, but one gene encoded a LysM effector domain (Ety_004233) and this gene showed evidence of selective sweeps in all populations (CLR sweeps in Aub, Auv, El, Kew, MdR overlapping with iHS sweeps in Cev, Kew, MdR, So). In three populations, the sweep regions were fully overlapping an AT-rich region containing Ety_004235 with homology to an alkaloid biosynthesis gene (Figure S10). Within *E. clarkii* populations, we only detected two regions with sweep signatures inferred by both methods. One region contained two genes encoding pectate lyase superfamily proteins (PFAM:PF12708) and a third encoding a protein with shared similarity to a LysM domain-containing protein in *Metarhizium acridum* (84% query cover and 68% identity). None of the genes were significantly up- or downregulated *in planta*. We investigated the cross-population haplotype structure at this locus (Figures S11 and S12). Populations for which the iHS scan inferred a sweep contained a shared haplotype at intermediate frequencies and three other haplotypes were present at different frequencies in different populations. The homologous region in *E. typhina* showed no evidence for a selective sweep (Figure S13: Cross-population haplotype structure for homologous locus not under selection on chromosome four in *E. typhina*).

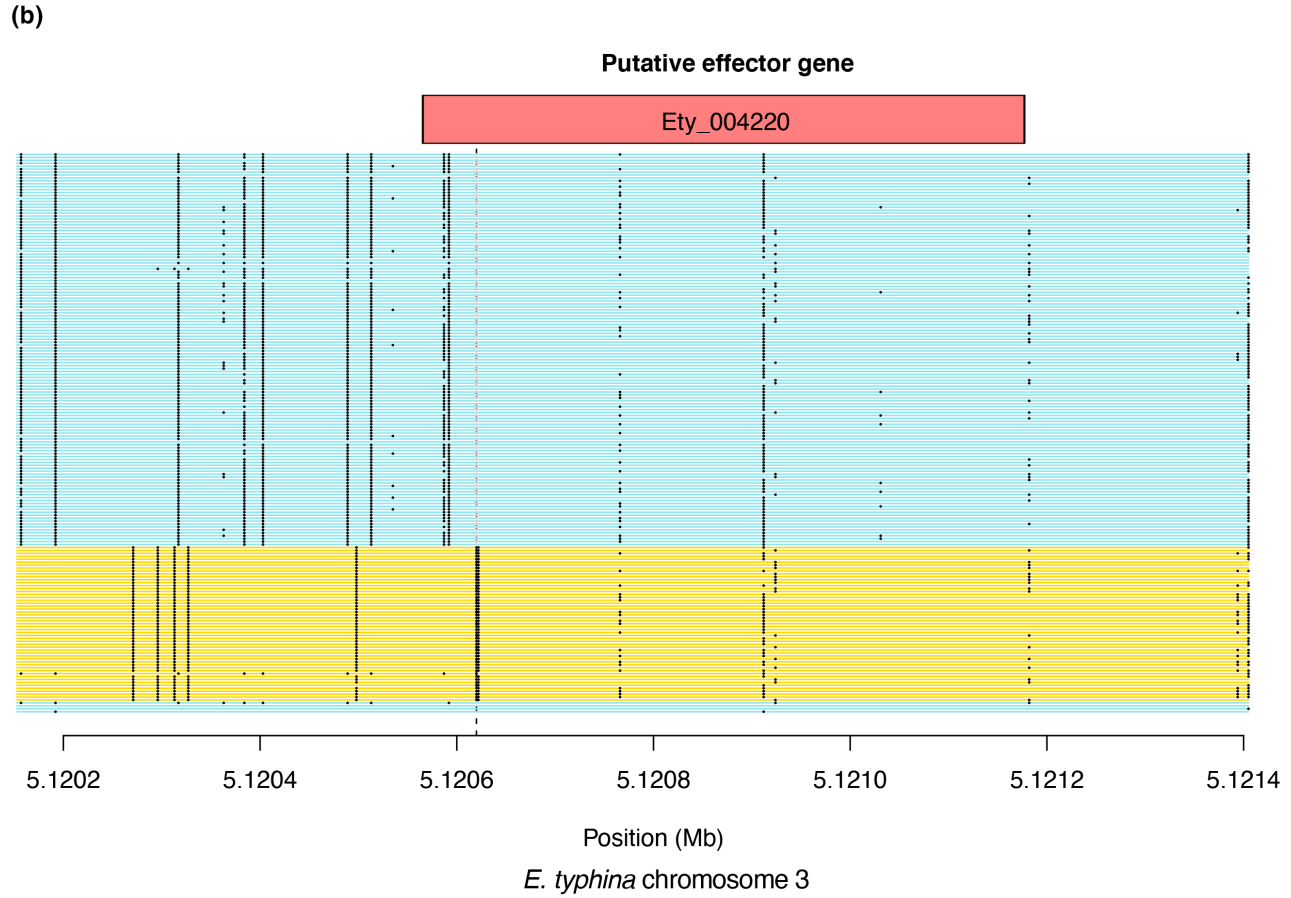
4 | DISCUSSION

4.1 | Contrasting signatures of selection between species

We identified genome-wide signatures of selective sweeps in two species sampled across seven sympatric locations in Europe. Selective sweeps were widespread across the genomes and were more abundant in *E. typhina* compared with its sibling species *E. clarkii*. The lower number of selective sweeps in *E. clarkii* populations agrees with expectations based on the species' demography—lower rates of gene flow, lower genetic diversity and smaller effective population sizes (Treindl et al., 2023). Due to the lower effective population size, selection is less efficient and drift more pervasive, so alleles are more likely to become fixed or be lost by chance. Selective sweeps were also rarer in the anther smut fungus *Microbotryum silenes-dioicae* compared with its sister species *M. lychnidis-dioicae* (Badouin et al., 2017) consistent with differences in effective population size. Similarly, genome scans identified fewer sweeps in less genetically diverse field populations of the wheat pathogen *Z. tritici* (Hartmann et al., 2018).

FIGURE 4 Cross-population haplotype structure for candidate effector loci. (a) Locus of a predicted effector gene (Ecl_000006) in *E. clarkii* chromosome 1. Haplotypes for all isolates from all populations are shown in rows ($n = 190$) and SNPs in columns, where black dots represent derived alleles. Haplotypes are coloured relative to the first nonsynonymous SNP position (dashed line) located in the genic region of Ecl_000005 (166,495; blue = A, yellow = G. The depicted haplotype block is located in a CLR-sweep region and shared by six populations (except MdR). The gene was differentially expressed *in planta* and lies in a relatively short GC-rich region (4982 bp) sandwiched between two AT-rich regions at the start of chromosome 1 and contains two other predicted genes besides the effector, one of them encoding a protein kinase domain (Ecl_000004). Blue haplotypes at the bottom of the plot indicate nine individuals that have not been genotyped at this locus. Three different haplotypes can be distinguished at this locus: The most common blue haplotype is present in all populations except MdR where isolates are fixed for a different haplotype (also blue) as indicated in the plot. The yellow haplotype is present at intermediate frequency in So and at low frequency in Cev (see Figure S7a showing haplotypes at this locus for each population separately). The homologous locus was also located within a selective sweep in *E. typhina*. (b) Locus of a predicted effector gene (Ety_004220) in *E. typhina* chromosome 3. Haplotypes for all isolates from all populations are shown in rows ($n = 175$) and SNPs in columns, where black dots represent derived alleles. Haplotypes are coloured relative to the second nonsynonymous SNP position (dashed line) located in the genic region of Ety_004220 (5,120,620; blue = A, yellow = G) are shown. The depicted haplotype block is located in a CLR-sweep region in one population (So) which was also under selection in *E. clarkii*. The gene was significantly downregulated *in planta* and lies in a longer GC-rich region (31.5 kb) at the end of chromosome 3 (which is the homologous inversion of *E. clarkii* chromosome 1). Blue haplotypes at the bottom of the plot indicate four individuals that have not been genotyped at this locus. Two different haplotypes can be distinguished at this locus, both of which are present at different frequencies in all seven populations (see Figure S7b showing haplotypes at this locus for each population separately). The population for which the CLR scan inferred a sweep (So) is almost fixed for the blue haplotype with some low-frequency variants and one isolate of yellow haplotype present.





4.2 | Different methods detected mostly distinct sweeps

Overall, the concordance between selective sweeps detected by different genome scan methods in our study was relatively low and consistent with other studies (Hartmann et al., 2018; Pavlidis & Alachiotis, 2017). The two scan methods assess different signatures of selection and have increased power at different time points relative to the onset of selection: that is, SFS-based scans such as CLR perform better at detecting selected variants close to or after fixation (young sweeps), while LD-based scans such as iHS are best suited to identify ongoing sweeps with alleles at intermediate frequencies (Vitti et al., 2013; Weigand & Leese, 2018). By combining two complementary methods, we were able to provide a more integrated evaluation of the impact of selection in *Epichloe* genomes.

4.3 | Evidence of recent selection was more common in repeat-rich regions and largely population-specific

Sweep regions were not randomly distributed across *Epichloe* genomes but were more frequently overlapping AT-rich regions. Such repeat-rich genome compartments play a role in the rapid adaptation of several fungal plant pathogens. 'Two-speed' genomes in the form of accessory chromosomes [e.g. in *Fusarium oxysporum* (Ma et al., 2010) and *Z. tritici* (Plissonneau et al., 2018)], or scattered mosaic-like sequence blocks such as in *Leptosphaeria maculans* (Grandaubert et al., 2014; Rouxel et al., 2011; Van de Wouw et al., 2010) and *Verticillium dahliae* (Faino et al., 2016) harbour substantial structural variation and encode important effector genes. However, the signatures of selection in or nearby genomic regions have rarely been investigated mostly because characterization of genetic variation in highly polymorphic regions presents particular challenges (Eschenbrenner et al., 2020). This is largely because there are limitations to the accuracy of variant discovery based on the alignment of short reads in highly polymorphic repeat-rich regions (Pfeifer, 2017; Treangen & Salzberg, 2012). The more abundant polymorphism may result from higher mutation rates but also erroneous mapping and may bias inferences of selective sweeps positively by making outliers more prominent but also negatively by breaking down sweep signatures. Furthermore, which regions can be resolved strongly depends on the reference sequence and structural variation such as presence or absence polymorphisms. For example, when a particular AT-rich region is present in our reference genome but missing in some mapped individuals or vice versa, it can lead to erroneous inferences of SNPs. In order to correct these issues and achieve an accurate evaluation of genetic variation of high-confidence loci across all genomic regions, we developed a thorough variant filtering procedure. First, SNPs from ambiguous loci detectable as false 'heterozygous' sites were excluded. Second, we employed stringent filtering criteria that limited our downstream analyses to high-confidence SNPs with similar genotyping rates across all populations, thereby excluding genomic regions with large-scale presence/absence polymorphism (see Treindl et al., 2023 for detail). Finally, for the genome scans performed in this study we included only conserved SNPs that had high mapping quality of the outgroup species (see Methods section: 2.2 Variant detection and data polarization). In our previous study, we further demonstrate that the genetic structure among populations based on SNPs called in the conserved gene-rich regions is also resolved when including only the SNPs called in the dynamic AT-rich regions (see Treindl et al., 2023 Supplementary Materials). Here, we find more selective sweeps in or in proximity to repeat-rich genomic regions. While repeat-rich regions themselves contain only few genes and we did not detect a significant enrichment of candidate effector genes, there are many genes in the bordering regions, including interesting candidates for genes encoding pathogenicity-related functions. This pattern was consistent across the two species and all populations. Thus, although regions were not necessarily shared between populations, there was evidence that repeat-rich regions are consistently under selection.

Selective sweeps in *E. typhina* and *E. clarkii* were predominantly population-specific, suggesting that selection and neutral population genetic processes varied in space. This pattern could be explained by spatially variable selection as would be expected in heterogeneous environments such as natural ecosystems (Thompson, 2005), selection acting on population-specific events like RIP on TE bursts, asynchronous coevolutionary cycles, or neutral population processes and demographic history. Considering spatially variable selection first, variable selection pressures could result from differences in the genetic makeup of host populations and arms race coevolution with local host genotypes. In addition, adaptation to different abiotic environmental conditions independent of the host plant such as fluctuations in humidity or temperature could have played role. Environmental heterogeneity including both biotic and abiotic factors, reduced levels of gene flow and selection varying in space can drive coevolutionary dynamics (Amandine et al., 2022). At the local scale, such selection regimes can generate locally adapted populations (Croll & McDonald, 2016; Gandon & Nuismer, 2009; Laine, 2005). Empirical studies revealed in such scenarios that distinct loci can show signatures of selection among populations across a wide range of fungal pathogen systems, including in agricultural and natural ecosystems (Badouin et al., 2017; Branco et al., 2017; Ellison et al., 2011; Hartmann et al., 2018; Mohd-Assaad et al., 2018; Persoons et al., 2022). Alternatively, targets of selection could be the result of more stochastic events, including a population-specific burst of TEs and subsequent RIP. If RIP silencing is advantageous in one population, then this could result in little overlap in selected loci between populations. It could also produce the pattern we observed that selective sweeps were more frequent in proximity to repeat-rich regions more prone to RIP.

Despite most sweeps being unique to populations, we found some selective sweeps shared among multiple populations and large geographic distances. This could be the result of shared selection regimes stemming from shared coevolutionary dynamics at a larger

scale. However, shared signatures of selective sweeps could also reflect a process that occurred in the ancestral population and was retained after the populations split. This could be the result of selection in the past, neutral genetic processes or stochastic events. The findings of both unique and shared regions under selection across populations may also be explained by asynchronous coevolutionary cycles with different populations being at different stages (Amandine et al., 2022). Complex signatures of selection regimes underscore the multidimensional nature of the coevolutionary process and were called 'the geographic mosaic of coevolution' by Thompson (2005). Antagonistic selection pressures are asynchronized over space and time as they act locally within populations but are affected by among-population processes depending on the underlying species distribution, their dispersal rates (migration) and levels of gene flow (Ebert & Fields, 2020; Gandon & Nuismer, 2009; Smith et al., 2011). For example, a beneficial new mutation enabling the infection of a previously resistant host genotype may first appear in one local pathogen population, sweep to high frequency and given sufficient levels of gene flow, spread to other populations where the same host genotype is present. Such joint increases in frequency of beneficial mutations can generate shared signatures of selective sweeps across populations. Population-specific sweeps in this model would be generated by locally adaptive alleles and fail to spread to other populations. Given the joint effects of biotic (e.g. host genotype) and abiotic factors driving selection across a pathogen species range, disentangling coevolutionary dynamics is challenging (Amandine et al., 2022). Understanding whether coevolution between a pathogen with its host is driven by arms race or trench warfare dynamics would require first ascertaining whether a particular locus is indeed selected due to host–pathogen interactions.

4.4 | Candidate genes involved in recent pathogen adaptation

Selective sweeps contained genes encoding a broad range of different biological functions with a notable enrichment in protein phosphorylation in both species. Protein kinases catalyse the phosphorylation of proteins thereby acting as regulators of protein activity and gene expression. In fungal pathogens, these proteins govern a variety of cellular pathways related to plant infection, including processes involved in cell surface remodelling to avoid recognition by the host, differentiation of infection-related structures and invasive growth (Turrà et al., 2014). A protein kinase (sakA) in *E. festucae* has also been shown to be important for the regulation of hyphal growth *in planta* and maintenance of a mutualistic interaction with the host plant *Lolium perenne*, while a deletion of this gene leads to pathogenic interactions (Eaton et al., 2010). Our findings suggest that polymorphism near genes encoding protein kinases are often targets of recent selection and that protein kinases are important components of *Epichloe* infection strategies. Additional loci under recent selection encoded genes related to the biosynthesis of secondary metabolites. Genes involved in the metabolism of mycotoxins including alkaloids and sterols were enriched in selective sweeps in *E. clarkii* but not *E. typhina*, and mycotoxin genes were also highly expressed *in planta* in both species (Table S7). An extensive body of research has highlighted the structural diversity of alkaloid genes across the *Epichloe* genus and specifically emphasized the importance of alkaloids produced predominantly by nonpathogenic *Epichloe* in providing benefits to the host plant as herbivore deterrents (Schardl, Florea, et al., 2013; Young et al., 2015). Therefore, alkaloids were called the 'currencies maintaining mutualistic symbiosis' (Schardl, Young, et al., 2013).

Beyond functions related to protein regulation and metabolite production, we also identified a set of candidate effector genes encoded in selective sweep regions. Effectors play broad roles manipulating host physiology and can mediate pathogen recognition (Lo Presti et al., 2015). Interestingly, several predicted effectors encoded conserved functions such as the degradation of cell wall components including a cutinase which can facilitate penetration of the plant cuticle (Chen et al., 2013), pectate lyases and a pectinesterase. Multiple selective sweep regions encoded predicted effector proteins with CVNH domains. A further gene encoding a CVNH effector (Ety_004758/Ecl_005079) was identified as a candidate underpinning host specialization showing elevated levels of differentiation between *E. typhina* and *E. clarkii* (Schirrmann et al., 2018). CVNH effectors also promote infection and growth in the stem rot pathogen *Sclerotinia sclerotiorum* (Lyu et al., 2015; Seifbarghi et al., 2017). Finally, we identified several genes encoding LysM domains, which have conserved roles in plant pathogens manipulating chitin-triggered recognition during plant invasion (de Jonge & Thomma, 2009). Even though we identified a broad range of different gene functions encoded in selective sweep regions, a substantial portion seems at least indirectly involved in host–pathogen interactions.

Selective sweep regions identified through genome scans typically identify many different gene functions but inferring selection pressures on populations from such approaches remains challenging. An obvious caveat is that most gene functions can only be inferred by homology biasing interpretations towards evolutionarily well-conserved functions. Furthermore, interpreting gene functions is often influenced by a priori knowledge of the species' biology (e.g. pathogens) to infer plausible mechanisms (Pavlidis et al., 2012). Hence, biological interpretations of selection scan results should be treated cautiously in the absence of independent approaches such as understanding gene functions through experimentation or assessing pathogenicity trait variation among the analysed individuals.

Pathogenicity is a complex and rapidly evolving trait expressed in the context of a given host environment. Through our widespread sampling, we could study the shared and distinct signatures of selection across two species and across multiple populations giving our study unique spatial and taxonomic breadth. Our results demonstrate that selective sweeps in the genomes of *Epichloe* sibling species are widespread, spatially variable and may involve many different genes. Our study also demonstrates that selection is not uniform across the genome, but is biased to AT-rich regions,

suggesting that these dynamic genome compartments are important for adaptive evolution in *Epichloe* species. Together these data provide a comprehensive genomic perspective on the complexity of host–pathogen interactions in natural populations of fungal plant pathogens.

AUTHOR CONTRIBUTIONS

Artemis D. Treindl contributed to conceptualization (equal); formal analysis (equal); methodology (equal); writing—original draft preparation (lead); and writing—review and editing (equal). *Jessica Stapley* contributed to formal analysis (equal); methodology (equal); and writing—review and editing (equal). *Daniel Croll* contributed to writing—review and editing (equal) and supervision (equal). *Adrian Leuchtmann* contributed to conceptualization (equal); writing—review and editing (equal); supervision (equal); and funding acquisition.

ACKNOWLEDGEMENTS

We would like to kindly thank Claudia Michel and Beatrice Arnold for laboratory assistance and Niklaus Zemp for bioinformatic support and three anonymous reviewers for their helpful comments. Data presented and analysed here were generated in collaboration with the Genetic Diversity Centre (GDC), ETH Zurich and the Functional Genomics Center Zurich (FGCZ), UZH.

FUNDING INFORMATION

Swiss National Science Foundation (SNF), grant Number: 31003A_169269.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Genome assemblies are available on GenBank (Accession nos. PRJNA533210 PRJNA533212). Genome sequencing data have been deposited in the European Nucleotide Archive ENA (Accession no. PRJEB59262). Additional figures of selection scans are available on Zenodo (10.5281/zenodo.7573676).

BENEFIT-SHARING STATEMENT

Benefits from this research accrue from the sharing of our data and results on public databases as described above.

ORCID

Artemis D. Treindl https://orcid.org/0000-0001-7347-2891
Daniel Croll https://orcid.org/0000-0002-2072-380X
Adrian Leuchtmann https://orcid.org/0000-0002-9070-0902

REFERENCES

Aguileta, G., Refrégier, G., Yockteng, R., Fournier, E., & Giraud, T. (2009). Rapidly evolving genes in pathogens: Methods for detecting positive selection and examples among fungi, bacteria, viruses and protists. *Infection, Genetics and Evolution*, *9*, 656–670.

Alexa, A., & Rahnenführer, J. (2009). Gene set enrichment analysis with topGO. *Bioconductor Improv*, *27*, 1–26.

Amandine, C., Ebert, D., Stukenbrock, E., Rodríguez de la Vega, R. C., Tiffin, P., Croll, D., & Tellier, A. (2022). Unraveling coevolutionary dynamics using ecological genomics. *Trends in Genetics*, *38*, 1003–1012.

Ashby, B., & Gupta, S. (2014). Parasitic castration promotes coevolutionary cycling but also imposes a cost on sex. *Evolution*, *68*, 2234–2244.

Badouin, H., Gladieux, P., Gouzy, J., Siguenza, S., Aguileta, G., Snirc, A., Le Prieur, S., Jeziorski, C., Branca, A., & Giraud, T. (2017). Widespread selective sweeps throughout the genome of model plant pathogenic fungi and identification of effector candidates. *Molecular Ecology*, *26*, 2041–2062.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., & Sonnhammer, E. L. L. (2002). The Pfam protein families database. *Nucleic Acids Research*, *30*, 276–280.

Branco, S., Bi, K., Liao, H. L., Gladieux, P., Badouin, H., Ellison, C. E., Nguyen, N. H., Vilgalys, R., Peay, K. G., Taylor, J. W., & Bruns, T. D. (2017). Continental-level population differentiation and environmental adaptation in the mushroom *Suillus brevipes*. *Molecular Ecology*, *26*, 2063–2076.

Brockhurst, M. A., Chapman, T., King, K. C., Mank, J. E., Paterson, S., & Hurst, G. D. D. (2014). Running with the red queen: The role of biotic conflicts in evolution. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20141382.

Brown, J. K. M., & Tellier, A. (2011). Plant-parasite coevolution: Bridging the gap between genetics and ecology. *Annual Review of Phytopathology*, *49*, 345–367.

Bultman, T. L., & Leuchtmann, A. (2008). Biology of the *Epichloë-Botanophila* interaction: An intriguing association between fungi and insects. *Fungal Biology Reviews*, *22*, 131–138.

Charlesworth, B. (2009). Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*, 195–205.

Charlesworth, D. (2006). Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics*, *2*, 379–384.

Chen, S., Su, L., Chen, J., & Wu, J. (2013). Cutinase: Characteristics, preparation, and application. *Biotechnology Advances*, *31*, 1754–1767.

Chung, K.-R., & Schardl, C. L. (1997). Sexual cycle and horizontal transmission of the grass symbiont, *Epichloë typhina*. *Mycological Research*, *101*, 295–301.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, *6*, 80–92.

Clay, K., & Schardl, C. L. (2002). Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. *The American Naturalist*, *160*, S99–S127.

Croll, D., & McDonald, B. A. (2012). The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathogens*, *8*, e1002608.

Croll, D., & McDonald, B. A. (2016). The genetic basis of local adaptation for pathogenic fungi in agricultural ecosystems. *Molecular Ecology*, *26*, 2027–2040.

de Jonge, R., & Thomma, B. P. H. J. (2009). Fungal LysM effectors: Extinguishers of host immunity? *Trends in Microbiology*, *17*, 151–157.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*, 15–21.

Dong, S., Raffaele, S., & Kamoun, S. (2015). The two-speed genomes of filamentous pathogens: Waltz with plants. *Current Opinion in Genetics and Development*, *35*, 57–65.

Eaton, C. J., Cox, M. P., Ambrose, B., Becker, M., Hesse, U., Schardl, C. L., & Scott, B. (2010). Disruption of signaling in a fungal-grass symbiosis leads to pathogenesis. *Plant Physiology*, *153*, 1780–1794.

Ebert, D., & Fields, P. D. (2020). Host–parasite co-evolution and its genomic signature. *Nature Reviews Genetics*, *21*, 754–768.

Ellison, C. E., Hall, C., Kowbel, D., Welch, J., Brem, R. B., Glass, N. L., & Taylor, J. W. (2011). Population genomics and local adaptation in wild isolates of a model microbial eukaryote. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 2831–2836.

Eschenbrenner, C. J., Feurtey, A., & Stukenbrock, E. H. (2020). Population genomics of fungal plant pathogens and the analyses of rapidly evolving genome compartments. In J. Y. Dutheil (Ed.), *Statistical population genomics*, *methods in molecular biology* (pp. 337–355). Humana.

Faino, L., Seidl, M. F., Shi-Kunne, X., Pauper, M., Van Den Berg, G. C. M., Wittenberg, A. H. J., & Thomma, B. P. H. J. (2016). Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. *Genome Research*, *26*, 1091–1100.

Franceschetti, M., Maqbool, A., Jiménez-Dalmaroni, M. J., Pennington, H. G., Kamoun, S., & Banfield, M. J. (2017). Effectors of filamentous plant pathogens: Commonalities amid diversity. *Microbiology and Molecular Biology Reviews*, *81*, e00066-16.

Gandon, S., & Michalakis, Y. (2002). Local adaptation, evolutionary potential and host-parasite coevolution: Interactions between migration, mutation, population size and generation time. *Journal of Evolutionary Biology*, *15*, 451–462.

Gandon, S., & Nuismer, S. L. (2009). Interactions between genetic drift, gene flow, and selection mosaics drive parasite local adaptation. *American Naturalist*, *173*, 212–224.

Gautier, M., Klassmann, A., & Vitalis, R. (2017). Rehh 2.0: A reimplementation of the R package rehh to detect positive selection from haplotype structure. *Molecular Ecology Resources*, *17*, 78–90.

Gel, B., Díez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M. A., & Malinverni, R. (2016). regioneR: An R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics*, *32*, 289–291.

Gladieux, P., Ropars, J., Badouin, H., Branca, A., Aguileta, G., De Vienne, D. M., Rodríguez de la Vega, R. C., Branco, S., & Giraud, T. (2014). Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Molecular Ecology*, *23*, 753–773.

Grandaubert, J., Lowe, R. G. T., Soyer, J. L., Schoch, C. L., Van De Wouw, A. P., Fudal, I., Robbertse, B., Lapalu, N., Links, M. G., Ollivier, B., Linglin, J., Barbe, V., Mangenot, S., Cruaud, C., Borhan, H., Howlett, B. J., Balesdent, M. H., & Rouxel, T. (2014). Transposable element-assisted evolution and adaptation to host plant within the *Leptosphaeria maculans-Leptosphaeria biglobosa* species complex of fungal pathogens. *BMC Genomics*, *15*, 1–27.

Hartmann, F. E., McDonald, B. A., & Croll, D. (2018). Genome-wide evidence for divergent selection between populations of a major agricultural pathogen. *Molecular Ecology*, *27*, 2725–2741.

Kubicek, C. P., Starr, T. L., & Glass, N. L. (2014). Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. *Annual Review of Phytopathology*, *52*, 427–451.

Kurtz, J., Schulenburg, H., & Reusch, T. B. H. (2016). Host–parasite coevolution—rapid reciprocal adaptation and its genetic basis. *Zoology*, *119*, 241–243.

Laine, A. L. (2005). Spatial scale of local adaptation in a plant-pathogen metapopulation. *Journal of Evolutionary Biology*, *18*, 930–938.

Leuchtmann, A. (2003). Taxonomy and diversity of *Epichloë* endophytes. In J. F. White Jr., C. W. Bacon, N. L. Hywel-Jones, & J. W. Spatafora (Eds.), *Clavicipitalean fungi: Evolutionary biology, chemistry, biocontrol and cultural impacts* (pp. 169–194). Marcel Dekker.

Leuchtmann, A., & Clay, K. (1997). The population biology of grass endophytes. In G. C. Carroll & P. Tudzynski (Eds.), *The Mycota*, *Vol. V. Part B*, *Plant Relationships* (pp. 185–202). Springer-Verlag.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*, *47*, e47.

Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., & Kahmann, R. (2015). Fungal effectors and plant susceptibility. *Annual Review of Plant Biology*, *66*, 513–545.

Lyu, X., Shen, C., Fu, Y., Xie, J., Jiang, D., Li, G., & Cheng, J. (2015). Comparative genomic and transcriptional analyses of the carbohydrate-active enzymes and secretomes of phytopathogenic fungi reveal their significant roles during infection and development. *Scientific Reports*, *5*, 15565.

Ma, L.-J., Van Der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M.-J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., & Henrissat, B. (2010). Comparative analysis reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature*, *464*, 367–373.

McDonald, B. A., & Stukenbrock, E. H. (2016). Rapid emergence of pathogens in agro-ecosystems: Global threats to agricultural sustainability and food security. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*, 20160026.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., & Daly, M. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*, 1297–1303.

Mohd-Assaad, N., McDonald, B. A., & Croll, D. (2018). Genome-wide detection of genes under positive selection in worldwide populations of the barley scald pathogen. *Genome Biology and Evolution*, *10*, 1315–1332.

Möller, M., & Stukenbrock, E. H. (2017). Evolution and genome architecture in fungal plant pathogens. *Nature Reviews Microbiology*, *15*, 756–771.

Nielsen, R., Williamson, S., Kim, Y., Hubisz, M. J., Clark, A. G., & Bustamante, C. (2005). Genomic scans for selective sweeps using SNP data. *Genome Research*, *15*, 1566–1575.

Papkou, A., Guzella, T., Yang, W., Koepper, S., Pees, B., Schalkowski, R., Barg, M. C., Rosenstiel, P. C., Teotónio, H., & Schulenburg, H. (2019). The genomic basis of red queen dynamics during rapid reciprocal host–pathogen coevolution. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 923–928.

Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., Quail, M., Smith, F., Walker, D., Libberton, B., Fenton, A., Hall, N., & Brockhurst, M. A. (2010). Antagonistic coevolution accelerates molecular evolution. *Nature*, *464*, 275–278.

Pavlidis, P., & Alachiotis, N. (2017). A survey of methods and tools to detect recent and strong positive selection. *Journal of Biological Research-Thessaloniki*, *24*, 7.

Pavlidis, P., Jensen, J. D., Stephan, W., & Stamatakis, A. (2012). A critical assessment of storytelling: Gene ontology categories and the importance of validating genomic scans. *Molecular Biology and Evolution*, *29*, 3237–3248.

Pavlidis, P., Živković, D., Stamatakis, A., & Alachiotis, N. (2013). SweeD: Likelihood-based detection of selective sweeps in thousands of genomes. *Molecular Biology and Evolution*, *30*, 2224–2234.

Persoons, A., Hayden, K. J., Fabre, B., Frey, P., De Mita, S., Tellier, A., & Halkett, F. (2017). The escalatory red queen: Population extinction and replacement following arms race dynamics in poplar rust. *Molecular Ecology*, *26*, 1902–1918.

Persoons, A., Maupetit, A., Louet, C., Andrieux, A., Lipzen, A., Barry, K. W., Na, H., Adam, C., Grigoriev, I. V., Segura, V., Duplessis, S., Frey, P., Halkett, F., & De Mita, S. (2022). Genomic signatures of a major adaptive event in the pathogenic fungus *Melampsora larici-populina*. *Genome Biology and Evolution*, *14*(1), evab279.

Pfeifer, S. P. (2017). From next-generation resequencing reads to a high-quality variant data set. *Heredity*, *118*, 111–124.

Plissonneau, C., Benevenuto, J., Mohd-Assaad, N., Fouché, S., Hartmann, F. E., & Croll, D. (2017). Using population and comparative genomics to understand the genetic basis of effector-driven fungal pathogen evolution. *Frontiers in Plant Science*, *8*, 1–15.

Plissonneau, C., Hartmann, F. E., & Croll, D. (2018). Pangenome analyses of the wheat pathogen *Zymoseptoria tritici* reveal the structural basis of a highly plastic eukaryotic genome. *BMC Biology*, *16*, 5.

Raffaele, S., & Kamoun, S. (2012). Genome evolution in filamentous plant pathogens: Why bigger can be better. *Nature Reviews Microbiology*, *10*, 417–430.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*, 139–140.

Rouxel, T., Grandaubert, J., Hane, J. K., Hoede, C., van de Wouw, A. P., Couloux, A., Dominguez, V., Anthouard, V., Bally, P., Bourras, S., Cozijnsen, A. J., Ciuffetti, L. M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin, S. B., Gout, L., Glaser, N., … Howlett, B. J. (2011). Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by repeat-induced point mutations. *Nature Communications*, *2*, 202.

Schardl, C. L., Florea, S., Pan, J., Nagabhyru, P., Bec, S., & Calie, P. J. (2013). The epichloae: Alkaloid diversity and roles in symbiosis with grasses. *Current Opinion in Plant Biology*, *16*, 480–488.

Schardl, C. L., Young, C. A., Pan, J., Florea, S., Takach, J. E., Panaccione, D. G., Farman, M. L., Webb, J. S., Jaromczyk, J., Charlton, N. D., Nagabhyru, P., Chen, L., Shi, C., & Leuchtmann, A. (2013). Currencies of mutualisms: Sources of alkaloid genes in vertically transmitted epichloae. *Toxins*, *5*, 1064–1088.

Schirawski, J., Mannhaupt, G., Münch, K., Brefort, T., Schipper, K., Doehlemann, G., Di Stasio, M., Rössel, N., Mendoza-Mendoza, A., Pester, D., Müller, O., Winterberg, B., Meyer, E., Ghareeb, H., Wollenberg, T., Münsterkötter, M., Wong, P., Walter, M., Stukenbrock, E., … Kahmann, R. (2010). Pathogenicity determinants in smut fungi revealed by genome comparison. *Science*, *330*, 1546–1548.

Schirrmann, M. K., Zoller, S., Croll, D., Stukenbrock, E. H., Leuchtmann, A., & Fior, S. (2018). Genomewide signatures of selection in *Epichloë* reveal candidate genes for host specialization. *Molecular Ecology*, *27*, 3070–3086.

Seidl, M. F., & Thomma, B. P. H. J. (2017). Transposable elements direct the coevolution between plants and microbes. *Trends in Genetics*, *33*, 842–851.

Seifbarghi, S., Borhan, M. H., Wei, Y., Coutu, C., Robinson, S. J., & Hegedus, D. D. (2017). Changes in the *Sclerotinia sclerotiorum* transcriptome during infection of *Brassica napus*. *BMC Genomics*, *18*, 1–37.

Smith, D. L., Ericson, L., & Burdon, J. J. (2011). Co-evolutionary hot and cold spots of selective pressure move in space and time. *Journal of Ecology*, *99*, 634–641.

Stukenbrock, E. H., & Croll, D. (2014). The evolving fungal genome. *Fungal Biology Reviews*, *28*, 1–12.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*, 585–595.

Tellier, A., Moreno-Gámez, S., & Stephan, W. (2014). Speed of adaptation and genomic footprints of host–parasite coevolution under arms race and trench warfare dynamics. *Evolution*, *68*, 2211–2224.

Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., & Narechania, A. (2003). PANTHER: A library of protein families and subfamilies indexed by function. *Genome Research*, *13*, 2129–2141.

Thompson, J. N. (2005). *The geographic mosaic of coevolution*. University of Chicago Press.

Treangen, T. J., & Salzberg, S. L. (2012). Repetitive DNA and next-generation sequencing: Computational challenges and solutions. *Nature Reviews Genetics*, *13*, 36–46.

Treindl, A. D., Stapley, J., & Leuchtmann, A. (2021). Resolving the phylogeny of the *Epichloe typhina* species complex using genome-wide SNPs. Pages 177–192. *Into the wild – adaptation genomics of Epichloe grass pathogens in natural ecosystems*. Treindl, A. D., Diss. ETH No. 27662, Zürich. https://doi.org/10.3929/ethz-b-000521608

Treindl, A. D., Stapley, J., & Leuchtmann, A. (2023). Genetic diversity and population structure of *Epichloë* fungal pathogens of plants in natural ecosystems. *Frontiers in Ecology and Evolution*, *11*, 1129867. https://doi.org/10.3389/fevo.2023.1129867

Treindl, A. D., Stapley, J., Winter, D. J., Cox, M. P., & Leuchtmann, A. (2021). Chromosome-level genomes provide insights into genome evolution, organization and size in *Epichloe* fungi. *Genomics*, *113*, 4267–4275.

Turrà, D., Segorbe, D., & Di Pietro, A. (2014). Protein kinases in plant-pathogenic fungi: Conserved regulators of infection. *Annual Review of Phytopathology*, *52*, 267–288.

Van de Wouw, A. P., Cozijnsen, A. J., Hane, J. K., Brunner, P. C., McDonald, B. A., Oliver, R. P., & Howlett, B. J. (2010). Evolution of linked avirulence effectors in *Leptosphaeria maculans* is affected by genomic environment and exposure to resistance genes in host plants. *PLoS Pathogens*, *6*, e1001180.

van der Does, H. C., & Rep, M. (2017). Adaptation to the host environment by plant-pathogenic fungi. *Annual Review of Phytopathology*, *55*, 427–450.

Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013). Detecting natural selection in genomic data. *Annual Review of Genetics*, *47*, 97–120.

Weigand, H., & Leese, F. (2018). Detecting signatures of positive selection in non-model species using genomic data. *Zoological Journal of the Linnean Society*, *184*, 528–583.

White, J. F., Jr., & Bultman, T. L. (1987). Endophyte-host associations in forage grasses. VIII. Heterothallism in *Epichloë typhina*. *American Journal of Botany*, *74*, 1716–1721.

Winter, D. J., Ganley, A. R. D., Young, C. A., Liachko, I., Schardl, C. L., Dupont, P. Y., Berry, D., Ram, A., Scott, B., & Cox, M. P. (2018). Repeat elements organise 3D genome structure and mediate transcription in the filamentous fungus *Epichloë festucae*. *PLoS Genetics*, *14*, 1–29.

Woolhouse, M. E. J., Webster, J. P., Domingo, E., Charlesworth, B., & Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics*, *32*, 569–577.

Young, C. A., Schardl, C. L., Panaccione, D. G., Florea, S., Takach, J. E., Charlton, N. D., Moore, N., Webb, J. S., & Jaromczyk, J. (2015). Genetics, genomics and evolution of ergot alkaloid diversity. *Toxins*, *7*, 1273–1302.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Treindl, A. D., Stapley, J., Croll, D., & Leuchtmann, A. (2024). Two-speed genomes of *Epichloe* fungal pathogens show contrasting signatures of selection between species and across populations. *Molecular Ecology*, *33*, e17242. https://doi.org/10.1111/mec.17242