WILEY

Molecular Plant Pathology

ORIGINAL ARTICLE OPEN ACCESS

Comparative Genomics Reveals Sources of Genetic Variability in the Asexual Fungal Plant Pathogen *Colletotrichum lupini*

Joris A. Alkemade[1,2,3] 📧 | Pierre Hohmann[3,4] 📧 | Monika M. Messmer[3] 📧 | Timothy G. Barraclough[1,2] 📧

[1]Department of Biology, University of Oxford, Oxford, UK | [2]Calleva Research Centre for Evolution and Human Science, Magdalen College, Oxford, UK | [3]Department of Crop Sciences, Research Institute of Organic Agriculture (FiBL), Frick, Switzerland | [4]Department of Biology, Healthcare and the Environment, Faculty of Pharmacy and Food Sciences, Universitat de Barcelona, Barcelona, Spain

Correspondence: Joris A. Alkemade (alkemadejoris@gmail.com; joris.alkemade@magd.ox.ac.uk)

ABSTRACT

Fungal plant pathogens cause major crop losses worldwide, with many featuring compartmentalised genomes that include both core and accessory regions, which are believed to drive adaptation. The highly host-specific fungus *Colletotrichum lupini* greatly impacts lupin (*Lupinus* spp.) cultivation. This pathogen is part of clade 1 of the *C. acutatum* species complex and comprises four genetically uniform, presumably clonal, lineages (I–IV). Despite this, variation in virulence and morphology has been observed within these lineages. To investigate the potential sources of genetic variability in this asexual fungus, we compared the genomes of 16 *C. lupini* strains and 17 related *Colletotrichum* species. Phylogenomics confirmed the presence of four distinct lineages, but further examination based on genome size, gene content, transposable elements (TEs), and deletions revealed that lineage II could be split into two groups, II-A and II-B. TE content varied between lineages and correlated strongly with genome size variation, supporting a role for TEs in genome expansion in this species. Pangenome analysis revealed a highly variable accessory genome, including a minichromosome present in lineages II, III, and IV, but absent in lineage I. Accessory genes and effectors appeared to cluster in proximity to TEs. Presence/absence variation of putative effectors was lineage-specific, suggesting that these genes play a crucial role in determining host range. Notably, no effectors were found on the TE-rich minichromosome. Our findings shed light on the potential mechanisms generating genetic diversity in this asexual fungal pathogen that could aid future disease management.

1 | Introduction

The interaction between fungi and plants is ancient and spans a continuum from mutualistic to parasitic and from epiphytic to endophytic (Naranjo-Ortiz and Gabaldón 2020). To facilitate successful colonisation of a host, fungi secrete effector molecules to manipulate the plant immune system or ward off other microbes (Snelders, Rovenich, and Thomma 2022;

Cook, Mesarich, and Thomma 2015; Plett and Martin 2015; Redkar et al. 2022). For pathogenic fungi, there is strong evolutionary pressure to prevent host recognition (Möller and Stukenbrock 2017). Across eukaryotes, rapid evolution is generally enabled by sexual reproduction, which generates new gene combinations in a population (Tellier, Moreno-Gámez, and Stephan 2014). Yet, sexual reproduction in fungi is often rare, with one fifth of described species being presumed

predominantly asexual (Naranjo-Ortiz and Gabaldón 2020; Taylor, Jacobson, and Fisher 1999), including important plant pathogens such as *Fusarium oxysporum* (McTaggart et al. 2021; Ordonez et al. 2015), *Magnaporthe oryzae* (Latorre et al. 2023) and *Verticillium dahliae* (de Jonge et al. 2013). Understanding how they generate the variability required to keep pace with coevolving host immune defences is important for improving disease management.

Many filamentous plant pathogens have segmented genomes, often referred to as the two-speed genome (Dong, Raffaele, and Kamoun 2015), with effector genes mostly localised within highly variable repeat-rich regions, enabling fast diversification (Torres, Thomma, and Seidl 2021; van Westerhoven et al. 2024; Sánchez-Vallet et al. 2018). Pangenome analyses have shown that effector gene repertoires greatly fluctuate between closely related lineages and are often crucial for determining host range (Badet et al. 2020; van Dam et al. 2016; Le Naour-Vernet et al. 2023). Alongside diverse effector repertoires, many fungal genomes exhibit structural variation within species, such as the occurrence of large duplication and deletion events and varying numbers of accessory chromosomes (ACs; Todd, Forche, and Selmecki 2017; Möller and Stukenbrock 2017), highlighting the diversification potential of (asexual) plant-pathogenic fungi.

Colletotrichum species are notorious plant pathogens, causing devastating pre- and postharvest disease in numerous crops (Cannon et al. 2012). *Colletotrichum* is traditionally recognised as asexual (Cannon et al. 2012), but sexual reproduction has been observed in a few rare cases (Damm et al. 2012; Wilson et al. 2021; Rogério et al. 2022). Despite the rarity of sexual recombination, *Colletotrichum* species are highly diverse in both lifestyle and host range (Talhinhas and Baroncelli 2021). The mechanisms behind the emergence of this diversity are not yet fully understood (da Silva et al. 2020) but could include the mechanisms outlined above. For example, the evolution of ACs seems to play a significant role in host adaptation by serving as vehicles for horizontal gene transfer (HGT) to acquire novel pathogenicity genes (Wang et al. 2023). In *C. lentis* and *C. graminicola*, ACs were shown to be crucial to achieve full virulence in their respective hosts (Bhadauria et al. 2019; Ma et al. 2023). Transposable elements (TEs) were shown to generate genome variability and diversity in *C. higginsum* (Tsushima et al. 2019), a pathogen of several brassica species. Furthermore, TE activity in various *Colletotrichum* species influenced the gain and loss of gene families, which contributed to the adaptation of different lifestyles (Gan et al. 2016), such as host jumps from monocot to eudicot plants (Baroncelli et al. 2024).

Lupin anthracnose, caused by *C. lupini*, is hampering sustainable cultivation of the high-quality protein crops of white (*Lupinus albus*), blue (*L. angustifolius*), yellow (*L. luteus*), and Andean lupin (*L. mutabilis*). These crops have high potential to provide a locally produced alternative to soybean in temperate regions. *C. lupini* belongs to clade 1 of the *C. acutatum* species complex (CaSC), which harbours many devastating plant pathogens (Damm et al. 2012). In previous work, 67 *C. lupini* isolates, collected from lupin production areas across the world, were characterised through 3D-RAD sequencing (Alkemade

et al. 2023). Population genetics showed that, like many other CaSC species (Damm et al. 2012), *C. lupini* seems to be asexual, with four (I–IV) distinct and genetically uniform lineages displaying low admixture and high divergence (Alkemade et al. 2023). Despite the high genetic uniformity, differences in virulence and morphology were observed amongst and within those four lineages (Alkemade et al. 2021). All four lineages were present in the Andes region of South America, which is the presumed centre of origin of *C. lupini* (Riegel et al. 2010) and other CaSC clade 1 species (Bragança et al. 2016). Isolates found outside South America after 1990 all belong to lineage II, which contains strains that are highly aggressive on white and Andean lupin (Alkemade et al. 2021). The escape and global spread of a genetically uniform lineage, as seen for *C. lupini*, is typical for invasive fungi (Gladieux et al. 2016) and has been observed for numerous plant pathogens such as *Fusarium odoratissimum* on banana (van Westerhoven et al. 2022) and *M. oryzae* on wheat (Latorre et al. 2023). Yet, the potential genetic basis for differences within and between lineages in *C. lupini* remains unexplored.

In this study, we investigate inter- and intralineage genetic variability within *C. lupini* by sequencing the genomes of 16 strains, collected from five continents and 10 countries, representing all four lineages, and comparing them to each other and to other CaSC species. We quantify the extent to which genome structure, TE landscape, the gain or loss of effectors, and selection on different gene categories, in turn, contribute towards genetic diversity within and between *C. lupini* lineages. The findings should provide insights on the evolution and genetic variability of a putatively asexual fungal plant pathogen.

2 | Results

2.1 | Genomic Diversity Within the *C. acutatum* Species Complex

To study *C. lupini* evolution, 14 strains that were collected from across the globe and represented all four lineages (Figure S1) were sequenced and assembled using short-read data at an average of 141× coverage (Table S1). Two publicly available *C. lupini* long-read assemblies (CBS 109225 and IMI 504893, referred to as CLUP01 and CLUP02, respectively) were included as well (Baroncelli et al. 2021; Baroncelli et al. 2024). Seventeen publicly available genomes of species representing the five different clades of the CaSC were included to assess the genetic diversity of *C. lupini* in relation to the CaSC (Baroncelli et al. 2024). Although short-read assemblies were more fragmented (484–2499 contigs) than long-read assemblies (11–116 contigs), all *C. lupini* assemblies contained at least 96% of the single-copy BUSCO genes (Table S1), indicating nearly complete coverage of the conserved genes. *C. lupini* genome sizes ranged between 54.5 and 63.4 Mb (Figure 1b), corresponding to total K-mer counts (Table S1). The average *C. lupini* genome size of 59.2 Mb was larger than other CaSC species, except for *C. filicis* (62.96 Mb) and the very large *C. cuscutae* (80.45 Mb). De novo gene annotation on all *C. lupini* genomes resulted in a predicted gene count ranging between 16,077 and 17,181 (Figure 1c). The number of predicted secreted proteins within *C. lupini* ranged from 998 to 1107 and predicted effectors ranged from 418 to 480.
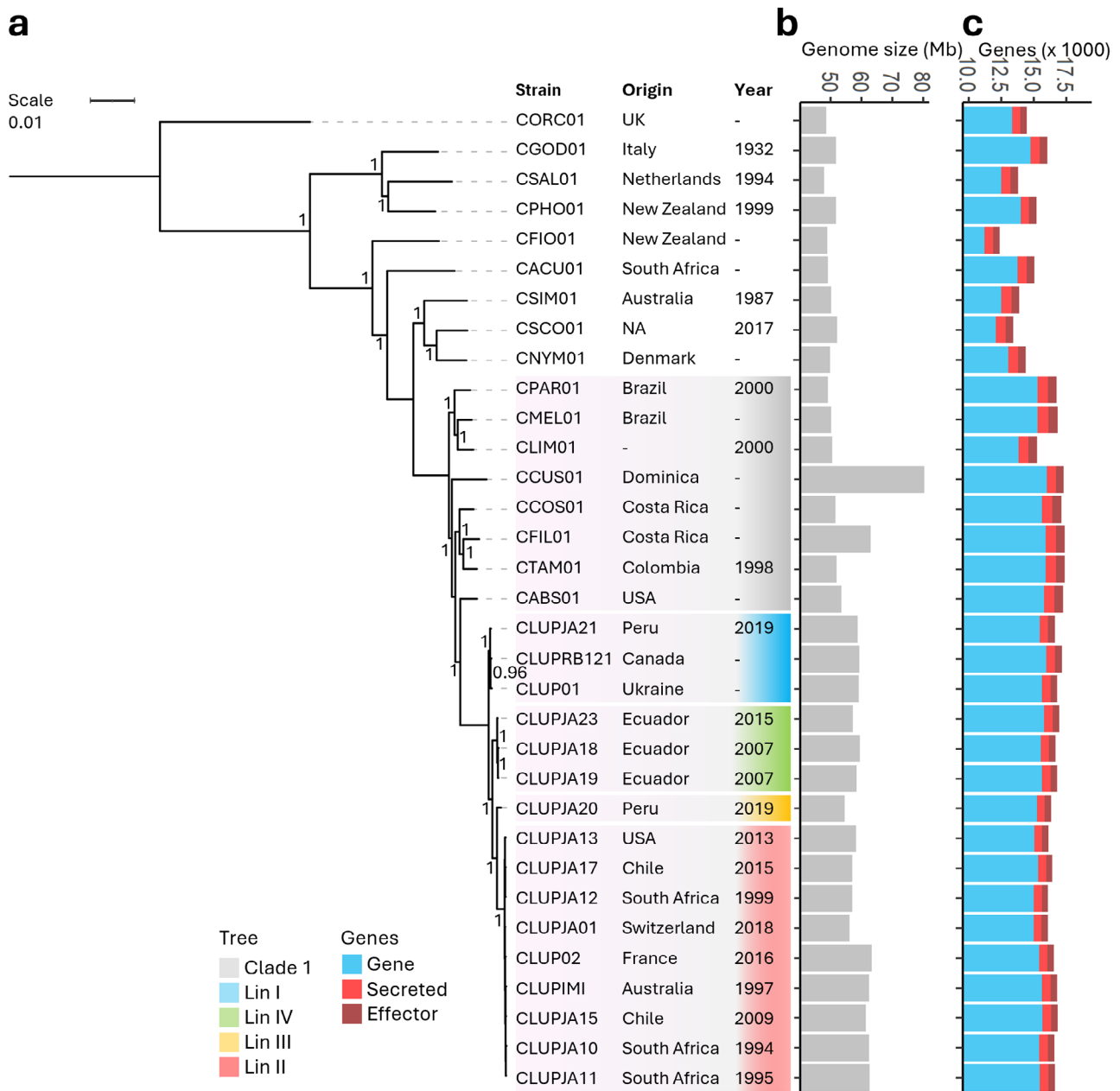
FIGURE 1 | Phylogenomic tree, genome size and gene content of the *Colletotrichum acutatum* species complex (CaSC). (a) Tree constructed based on 6350 single-copy orthologue genes (3,520,830 amino acids) using FastTree. Colours in the tree indicate *C. lupini* lineage or CaSC clade 1. Bootstrap support values are given at each node. Columns indicate strain name, origin, and year of collection. (b) Genome size (Mb) of CaSC species. (c) Predicted (secreted) gene and effector count of CaSC species. Colours indicate gene (blue), secreted (red), or effector (dark red).

Average predicted gene content for *C. lupini* (16,600) was similar to clade 1 (16,912) but higher than other CaSC species (14,294), whereas predicted effectors were lowest for *C. lupini* (497), followed by CaSC (562) and other clade 1 species (669; Figure S2). To assess genetic diversity, a phylogenomic tree was constructed from 6350 single-copy orthologue genes (3,520,830 amino acids [AA], PID = 94.9%). As expected, *C. lupini* was placed in clade 1 and consisted of four distinct but highly uniform lineages (Figure 1a). These four highly uniform lineages were also observed with a principal component analysis (PCA) based on 310,498 single-nucleotide polymorphisms (SNPs; Figure 2a). However, within lineage (Lin) II clear differences in genome sizes were observed (Figure 1b).

2.2 | Substantial Genome Variation Within *C. lupini*

To identify core (present in all strains) and accessory (present in some strains) chromosomes, Ragtag assemblies based on reference CLUP02 were used. Ten core and one accessory chromosome (AC), absent in Lin I, were identified in *C. lupini*. Core chromosomes showed an average size variation of 9.08%, with the highest variation of 22.54% observed for chromosome 8 (Figure 3a). AC 11 size ranged from 0.34 to 0.52 Mb, indicating an average variation of 14.89%, with the highest deviation compared to CLUP02 of 35.87% seen in CLUPJA19. Comparing genome sizes within *C. lupini* showed a variation
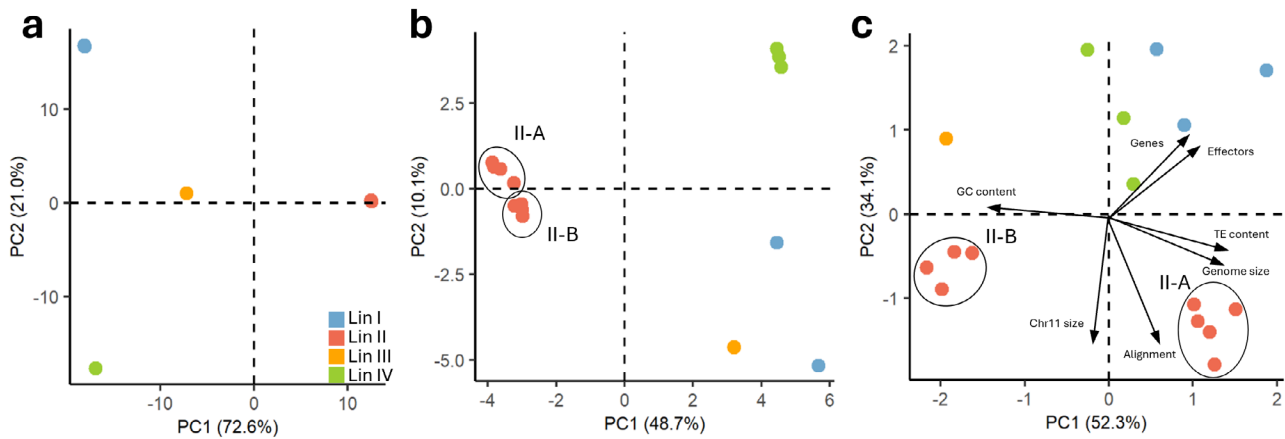
FIGURE 2 | Global genetic structure of 16 *Colletotrichum lupini* strains based on (a) 310,498 single-nucleotide polymorphisms (SNPs). (b) Presence/absence matrix of transposable element (TE) insertions. (c) Overall genome statistics: genome size, GC content, gene content, effector content, chromosome 11 size, TE content, and proportion of alignment to CLUP02. Principal components 1 and 2 are plotted. Blue indicates lineage I, red: lineage II, orange: lineage III and green: lineage IV. Lineage II-A and B are indicated with a circle.

of 5.58% within Lin II, compared to a variation of 0.5% for Lin I and 1.85% for Lin IV (Figure 1, Figure S3). Similarly, there was greater GC content (Lin II = 2.96%, Lin I = 0.24%, and Lin IV = 0.90%), gene content (Lin II = 2.25%, Lin I = 1.83%, and Lin IV = 0.92%), and effector content variation (Lin II = 4.75%, Lin I = 2.02%, and Lin IV = 0.73%) in Lin II compared to the other lineages. When plotting these variables, together with AC 11 size, TE content, and proportion of alignment to CLUP02 in a PCA, two distinct Lin II groups were identified (Figure 2c). These groups will be referred to as Lin II-A, including CLUP02, -IMI, -JA10, -JA11, and -JA15, and Lin II-B, including CLUPJA01, -JA12, -JA13, and -JA17. The total size of unmapped contigs ranged between 22.6 and 1165 kb and had a GC content ranging between 23.1% and 54.7%, with an average of 30.8%. The average size of unmapped contigs for Lin II-A was 67 kb, 447 kb for II-B, 400 kb for I, 607 kb for III, and 719 kb for IV (Figure S3).

Deletions were identified through whole-genome alignments (>80% identity) against the reference genome CLUP02 (Figure 3d). Two large deletions, ranging between 350 and 100 kb, were found in gene-poor regions of chromosome 4 (Figure 3e), and another 200 kb deletion was observed in a gene-poor region of chromosome 2 (Figure 3d). AC 11 showed the most variability, being completely absent in Lin I strains, with Lin IV strains only aligning to 70.5% and Lin III strain CLUPJA20 aligning to 82.2% of the reference (Figure 4b). Substantial diversity was noted within Lin II, where group A aligned to 90% and group B aligned to 86.5% of AC 11 of CLUP02. The observed deletions matched strain relatedness, with Lin I, III, and IV showing an average alignment of 97.7%. Within Lin II, group A aligned to 99.83% of the reference genome, while group B aligned to 98.91%. Synteny analysis between long-read assemblies CLUP01 (Lin I) and CLUP02 (Lin II) revealed high synteny apart from two major rearrangements between chromosomes 5 and 7 (621 kb) and chromosomes 2 and 6 (209 kb) in CLUP02 and CLUP01, respectively (Figure S4).

2.3 | Distinct Lineage-Specific TE Landscapes

TEs were predicted based on a consensus repeat library of known *Colletotrichum* repeats (Dallery et al. 2017), de novo predicted repeats from CLUP02 and CLUP01, and a combined DFAM and Repbase library. Total TE content ranged from 9.79% (CLUPJA20) to 20.59% (CLUP02, Figure 5b). The highest TE content was observed in AC 11 and ranged from 26.26% (CLUPJA20) to 42.30% (CLUPJA19; Figures 3b and 5c). High variability in TE content was observed within Lin II, with an average TE content of 19.68% for II-A and 13.38% for II-B (Figure S3). Total TE content (%) was strongly correlated with genome size ($r = 0.99$, $p = 5e-13$; Figure 5f) and core chromosomes ($r = 0.94$ to $r = 0.7$; Figure S5), whereas AC 11 TE content did not correlate with its size ($r = -0.024$, $p = 0.94$; Figure 5g). The long terminal repeat (LTR)-retrotransposons Copia and Gypsy were most common, representing on average 44.98% and 25.48% of the total TE content, respectively (Figure 5a). Unknown repeats were the third most common, representing 16.57% of the total TE content, and the most abundant DNA transposon, IS3EU, represented 3.15%. LTR-Copia ($r = 0.98$, $p = 2e-12$) and unknown repeat content ($r = 0.97$, $p = 2.7e-10$) were highly correlated with total genome size, whereas LTR-Gypsy content was not ($r = 0.43$, $p = 0.1$; Figure S5), indicating that LTR-Copia has the highest impact on observed genome size variations. The long interspersed nuclear elements (LINE)-Tad1 transposons were on average the largest (3495 bp), followed by the LTR retrotransposons Copia (2945 bp) and Gypsy (2831 bp; Figure 5d). On AC 11, DNA transposon IS3EU was only present in Lin III and IV, and rolling circle (RC) helitrons were almost absent in CLUPJA20 (Lin III) representing only 0.22% compared to an average of 1.65%. To further analyse TE diversity within *C. lupini*, all genomes were screened for TE insertions, which were classified as present (1) or absent (0) in the reference genome. TE insertions resembled a normal distribution, with most insertions being present in half of all the strains. Only 5% of the insertions was present in only one of the isolates (Figure S6). A PCA was created based on the
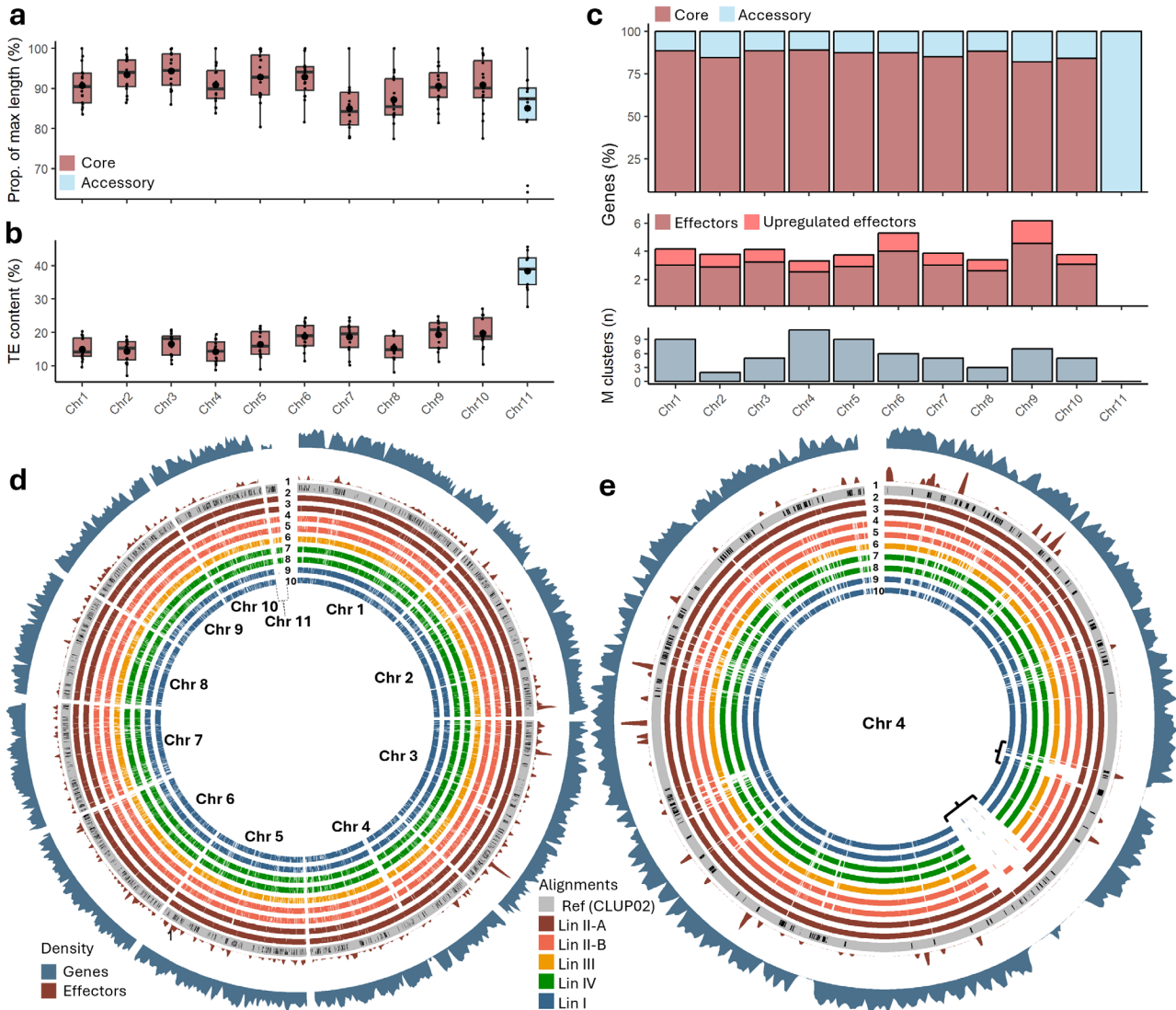
FIGURE 3 | Genome variability of *Colletotrichum lupini*. (a) Chromosome length variation expressed as the percentage of the maximum observed length. Lineage I strains are excluded for chromosome 11. (b) Transposable element (TE) content (%) for each chromosome. (c) Top, percentage of core (red), and accessory (blue) genes per chromosome. Middle, percentage of upregulated (red) and stably expressed (dark red) effectors per chromosome. Bottom, number of metabolic gene clusters per chromosome. (d) Whole-genome alignments with gene (blue) and effector (red) abundance (min = 0.2, max = 3.5/kb). (e) Chromosome 4 alignment with gene and effector abundance (min = 0, max = 2.5/kb), gaps of 350 and 100 kb are indicated. Within alignment plots: 1: reference genome CLUP02 with annotated transposons, 2: CLUPJA10, 3: CLUPJA15, 4: CLUPJA01, 5: CLUPJA17, 6: CLUPJA20, 7: CLUPJA18, 8: CLUPJA23, 9: CLUP01, and 10: CLUPJA21.

presence/absence matrix of TE insertions, showing the four expected lineages but with a higher interlineage variability than was shown with SNPs (Figure 2b).

2.4 | Pangenome Analysis Reveals Dynamic Accessory Genome

Despite previous observations of limited within-lineage genetic variation from SNP data (Alkemade et al. 2023), variation in accessory genes/chromosomes can generate genetic variation within a species. To analyse gene-content diversity, all 265,271 predicted protein-coding genes across the 16 *C. lupini* strains were grouped into 17,535 orthologous groups (OGs) based on protein homology. The pangenome included 14,107 (80.45%) core groups, with 13,004 (74.2%) consisting of single-copy

orthologous genes (Figure S7). Only one core group had genes with a copy number above 10, and only three had more than 5 (Figure S8). The remaining 3426 (19.54%) were considered accessory, shared among some but not all strains. Only two groups, containing six genes in total, were found in a single genome (Table S2).

When compared with the other CaSC genomes, 6350 single-copy orthologous gene groups were shared across all species, representing core CaSC genes, while 1340 OGs were unique to *C. lupini*. Amongst the 593 OGs containing 7951 genes identified as effectors, 358 (60.37%) were part of the core genome, and the rest were accessory (Figure S7). A total of 11,087 carbohydrate-active enzymes (CAZymes) were identified, which clustered into 707 OGs. Of these, 4598 (41%) were secreted, and a total of 586 CAZyme orthogroups (82.9%) were conserved among
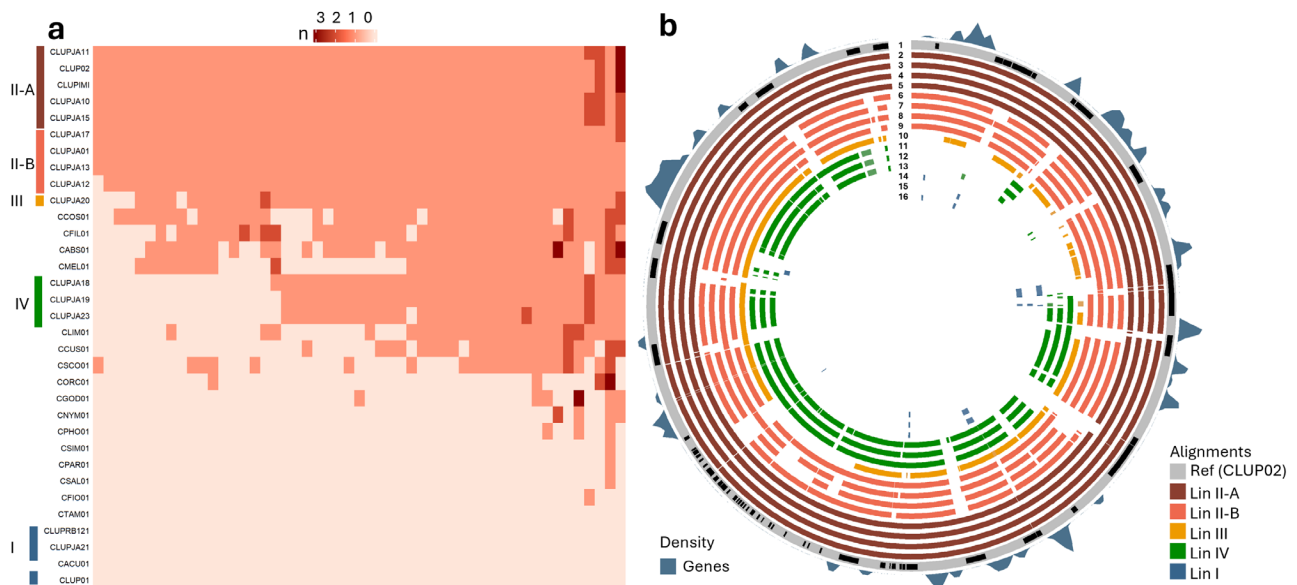
FIGURE 4 | Overview of accessory chromosome 11: (a) orthogroups (columns) shared amongst *Colletotrichum acutatum* species complex (CaSC) species (rows). (b) Whole-genome alignment (>80%) with gene (blue) abundance (min = 0, max = 0.3/kb). From outer ring to inner: 1: reference CLUP02 with annotated transposons shown in black, 2: CLUPIMI, 3: CLUPJA10, 4: CLUPJA11, 5: CLUPJA15, 6: CLUPJA01, 7: CLUPJA12, 8: CLUPJA13, 9: CLUPJA17, 10: CLUPJA20, 11: CLUPJA18, 12: CLUPJA19, 13: CLUPJA23, 14: CLUP01, 15: CLUPJA21, and 16: CLUPRB121. Total length of outer ring is 0.52 Mb.

all *C. lupini* genomes (Figure S7). Most of the CAZymes could be grouped in five different families, with the biggest family being glycoside hydrolases (GH), representing 52.5% and approximately 363 genes per genome (Figure S9). Core gene copy number variation showed that South American strains clustered together, whereas accessory, CAZyme, and effector OGs conformed to previously described lineages (Figure S8).

Higher protein conservation, expressed as OG identity, was observed for effectors (97.8%), CAZymes (97.0%), and secreted genes (97.3%) compared to *C. lupini* core (96.9%) and randomly selected genes (95.8%; Figure 6c). Lowest conservation was observed for *C. lupini*-specific (94.7%), accessory (93.4%), and AC 11 (93.0%) genes. Additionally, *C. lupini*-specific (dN/dS = 0.67 [median], 250 AA), AC 11 (dN/dS = 0.66, 332 AA), and accessory (dN/dS = 0.51, 489 AA) genes appear to evolve under relaxed selective pressure and are shorter compared to core CaSC (dN/dS = 0.14, 601 AA), core *C. lupini* (dN/dS = 0.26, 521 AA), and randomly selected genes (dN/dS = 0.32, 513 AA; Figure 6a,b).

When plotted on the reference genome, effectors were not dispersed equally across the genome but were segmented in effector-rich but gene-poor regions (Figure 3d,e), showing a distribution coefficient of variation (CV) of 1.25 compared to a CV of 0.29 for all genes (Figure S10). Despite the high protein conservation observed in effectors, CAZymes, and secreted genes, they tend to cluster near transposable elements (TEs). Their average distances to the nearest TE and the proportions of genes within 10 kb of a TE are 14 kb (65%), 19 kb (57%), and 15 kb (64%), respectively (Figure 6c, Figure S11). This is closer than core CaSC genes (30 kb, 32%) and randomly selected genes (24 kb, 45%) but similar to accessory genes (15 kb, 67%). Genes on AC 11 are much closer, with an average distance of 2 kb and all genes within 10 kb of a TE. Whole-genome selective sweep analysis

across the 16 *C. lupini* isolates indicated 128 regions that have been subjected to positive selection (composite-likelihood ratio [CLR] > 10; Figure S12). These regions, with the majority (95%) overlapping with TEs, contained 2.81% of accessory, 2.56% of effector, and only 1.86% of CaSC core genes. Across all lineages no selective sweep signatures were found on AC 11. Within Lin I, II, and IV, no selective sweeps were found at all, suggesting the absence of intralineage recombination. These results indicate that regions containing effector, accessory or *C. lupini*-specific genes tend to serve as hotspots for diversification.

Expression per gene category was based on data of CLUP02 growing for 24 h in liquid culture and 24 and 84 h after infection of white lupin (*L. albus*; Dubrulle et al. 2020). Analysis of differentially expressed genes (DEGs) showed that CAZyme (40.22%), secreted (33.72), effector (31.42%), and accessory (29.3%) genes had highest fractions of upregulated genes, in contrast to *C. lupini* core (21.8%), AC 11 (15.79%), and *C. lupini*-specific (13.76%) genes (Figure 6e). Average expression ($\log_{10}$[normalised counts +1]) was highest for CaSC core genes (1.33), followed by *C. lupini* core genes (0.93; Figure 6f). Expression was lowest for accessory (0.47), *C. lupini*-specific (0.37), and AC 11 genes (0.23), and decreased in the early infection stage (Figure 6g). CAZymes showed an overall expression of 0.87 and expression increased during plant infection. Overall expression of effectors (0.57) and secreted (0.63) proteins was lower but increased 84 h after infection, indicating host colonisation-related activity.

2.5 | Species- and Lineage-Specific Effector Repertoires in *C. lupini*

Distinct virulence patterns for each strain were observed on white and Andean lupin (*L. mutabilis*; Alkemade et al. 2023; Alkemade et al. 2021). Virulence patterns on Andean lupin, however, did not
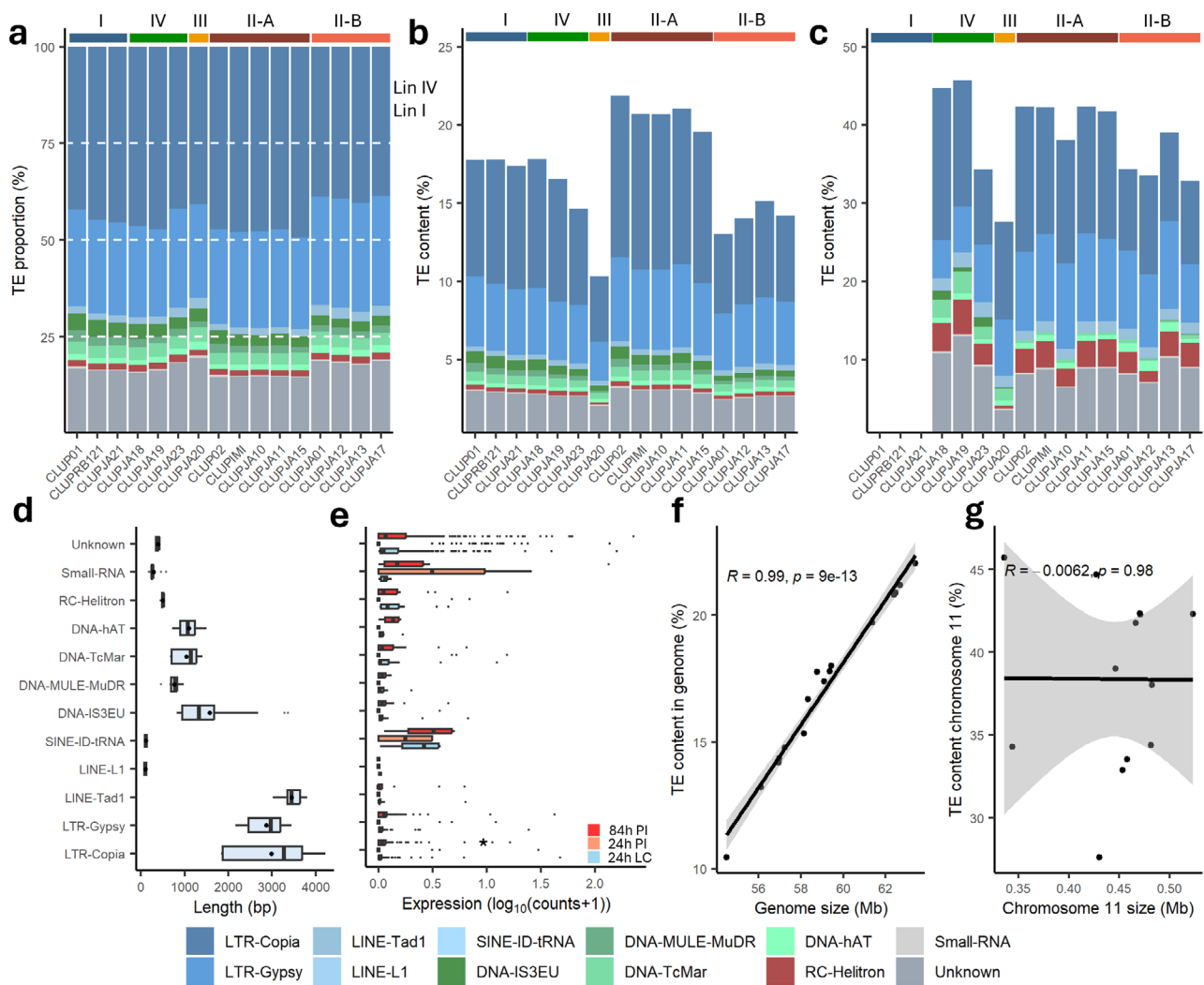
FIGURE 5 | Transposable element (TE) landscape of *Colletotrichum lupini*. (a) Relative frequency of TE superfamilies across all genomes with 100% referring to the total TE content of the respective genome. (b) Contribution of TE superfamilies to core genome size. (c) Contribution of TE superfamilies to accessory chromosome 11 size. (d) TE superfamily length in bp, black dots indicate means. (e) TE de-repression shown as expression at 24 h in liquid culture (LC), 24 and 84 h after white lupin inoculation (PI), asterisk (*) indicates significant difference between 24 h in LC and infection treatment (Kruskal–Wallis, $p < 0.05$). (f) Correlation (Pearson) of TE content (%) to genome size. (g) Correlation of TE content (%) to chromosome 11 size. Bottom: transposon superfamily legend.

correspond to assigned lineages, with great variation within and no difference between lineages for virulence on Andean lupin (Figure S3h–j). On white lupin, Lin I, III, and IV showed weak virulence, whereas virulence levels of Lin II-A and B were high. Effectors play an important role in virulence and were upregulated upon infection (Figure 6g). To explain observed virulence patterns and to identify *C. lupini*- and lineage-specific effector-containing orthologous genes, a genome-wide association study (GWAS)-like approach was used. Predicted orthologous genes of mature effector proteins were used as input, and species, lineage, and virulence data were used as "phenotype" input (Alkemade et al. 2021). This resulted in four effector orthogroups appearing to be unique for *C. lupini* (Table S3, Figure S13). These four orthogroups (OG0000520, −539, −540, and −542) contained XP_049136846.1, XP_049146329.1, XP_049145599.1, and XP_049136058.1 of the reference genome CLUP02. The predicted effector XP_049136846.1, encoding a peptidase A1 domain and upregulated upon plant infection (Table S3), only appears to match with effectors of *Seridium cardinale*, the

cereal pathogen *Pyrenophora teres* and *Microdochium nivale* (Figure S14). XP_049146329.1 was only shared with closely related *C. tamarilloi*, *C. costaricense*, and *C. filicis*, whereas XP_049145599.1 and XP_049136058.1 yielded no matches to any other species. In Lin I, seven unique effector orthogroups were identified, nine in II, one in II-A, one in II-B, three in III, and 10 in IV (Figure 7, Table S3). From the nine Lin II unique orthogroups, four (OG0000453, −016, −632 and −506) contained CLUP02 genes that were upregulated in planta, encoding a LysM domain, a fungal lipase, a protein kinase and a EC51a protein, respectively. One Lin II-specific orthogroup (OG0000536) contains secreted-in-xylem (SIX) proteins, which are commonly associated with pathogenicity. Two Lin I-specific orthogroups (OG0000537 and −745) contain KAK1721939.1 and KAK1709830.1, which encode a necrosis-inducing Ecp2 effector and a biotrophy-associated secreted protein 3 with a Zn(2)-C6 fungal-type domain, respectively. Lin III-specific OG0000526 contains a protein homologous to a CEC3 effector of *C. trifolium*. Based on virulence on white lupin (cultivar Feodora), only one orthogroup (OG0000033)
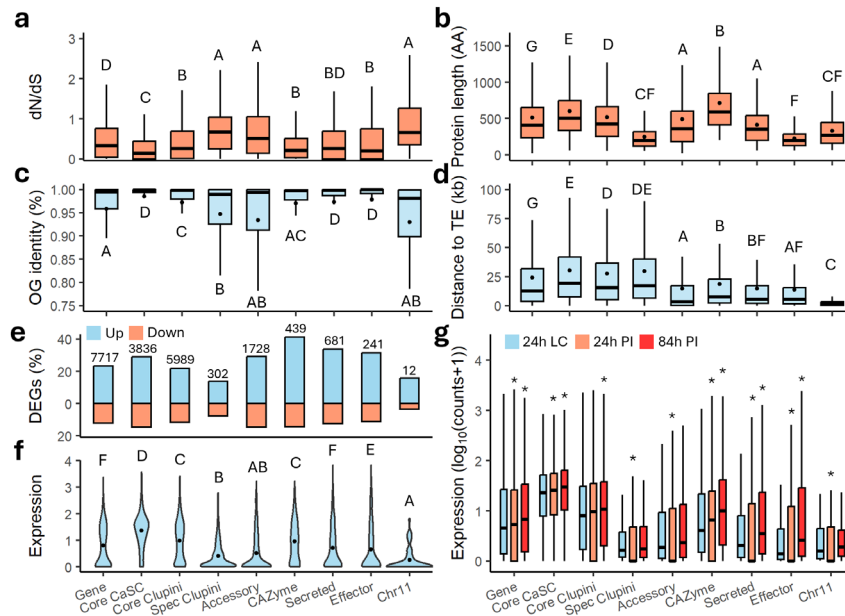
FIGURE 6 | Evolutionary dynamics of different gene categories. (a) Sequence conservation estimated by the number of nonsynonymous and synonymous substitutions (dN/dS) values. (b) Percent identity given by the multiple protein sequence alignments for each orthogroup. (c) Protein length in amino acids (AA). (d) Distance (kb) to closest transposable element (TE), mapped on reference genome CLUP02. (e) Differentially expressed genes (DEGs), numbers indicate total amount of DEGs. (f) Overall expression of CLUP02 genes per category. (g) Expression of CLUP02 genes at 24 h in liquid culture (LC), 24 and 84 h after inoculation (PI), asterisks (*) indicate significant difference between 24 h in LC and infection treatment (Kruskal–Wallis, $p < 0.05$). Uppercase letter within plots indicate significant differences between gene categories (Dunns's test, $p < 0.05$). Big black dots in (b, c, d and f) indicate means.
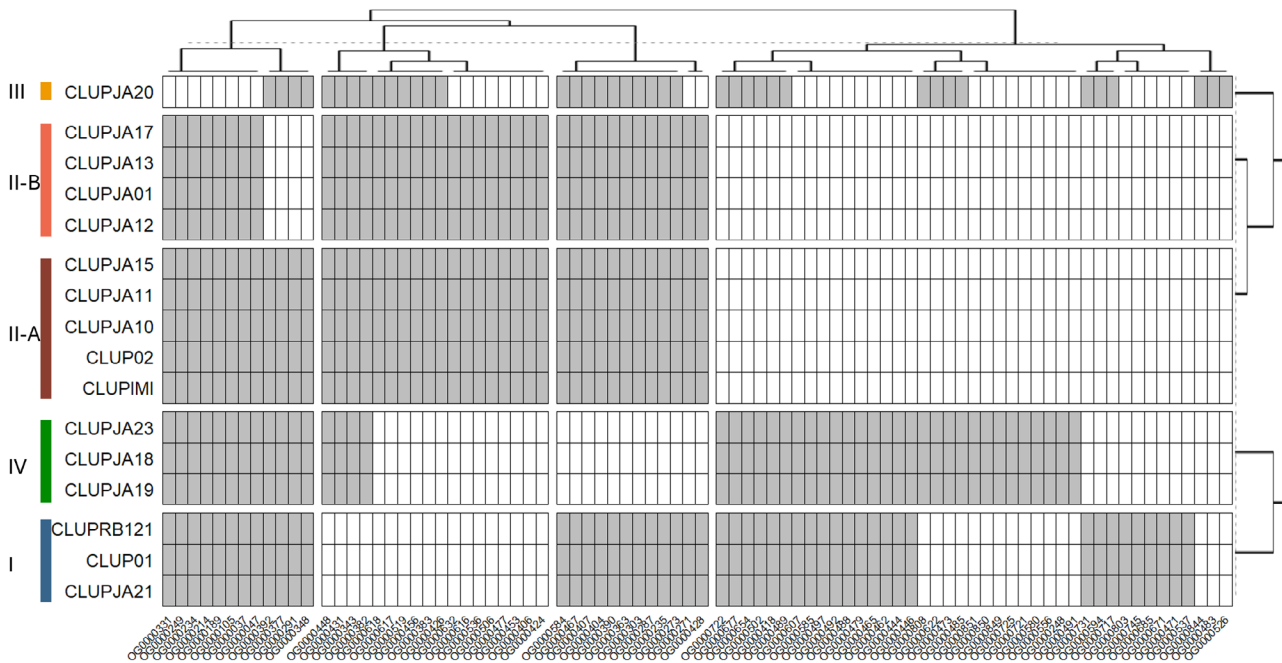

FIGURE 7 | Lineage-specific predicted effector-containing orthogroups. See Table S3 for predicted gene functions. Grey is present, white is absent. Euclidian distance dendrograms are shown above and right of the plot.

could be identified, encoding a protein with a cellulose-binding domain. No orthogroups were identified based on virulence on Andean lupin. Using the same approach, two Lin II- and one Lin IV-specific CAZyme orthogroups were identified, corresponding to a glycoside hydrolase, calcium-translocating P-type ATPase, and chitinase, respectively (Table S3).

2.6 | De-Repression of LTR-Copia Upon Infection

As TEs have been linked to virulence in various fungal plant pathogens, we compared TE content to *C. lupini* virulence. While no correlation was observed between total TE content and virulence on white lupin ($r = 0.28$, $p = 0.3$), significant correlations

were identified for the specific TE superfamilies LTR-Gypsy ($r = 0.5$, $p = 0.05$), LINE-Tad1 (0.83, $p = 7.5e-5$) and DNA-hAT ($r = 0.69$, $p = 0.003$; Figure S5). No correlation between TE content and virulence on Andean lupin was observed. Expression data showed that upon infection of white lupin, apart from a slight de-repression of LTR-Copia ($p = 0.009$), expression remained stable and there was no relaxation of TE repression (Figure 5e).

2.7 | Accessory Chromosome Not Directly Involved in Virulence

ACs are often involved in virulence (Bertazzoni et al. 2018), so we had a closer look at chromosome 11. AC 11 is conserved amongst Lin II isolates, partly present in Lin III and IV, and absent in Lin I (Figure 4b). A total of 54 genes was annotated on AC 11 of CLUP02, with none predicted to encode for a secreted, effector or CAZyme protein. No metabolic gene clusters were identified either (Figure 3c). The gene-poor AC 11 was clearly divided in gene- and TE-rich regions (Figure 4b). Orthologues were shared amongst lineages II, III, and IV and most clade 1 species (CCOS01, CFIL01, CABS01, CLIM01, CCUS01, and CSCO01), indicating that those species probably contain an AC as well (Figure 4a). Most AC 11 genes encode uncharacterized proteins, but one gene encodes a zinc finger C2H2-type (XP_049138510.1; Table S4). Two proteins, XP_049138504.1 and XP_049138485.1, were unique for Lin II, encoding an uncharacterized protein and a ubiquitin-like protease family profile domain-containing protein. The absence of secreted proteins, effectors, CAZymes, or metabolic gene clusters, the low expression upon plant infection (Figure 6f,g), the fact that Lin I did not show reduced virulence on Andean lupin, and that Lin I virulence on white lupin was not different compared to AC 11-containing Lin III and IV (Figure S3), suggests that AC 11 is not directly involved in lupin colonisation.

3 | Discussion

A global collection of *C. lupini* strains was sequenced to gain insights into the evolution and genetic variability of this asexual fungal plant pathogen. Based on conserved core genes, phylogenomics clearly showed that *C. lupini* was divided in four highly uniform lineages, matching the population analysis based on SNP profiling by Alkemade et al. (2023). The lack of genetic diversity, however, appeared at odds with the numerous differences in morphology and virulence observed within the species (Alkemade et al. 2021). Differences in TE content, gene content, whole-genome alignment, and AC presence revealed that *C. lupini* is more diverse than previously detected, with Lin II clearly splitting into two distinct groups, II-A and II-B.

Pangenome analysis of *C. lupini* revealed significant genomic diversity, characterised by a stable core genome (80%) and dynamic accessory genome. This level of conservation is similar to three human fungal pathogens and *Saccharomyces cerevisiae*, where core gene estimates ranged between 80% and 90% (McCarthy and Fitzpatrick 2019), but is more conserved than the plant pathogens *Colletotrichum graminicola* (60%; Becerra et al. 2023), *F. oxysporum* f. sp. *cubense* (53%; van Westerhoven et al. 2024), and *Zymoseptoria tritici* (60%; Badet et al. 2020). While the core genes show limited copy number variation and high protein

conservation, the accessory genome, including AC 11, exhibits great variability, with relaxed selective pressure and closer proximity to TEs. Genes found in all *C. lupini* strains but not in the rest of the CaSC also showed low conservation and relaxed selection pressure. This variability suggests that the accessory genome and *C. lupini*-specific genes play a key role in generating genetic diversity, as is seen for other *Colletotrichum* species (Lapalu et al. 2024; Liang et al. 2024) and asexual fungal plant pathogens (van Westerhoven et al. 2024; Langner et al. 2021). Pangenome analysis, however, greatly depends on the diversity of strains included in the analysis and should be interpreted carefully.

TEs contribute to intraspecific variation and play an important role in the evolution of fungi (Wells and Feschotte 2020; Grandaubert, Balesdent, and Rouxel 2014; Faino et al. 2016; Seidl and Thomma 2017). Especially in partially clonal fungal plant pathogens, such as *Magnaporthe*, *Fusarium*, and *Verticillium*, TEs have been shown to be crucial for host adaptation (Badet et al. 2020; Faino et al. 2016; Nakamoto, Joubert, and Krasileva 2023). In *Colletotrichum*, TE content was shown to be responsible for genome size variation and was linked to diversification (Tsushima et al. 2019; Gan et al. 2021). The most common TEs in *C. lupini* are the LTR retrotransposons Copia and Gypsy, which is similar to that found for other *Colletotrichum* species (Rao et al. 2018). In *C. lupini*, Copia elements especially, and Gypsy elements to a lesser extent, seem to have contributed most to genome expansion within the species. In the pathogens with clonal lineages *Rhynchosporium commune*, *Blumeria graminis*, and *Z. tritici*, Gypsy elements make up about half of the TE copies in the genome and have contributed to recent genome expansions and variability (Stalder et al. 2023; Frantzeskakis et al. 2018; Oggenfuss et al. 2021). This highlights that TE-mediated variation can provide an important source for adaptation when meiosis is rare or lacking. Retrotransposons have been shown to be de-repressed upon stress (Fouché et al. 2020; Torres, Thomma, and Seidl 2021; Gupta et al. 2023). In *C. lupini* Copia elements were the only TEs being de-repressed upon infection. This has also been observed in *Z. tritici*, in which the other TE elements were only upregulated during the less stressful saprophytic stages of the pathogen's life cycle (Fouché et al. 2020). In *Botrytis cinerea*, Copia elements have even been directly linked to increasing virulence (Porquier et al. 2021). Besides de-repression, correlations between white lupin virulence and LTR-Gypsy, LINE-Tad1, and DNA-hAT TE content were observed. These TE types have been indicated to improve the adaptability of fungal pathogens (Muszewska et al. 2019). Besides the benefits, TE de-repression increases the risk of TE insertions into essential genes and the rate of deleterious rearrangements (Fouché et al. 2022). To increase accuracy of transposon prediction, chromosome-level assemblies could be used to provide a more complete understanding of TE biology in *C. lupini*.

The *Colletotrichum* genus is thought to consist of 10 core chromosomes and a variable number (0–8) of ACs (Wang et al. 2023). In *C. lupini*, a single AC has been identified that was TE dense and highly variable in size, which is common for ACs seen in other fungal species (Habig and Stukenbrock 2020; Bertazzoni et al. 2018). As this analysis has been based on alignments to a reference genome, long-read assemblies will be required to confirm the presence of one or more ACs in *C. lupini*. In fungal plant pathogens, ACs are often associated with pathogenicity (Bertazzoni et al. 2018).

In *F. oxysporum* f. sp. *lycopersici*, transferring AC 14 converted nonpathogenic strains into pathogens (Ma et al. 2010). In *Colletotrichum* species *C. asianum* (Wang et al. 2023), *C. higginsianum* (Dallery et al. 2017; Plaumann et al. 2018) and *C. lentis* (Bhadauria et al. 2019), ACs are effector-rich and involved in mediating virulence. In *C. lupini*, however, AC 11 did not contain any effector, secreted or CAZyme-encoding genes or any metabolic gene clusters. Although Lin I strains showed low virulence on white lupin, the absence of AC 11 did not hamper virulence on Andean lupin (Alkemade et al. 2021; Alkemade et al. 2023), indicating no direct function related to virulence. The lack of effectors on ACs was also shown for *C. graminicola* (Becerra et al. 2023), but deletion of its AC 12 still hindered full virulence (Ma et al. 2023). In *B. cinerea* and *Z. tritici*, no direct link to pathogenicity was found and ACs are suggested to be required for niche adaptation rather than pathogenicity (Van Kan et al. 2017; Habig, Quade, and Stukenbrock 2017). While virulence factors are a prominent feature of ACs in fungal plant pathogens, these chromosomes can serve diverse functions beyond pathogenicity, reflecting the dynamic nature of fungal genomes and their adaptation strategies in various ecological environments (Bertazzoni et al. 2018; Habig and Stukenbrock 2020). Performing competition experiments and infection trials on a broader spectrum of lupin species with strains with and without AC 11 might further elucidate its function.

Effectors are not randomly distributed across the *C. lupini* genome but are mostly localised in gene-poor and TE-rich regions. This has been observed for other fungal plant pathogens such as *C. higginsianum* (Tsushima et al. 2019) and *F. oxysporum* (van Westerhoven et al. 2024) and is considered common but not ubiquitous across fungi (Dong, Raffaele, and Kamoun 2015; Torres et al. 2020). This co-localisation facilitates variability of effectors, which can contribute to niche adaptation. Effectors are often less conserved than core genes (Badet et al. 2020), but in *C. lupini*, pairwise identity amongst effectors was similar to core genes. This suggests that the presence/absence of effectors might be more common than amino acid substitutions to alter virulence and host specificity. In *F. oxysporum*, for example, the presence or absence of specific effectors greatly influences host range (van Westerhoven et al. 2024; Batson et al. 2021). Even though observed virulence on the tested Andean lupin varieties did not correspond to described lineages, presence/absence variation of predicted effectors did, indicating adaptive lineage-specific effector repertoires. One of the four effectors unique to *C. lupini* was upregulated during host colonisation and encodes a peptidase A1 domain, which has been linked to virulence in various fungal pathogens (Qian et al. 2022; Krishnan et al. 2018). In Lin II, two lineage-specific effectors were highly upregulated, one encoding a fungal lipase, shown to contribute to virulence in *F. graminearum* on wheat (Voigt, Schäfer, and Salomon 2005), and an EC51a protein homologous to a candidate effector in *C. higginsianum* (Gan et al. 2021). The gene function of accessory effectors identified in this study and its relevance for host specificity should be further characterised by knockouts or gene silencing.

In conclusion, we discovered considerable underlying genetic variability in *C. lupini* lineages that was not apparent from earlier surveys on sequence variation of SNPs. Our analyses do not rule out the possibility of rare sexual reproduction in *C. lupini*. Yet, the key components of diversity in chromosome structure, TE activity, and the presence or absence of putative effector genes that we detected can all arise through mechanisms independent of sexual reproduction. These findings now open up the potential to identify specific pathogenicity factors behind colonisation of lupins, disease symptoms, and crop losses. Understanding mechanisms that generate genetic variability in fungal plant pathogens, especially in genes that cause disease, is crucial for designing durable control and breeding strategies and is vital to sustainably reduce their impact on global food production.

4 | Experimental Procedures

4.1 | Culture Collection and DNA Extraction

A total of 16 *C. lupini* strains were collected from public culture collections and lupin plants with symptoms by collaborators worldwide, representing 10 countries across five continents (Table S1, Figure S1). All isolates were single-spored and maintained on potato dextrose agar (PDA; Carl Roth) at 22°C in the dark as working cultures. Isolates were stored in 25% glycerol at −80°C for long-term storage. To extract DNA, mycelium from single-spore cultures was collected after 10 days of growing on PDA at 22°C with a sterile spreader after flooding the Petri dish with 2 mL of sterile double-distilled water. Genomic DNA was isolated with a CTAB extraction protocol described in Minas et al. (2011).

4.2 | Sequencing and Genome Assembly

Whole-genome shotgun sequences were obtained through 150 bp paired-end sequencing at a depth of >115× coverage on an Illumina NovaSeqX Plus platform by Genome Quebec (Quebec, Canada). Raw reads were trimmed for adapter sequences and filtered for a phred quality of 20 using FastP v. 0.23.4 (Chen et al. 2018). Quality was evaluated using FastQC v. 0.12.1 (Andrews 2010) and summarised using MultiQC v. 1.14 (Ewels et al. 2016). Reads were assembled using SPAdes v. 3.15.5 (--isolate; Bankevich et al. 2012), and resulting scaffolds were filtered for minimum length (350 bp) and min (30×) and max (1000×) coverage. Assemblies were screened and cleaned from potential contamination using BLASTn (nt database) and BlobTools v. 1.1.1 (Laetsch and Blaxter 2017). QUAST v. 5.2.0 (Gurevich et al. 2013) was used to assess assembly quality and length, and BUSCO v. 5.4.7 (Manni et al. 2021) with the glomerellales_odb10 dataset was used to assess completeness. Kmers were counted using Jellyfish v. 2.3.0 (Marçais and Kingsford 2011). The resulting assemblies were compared to publicly available *C. lupini* and CaSC assemblies (Table S1). To identify core and accessory chromosomes, homology-based scaffolding, with CLUP02 as a reference, was performed using RagTag v. 2.1.0 (Alonge et al. 2022).

4.3 | Transposable Elements and Repeat Annotation

Repetitive elements were identified in the complete genomes of CLUP02 and CLUP01 using Repeatmodeler v. 2.0.4 (Flynn et al. 2020) with options "-engine ncbi" and "-LTRStruct". A consensus library of the predicted repetitive sequences and previously

described *Colletotrichum* transposons (Dallery et al. 2017) was created, filtering for identical sequences (identity and coverage >80%). Repeats were classified using the Repeatclassifier of RepeatMasker v. 4.1.5 (Flynn et al. 2020) and the created consensus library was split into known and unknown repetitive elements. Genomes were annotated for TEs and masked in four steps using RepeatMasker with a cut-off value of 250. First, simple repeats were identified and soft-masked based on fungal repeats present in a combined Rebase (Bao, Kojima, and Kohany 2015; release 20181026) and Dfam v. 3.7 (Storer et al. 2021) database, followed by the identification and hard masking of complex repeats based on the *Glomerella* repeats present in the database and known and unknown repeats of the consensus library, ignoring short (<100 bp) repeats. Masked short-read assemblies were used for gene annotations, whereas RagTag assemblies were used to determine the TE landscape.

4.4 | SNP and TE Insertion Analysis

SNPs were called by mapping reads to reference genome CLUP02 using BWA v. 0.7.17-r1188 (Li 2013). Resulting SAM files were converted to BAM files that were indexed and sorted using SAMtools v. 1.17 (Danecek et al. 2021). Variant calling was performed using mpileup of BCFtools v. 1.14 and variants were filtered for quality (Q20), minimum sequencing depth (2), mean sequencing depth (5), minor allele count (2), minor allele frequency (0.01), and missing data (0.95) using VCFtools v. 0.1.16 (Auton & Marcketta, 2009). TE insertions were called using ngs-te-mapper2 (Linheiro and Bergman 2012). Sequencing reads were mapped against the consensus TE library, using a window of 100 bp, to identify reference and non-reference TEs. A PCA was created from whole-genome SNPs and reference TE insertions using the *prcomp* function in R v. 4.3.1 (R Core Team 2020). Selective sweep analysis was performed using SweeD v. 4.0.0 (Pavlidis et al. 2013) with a 10 kb sliding window using the above-mentioned SNP dataset. The analysis was performed for all isolates together and for each lineage (I, II, and IV) separately. Results are expressed as composite-likelihood ratio (CLR).

4.5 | Gene Prediction and Functional Annotation

Gene annotation was performed on masked genomes with four rounds of the MAKER v. 3.01.04 annotation pipeline. In the first round, transcript sequences of CLUP02 were used as EST evidence, and CSCO01 and CFIO01 transcripts were used as alternative EST evidence, protein sequences of CLUP02, CLUP01, CPAR01, CSCO01, and CFIO01 were used for protein homology evidence (Table S1). The following three rounds were performed with a CLUP02-trained version of AUGUSTUS v. 3.5.0 (Keller et al. 2011) and SNAP v. 2006-07-28 (Korf 2004), which was trained each consecutive annotation round. Predicted proteomes were assessed for completeness using BUSCO (glomerellales_odb10).

The secretome was defined by proteins with a signal peptide but no transmembrane domain as predicted by Phobius v. 1.01 (Käll, Krogh, and Sonnhammer 2004), TMHMM v. 2.0 (Krogh et al. 2001), WoLF PSORT (Horton et al. 2007), and SignalP v. 6.0 (Teufel et al. 2022) integrated within the EffHunter v. 1.0 (Carreón-Anguiano et al. 2020) pipeline with a protein length

range of 0–13,000 amino acids and a minimum of 0 cysteine residues. Predicted secretome was screened for effector candidates by EffectorP v. 3.0 (Sperschneider and Dodds 2022). Protein functions were predicted using Interproscan v. 5.63–95.0 (Jones et al. 2014), adding -goterms -iprlookup and -pathway information. Carbohydrate-active enzymes (CAZymes) were identified using the dbCAN3 server with the HMMER:dbcan, HMMER:dbcansub, and DIAMOND:CAZy tools (Buchfink, Xie, and Huson 2015; Finn, Clements, and Eddy 2011; Zheng et al. 2023). Proteins were classified as a CAZyme if predicted by each of the three tools. Secondary metabolite gene clusters were predicted using the online fungal version of antiSMASH v. 7.0 (Blin et al. 2023).

4.6 | Comparative Genomics and Pangenome Analysis

Whole-genome alignments were performed using nucmer (options --maxmatch -c 80) in MUMmer v. 3.1 (Kurtz et al. 2004). Alignment was performed against the repeat-masked reference genome of CLUP02. Resulting files were filtered for 1 to 1 alignments with a sequence identity of >80% and an alignment length of >500 bp. Alignment plots were created using Circos (Krzywinski et al. 2009) and R package *circlize* (Gu et al. 2014). Phylogeny and orthologous groups (OGs) were identified using OrthoFinder v. 2.5.4 (Emms and Kelly 2019) with the "-M msa" option on predicted protein-coding genes of *C. lupini* and CaSC species. OGs present in all isolates were considered core CaSC, and genes only present in *C. lupini* were considered *C. lupini*-specific. Within *C. lupini*, OGs present in all isolates were considered core and genes present in fewer isolates were considered accessory. Genes were further divided into the following categories: CAZyme, secreted, effector and Chr11 genes. Percent identity of multiple protein sequence alignments was determined for each OG using MAFFT v. 7.505 (Katoh and Standley 2013). The number of nonsynonymous and synonymous substitutions (dN/dS) per OG was determined by using the aligned proteins, followed by a codon-guided nucleotide alignment using PAL2NAL v. 14.1 (Suyama, Torrents, and Bork 2006). This codon-guided alignment was used together with OG gene trees generated by Orthofinder to infer dN/dS values using CODEML from PAML v. 4.9 (Yang 2007). For each category, intersection, or distance to a selective sweep (>CLR 10) or TE was determined with the "closest" and "intersect" option of BEDTools v. 2.30 (Quinlan and Hall 2010). Gene distribution along the genome for each gene category was assessed by coefficients of variation (CV), and it was tested if genes followed a Poisson distribution. A fictional dataset of randomised genes was used as control. To identify *C. lupini*- and lineage-specific effectors, a GWAS-like approach, using *statgenGWAS* (van Rossum et al. 2020) and Orthofinder generated gene-count matrices, was performed. Identified genes were manually checked for specificity using BLAST.

4.7 | RNA-Seq Analysis

Expression data of *C. lupini* growing for 24 h in liquid medium (Czapek-Dox 0.5 g/L) and during white lupin infection 24 and 84 h after inoculation (Dubrulle et al. 2020) was downloaded from NCBI (Table S5). Raw reads were trimmed using

Trimmomatic v. 0.39 (Bolger, Lohse, and Usadel 2014), with phred score of 33, a minimum length of 50, and a sliding window of 5:10. Quality was assessed using FastQC. Reads were mapped to the reference genome (CLUP02), using STAR v. 2.7.10b (Dobin et al. 2013), allowing for multiple read mapping (parameters set as --outFilterMultimapNmax 100 --winAnchorMultimapNmax 200 --outFilterMismatchNmax 3). HTSeq-count v. 2.0.5 (Anders, Pyl, and Huber 2015) was used to retrieve counts per gene and TE. Counts were normalised, and DEGs were identified using the R package *DESeq2* (Love, Huber, and Anders 2014) by analysing all treatments together.

4.8 | Data Analysis

Statistical analyses were performed with R using the packages *stats* and *multcomp* (Hothorn, Bretz, and Westfal 2012). For genome characteristics that did not follow assumptions of normality of residuals and homogeneity of variance, square root or logit (percentage data) were transformed. A Tukey HSD test ($p \leq 0.05$) was applied for pairwise mean comparisons. Differences between gene categories were compared using a Bonferroni-adjusted Dunn's test. Expression data was analysed through mean comparison to the control (24 h in liquid medium) using a Kruskal–Wallis test because a normal distribution of the residuals could not be achieved.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in the Sequence Read Archive at https://www.ncbi.nlm.nih.gov/sra. The data sequenced in this study are available under BioProject PRJNA1099210. Accession, BioSample, and assembly identifiers are listed in Table S1.

References

Alkemade, J. A., R. Baroncelli, M. M. Messmer, and P. Hohmann. 2023. "Attack of the Clones: Population Genetics Reveals Clonality of *Colletotrichum lupini*, the Causal Agent of Lupin Anthracnose." *Molecular Plant Pathology* 24: 616–627. https://doi.org/10.1111/mpp.13332.

Alkemade, J. A., M. M. Messmer, R. T. Voegele, M. R. Finckh, and P. Hohmann. 2021. "Genetic Diversity of *Colletotrichum lupini* and Its Virulence on White and Andean Lupin." *Scientific Reports* 11: 13547. https://doi.org/10.1038/s41598-021-92953-y.

Alonge, M., L. Lebeigle, M. Kirsche, et al. 2022. "Automated Assembly Scaffolding Using RagTag Elevates a New Tomato System for High-Throughput Genome Editing." *Genome Biology* 23: 258. https://doi.org/10.1186/s13059-022-02823-7.

Anders, S., P. T. Pyl, and W. Huber. 2015. "HTSeq—a Python Framework to Work with High-Throughput Sequencing Data." *Bioinformatics* 31: 166–169. https://doi.org/10.1093/bioinformatics/btu638.

Andrews, S. 2010. *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Cambridge, UK: Babraham Bioinformatics, Babraham Institute. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Badet, T., U. Oggenfuss, L. Abraham, B. A. McDonald, and D. Croll. 2020. "A 19-Isolate Reference-Quality Global Pangenome for the Fungal Wheat Pathogen *Zymoseptoria Tritici*." *BMC Biology* 18: 1–18. https://doi.org/10.1186/s12915-020-0744-3.

Bankevich, A., S. Nurk, D. Antipov, et al. 2012. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing." *Journal of Computational Biology* 19: 455–477. https://doi.org/10.1089/cmb.2012.0021.

Bao, W., K. K. Kojima, and O. Kohany. 2015. "Repbase Update, a Database of Repetitive Elements in Eukaryotic Genomes." *Mobile DNA* 6: 1–6. https://doi.org/10.1186/s13100-015-0041-9.

Baroncelli, R., J. F. Cobo-Díaz, T. Benocci, et al. 2024. "Genome Evolution and Transcriptome Plasticity is Associated with Adaptation to Monocot and Dicot Plants In *Colletotrichum* Fungi." *GigaScience* 13: giae036. https://doi.org/10.1093/gigascience/giae036.

Baroncelli, R., F. Pensec, D. Da Lio, et al. 2021. "Complete Genome Sequence of the Plant Pathogenic Fungus *Colletotrichum Lupini*." *Molecular Plant-Microbe Interactions* 34: 1461–1464. https://doi.org/10.1094/MPMI-07-21-0173-A.

Batson, A. M., L. Fokkens, M. Rep, and L. J. du Toit. 2021. "Putative Effector Genes Distinguish Two Pathogenicity Groups of *Fusarium Oxysporum* f. sp. *Spinaciae*." *Molecular Plant-Microbe Interactions* 34: 141–156. https://doi.org/10.1094/MPMI-06-20-0145-R.

Becerra, S., R. Baroncelli, T. R. Boufleur, S. A. Sukno, and M. R. Thon. 2023. "Chromosome-Level Analysis of the *Colletotrichum graminicola* Genome Reveals the Unique Characteristics of Core and Minichromosomes." *Frontiers in Microbiology* 14: 1129319. https://doi.org/10.3389/fmicb.2023.1129319.

Bertazzoni, S., A. H. Williams, D. A. Jones, R. A. Syme, K.-C. Tan, and J. K. Hane. 2018. "Accessories Make the Outfit: Accessory Chromosomes and Other Dispensable DNA Regions in Plant-Pathogenic Fungi." *Molecular Plant-Microbe Interactions* 31: 779–788. https://doi.org/10.1094/MPMI-06-17-0135-FI.

Bhadauria, V., R. MacLachlan, C. Pozniak, et al. 2019. "Genetic Map-Guided Genome Assembly Reveals a Virulence-Governing Minichromosome in the Lentil Anthracnose Pathogen *Colletotrichum lentis*." *New Phytologist* 221: 431–445. https://doi.org/10.1111/nph.15369.

Blin, K., S. Shaw, H. E. Augustijn, et al. 2023. "antiSMASH 7.0: New and Improved Predictions for Detection, Regulation, Chemical Structures and Visualisation." *Nucleic Acids Research*: gkad344. https://doi.org/10.1093/nar/gkad344.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bragança, C. A., U. Damm, R. Baroncelli, N. S. M. Júnior, and P. W. Crous. 2016. "Species of the *Colletotrichum acutatum* Complex Associated with

Anthracnose Diseases of Fruit in Brazil." *Fungal Biology* 120: 547–561. https://doi.org/10.1016/j.funbio.2016.01.011.

Buchfink, B., C. Xie, and D. H. Huson. 2015. "Fast and Sensitive Protein Alignment Using DIAMOND." *Nature Methods* 12: 59–60. https://doi.org/10.1038/nmeth.3176.

Cannon, P., U. Damm, P. Johnston, and B. Weir. 2012. "*Colletotrichum* – Current Status and Future Directions." *Studies in Mycology* 73: 181–213. https://doi.org/10.3114/sim0014.

Carreón-Anguiano, K. G., I. Islas-Flores, J. Vega-Arreguín, L. Sáenz-Carbonell, and B. Canto-Canché. 2020. "Effhunter: A Tool for Prediction of Effector Protein Candidates in Fungal Proteomic Databases." *Biomolecules* 10: 712. https://doi.org/10.3390/biom10050712.

Chen, S., Y. Zhou, Y. Chen, and J. Gu. 2018. "Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor." *Bioinformatics* 34: i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Cook, D. E., C. H. Mesarich, and B. P. Thomma. 2015. "Understanding Plant Immunity as a Surveillance System to Detect Invasion." *Annual Review of Phytopathology* 53: 541–563. https://doi.org/10.1146/annurev-phyto-080614-120114.

da Silva, L. L., H. L. A. Moreno, H. L. N. Correia, M. F. Santana, and M. V. de Queiroz. 2020. "Colletotrichum: Species Complexes, Lifestyle, and Peculiarities of Some Sources of Genetic Variability." *Applied Microbiology and Biotechnology* 104: 1891–1904. https://doi.org/10.1007/s00253-020-10363-y.

Dallery, J.-F., N. Lapalu, A. Zampounis, et al. 2017. "Gapless Genome Assembly of *Colletotrichum higginsianum* Reveals Chromosome Structure and Association of Transposable Elements with Secondary Metabolite Gene Clusters." *BMC Genomics* 18: 1–22. https://doi.org/10.1186/s12864-017-4083-x.

Damm, U., P. F. Cannon, J. H. C. Woudenberg, and P. W. Crous. 2012. "The Colletotrichum acutatum Species Complex." *Studies in Mycology* 73: 37–113. https://doi.org/10.3114/sim0010.

Danecek, P., J. K. Bonfield, J. Liddle, et al. 2021. "Twelve Years of SAMtools and BCFtools." *GigaScience* 10: giab008. https://doi.org/10.1093/gigascience/giab008.

de Jonge, R., M. D. Bolton, A. Kombrink, G. C. van den Berg, K. A. Yadeta, and B. P. Thomma. 2013. "Extensive Chromosomal Reshuffling Drives Evolution of Virulence in an Asexual Pathogen." *Genome Research* 23: 1271–1282. https://doi.org/10.1101/gr.152660.112.

Dobin, A., C. A. Davis, F. Schlesinger, et al. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29: 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Dong, S., S. Raffaele, and S. Kamoun. 2015. "The Two-Speed Genomes of Filamentous Pathogens: Waltz with Plants." *Current Opinion in Genetics & Development* 35: 57–65. https://doi.org/10.1016/j.gde.2015.09.001.

Dubrulle, G., A. Picot, S. Madec, et al. 2020. "Deciphering the Infectious Process of *Colletotrichum lupini* in Lupin Through Transcriptomic and Proteomic Analysis." *Microorganisms* 8: 1621. https://doi.org/10.3390/microorganisms8101621.

Emms, D. M., and S. Kelly. 2019. "OrthoFinder: Phylogenetic Orthology Inference for Comparative Genomics." *Genome Biology* 20: 1–14. https://doi.org/10.1186/s13059-019-1832-y.

Ewels, P., M. Magnusson, S. Lundin, and M. Käller. 2016. "MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report." *Bioinformatics* 32: 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

Faino, L., M. F. Seidl, X. Shi-Kunne, et al. 2016. "Transposons Passively and Actively Contribute to Evolution of the Two-Speed Genome of a Fungal Pathogen." *Genome Research* 26: 1091–1100. https://doi.org/10.1101/gr.204974.116.

Finn, R. D., J. Clements, and S. R. Eddy. 2011. "HMMER Web Server: Interactive Sequence Similarity Searching." *Nucleic Acids Research* 39: W29–W37. https://doi.org/10.1093/nar/gkr367.

Flynn, J. M., R. Hubley, C. Goubert, et al. 2020. "RepeatModeler2 for Automated Genomic Discovery of Transposable Element Families." *Proceedings of the National Academy of Sciences of the United States of America* 117: 9451–9457. https://doi.org/10.1073/pnas.1921046117.

Fouché, S., T. Badet, U. Oggenfuss, C. Plissonneau, C. S. Francisco, and D. Croll. 2020. "Stress-Driven Transposable Element De-Repression Dynamics and Virulence Evolution in a Fungal Pathogen." *Molecular Biology and Evolution* 37: 221–239. https://doi.org/10.1093/molbev/msz216.

Fouché, S., U. Oggenfuss, E. Chanclud, and D. Croll. 2022. "A Devil's Bargain with Transposable Elements in Plant Pathogens." *Trends in Genetics* 38: 222–230. https://doi.org/10.1016/j.tig.2021.08.005.

Frantzeskakis, L., B. Kracher, S. Kusch, et al. 2018. "Signatures of Host Specialization and a Recent Transposable Element Burst in the Dynamic One-Speed Genome of the Fungal Barley Powdery Mildew Pathogen." *BMC Genomics* 19: 1–23. https://doi.org/10.1186/s12864-018-4750-6.

Gan, P., R. Hiroyama, A. Tsushima, et al. 2021. "Telomeres and a Repeat-Rich Chromosome Encode Effector Gene Clusters in Plant Pathogenic *Colletotrichum* Fungi." *Environmental Microbiology* 23: 6004–6018. https://doi.org/10.1111/1462-2920.15490.

Gan, P., M. Narusaka, N. Kumakura, et al. 2016. "Genus-Wide Comparative Genome Analyses of *Colletotrichum* Species Reveal Specific Gene Family Losses and Gains During Adaptation to Specific Infection Lifestyle." *Genome Biology and Evolution* 8: 1467–1481. https://doi.org/10.1093/gbe/evw089.

Gladieux, P., A. Feurtey, M. E. Hood, et al. 2016. "The Population Biology of Fungal Invasions." In *Invasion Genetics: The Baker and Stebbins Legacy*, 81–100. https://doi.org/10.1111/mec.13028.

Grandaubert, J., M.-H. Balesdent, and T. Rouxel. 2014. "Evolutionary and Adaptive Role of Transposable Elements in Fungal Genomes." In *Advances in Botanical Research*, 79–107. Elsevier.

Gu, Z., L. Gu, R. Eils, M. Schlesner, and B. Brors. 2014. "Circlize Implements and Enhances Circular Visualization in R." *Bioinformatics* 30: 2811–2812. https://doi.org/10.1093/bioinformatics/btu393.

Gupta, Y. K., F. C. Marcelino-Guimarães, C. Lorrain, et al. 2023. "Major Proliferation of Transposable Elements Shaped the Genome of the Soybean Rust Pathogen *Phakopsora Pachyrhizi*." *Nature Communications* 14: 1835. https://doi.org/10.1038/s41467-023-37551-4.

Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler. 2013. "QUAST: Quality Assessment Tool for Genome Assemblies." *Bioinformatics* 29: 1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

Habig, M., J. Quade, and E. H. Stukenbrock. 2017. "Forward Genetics Approach Reveals Host Genotype-Dependent Importance of Accessory Chromosomes in the Fungal Wheat Pathogen *Zymoseptoria tritici*." *mBio* 8: e01919-17. https://doi.org/10.1128/mBio.01919-17.

Habig, M., and E. H. Stukenbrock. 2020. "Origin, Function, and Transmission of Accessory Chromosomes." In *The Mycota. Genetics and Biotechnology*, edited by K. Esser, vol. 2, 25–47. Cham, Switzerland: Springer.

Horton, P., K.-J. Park, T. Obayashi, et al. 2007. "WoLF PSORT: Protein Localization Predictor." *Nucleic Acids Research* 35: W585–W587. https://doi.org/10.1093/nar/gkm259.

Hothorn, T., F. Bretz, and P. Westfal. 2012. "Multcomp: Simultaneous Tests and Confidence Intervals for General Linear Hypotheses in Parametric Models." *R Package Version* 1: 2–13.

Jones, P., D. Binns, H.-Y. Chang, et al. 2014. "InterProScan 5: Genome-Scale Protein Function Classification." *Bioinformatics* 30: 1236–1240. https://doi.org/10.1093/bioinformatics/btu031.

Käll, L., A. Krogh, and E. L. Sonnhammer. 2004. "A Combined Transmembrane Topology and Signal Peptide Prediction Method." *Journal of Molecular Biology* 338: 1027–1036. https://doi.org/10.1016/j.jmb.2004.03.016.

Katoh, K., and D. M. Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30: 772–780. https://doi.org/10.1093/molbev/mst010.

Keller, O., M. Kollmar, M. Stanke, and S. Waack. 2011. "A Novel Hybrid Gene Prediction Method Employing Protein Multiple Sequence Alignments." *Bioinformatics* 27: 757–763. https://doi.org/10.1093/bioinformatics/btr010.

Korf, I. 2004. "Gene Finding in Novel Genomes." *BMC Bioinformatics* 5: 1–9. https://doi.org/10.1186/1471-2105-5-59.

Krishnan, P., X. Ma, B. A. McDonald, and P. C. Brunner. 2018. "Widespread Signatures of Selection for Secreted Peptidases in a Fungal Plant Pathogen." *BMC Evolutionary Biology* 18: 1–10. https://doi.org/10.1186/s12862-018-1123-3.

Krogh, A., B. Larsson, G. Von Heijne, and E. L. Sonnhammer. 2001. "Predicting Transmembrane Protein Topology With a Hidden Markov Model: Application to Complete Genomes." *Journal of Molecular Biology* 305: 567–580. https://doi.org/10.1006/jmbi.2000.4315.

Krzywinski, M., J. Schein, I. Birol, et al. 2009. "Circos: An Information Aesthetic for Comparative Genomics." *Genome Research* 19: 1639–1645. https://doi.org/10.1101/gr.092759.109.

Kurtz, S., A. Phillippy, A. L. Delcher, et al. 2004. "Versatile and Open Software for Comparing Large Genomes." *Genome Biology* 5: 1–9. https://doi.org/10.1186/gb-2004-5-2-r12.

Laetsch, D. R., and M. L. Blaxter. 2017. "BlobTools: Interrogation of Genome Assemblies." *F1000Research* 6: 1287. https://doi.org/10.12688/f1000research.12232.1.

Langner, T., A. Harant, L. B. Gomez-Luciano, et al. 2021. "Genomic Rearrangements Generate Hypervariable Mini-Chromosomes in Host-Specific Isolates of the Blast Fungus." *PLoS Genetics* 17: e1009386. https://doi.org/10.1371/journal.pgen.1009386.

Lapalu, N., A. Simon, A. Lu, et al. 2024. "Complete Genome of the Medicago Anthracnose Fungus, *Colletotrichum destructivum*, Reveals a Mini-Chromosome-Like Region Within a Core Chromosome." *Microbial Genomics* 10: 001283. https://doi.org/10.1099/mgen.0.001283.

Latorre, S. M., V. M. Were, A. J. Foster, et al. 2023. "Genomic Surveillance Uncovers a Pandemic Clonal Lineage of the Wheat Blast Fungus." *PLoS Biology* 21: e3002052. https://doi.org/10.1371/journal.pbio.3002052.

Le Naour-Vernet, M., F. Charriat, J. Gracy, et al. 2023. "Adaptive Evolution in Virulence Effectors of the Rice Blast Fungus *Pyricularia oryzae*." *PLoS Pathogens* 19: e1011294. https://doi.org/10.1371/journal.ppat.1011294.

Li, H. 2013. "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM." https://doi.org/10.48550/arXiv.1303.3997.

Liang, X., W. Yu, Y. Meng, et al. 2024. "Genome Comparisons Reveal Accessory Genes Crucial for the Evolution of Apple Glomerella Leaf Spot Pathogenicity in *Colletotrichum* Fungi." *Molecular Plant Pathology* 25: e13454. https://doi.org/10.1111/mpp.13454.

Linheiro, R. S., and C. M. Bergman. 2012. "Whole Genome Resequencing Reveals Natural Target Site Preferences of Transposable Elements in *Drosophila melanogaster*." *PLoS One* 7: e30008. https://doi.org/10.1371/journal.pone.0030008.

Love, M. I., W. Huber, and S. Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data With DESeq2." *Genome Biology* 15: 1–21. https://doi.org/10.1186/s13059-014-0550-8.

Ma, L.-J., H. C. Van Der Does, K. A. Borkovich, et al. 2010. "Comparative Genomics Reveals Mobile Pathogenicity Chromosomes in *Fusarium*." *Nature* 464: 367–373. https://doi.org/10.1038/nature08850.

Ma, W., J. Yang, J. Ding, et al. 2023. "CRISPR/Cas9-Mediated Deletion of Large Chromosomal Segments Identifies a Minichromosome Modulating the *Colletotrichum graminicola* Virulence on Maize." *International Journal of Biological Macromolecules* 245: 125462. https://doi.org/10.1016/j.ijbiomac.2023.125462.

Manni, M., M. R. Berkeley, M. Seppey, F. A. Simao, and E. M. Zdobnov. 2021. "BUSCO Update: Novel and Streamlined Workflows Along With Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes." *Molecular Biology and Evolution* 38: 4647–4654. https://doi.org/10.1093/molbev/msab199.

Marçais, G., and C. Kingsford. 2011. "A Fast, Lock-Free Approach for Efficient Parallel Counting of Occurrences of k-Mers." *Bioinformatics* 27: 764–770. https://doi.org/10.1093/bioinformatics/btr011.

McCarthy, C. G., and D. A. Fitzpatrick. 2019. "Pan-Genome Analyses of Model Fungal Species." *Microbial Genomics* 5: e000243. https://doi.org/10.1099/mgen.0.000243.

McTaggart, A. R., T. James, R. Shivas, et al. 2021. "Population Genomics Reveals Historical and Ongoing Recombination in the *Fusarium oxysporum* Species Complex." *Studies in Mycology* 99: 100132. https://doi.org/10.1016/j.simyco.2021.100132.

Minas, K., N. R. McEwan, C. J. Newbold, and K. P. Scott. 2011. "Optimization of a High-Throughput CTAB-Based Protocol for the Extraction of qPCR-Grade DNA From Rumen Fluid, Plant and Bacterial Pure Cultures." *FEMS Microbiology Letters* 325: 162–169. https://doi.org/10.1111/j.1574-6968.2011.02424.x.

Möller, M., and E. H. Stukenbrock. 2017. "Evolution and Genome Architecture in Fungal Plant Pathogens." *Nature Reviews Microbiology* 15: 756–771. https://doi.org/10.1038/nrmicro.2017.76.

Muszewska, A., K. Steczkiewicz, M. Stepniewska-Dziubinska, and K. Ginalski. 2019. "Transposable Elements Contribute to Fungal Genes and Impact Fungal Lifestyle." *Scientific Reports* 9: 4307. https://doi.org/10.1038/s41598-019-40965-0.

Nakamoto, A. A., P. M. Joubert, and K. V. Krasileva. 2023. "Intraspecific Variation of Transposable Elements Reveals Differences in the Evolutionary History of Fungal Phytopathogen Pathotypes." *Genome Biology and Evolution* 15: evad206. https://doi.org/10.1093/gbe/evad206.

Naranjo-Ortiz, M. A., and T. Gabaldón. 2020. "Fungal Evolution: Cellular, Genomic and Metabolic Complexity." *Biological Reviews* 95: 1198–1232. https://doi.org/10.1111/brv.12605.

Oggenfuss, U., T. Badet, T. Wicker, et al. 2021. "A Population-Level Invasion by Transposable Elements Triggers Genome Expansion in a Fungal Pathogen." *eLife* 10: e69249. https://doi.org/10.7554/eLife.69249.

Ordonez, N., M. F. Seidl, C. Waalwijk, et al. 2015. "Worse Comes to Worst: Bananas and Panama Disease – When Plant and Pathogen Clones Meet." *PLoS Pathogens* 11: e1005197. https://doi.org/10.1371/journal.ppat.1005197.

Pavlidis, P., D. Živković, A. Stamatakis, and N. Alachiotis. 2013. "SweeD: Likelihood-Based Detection of Selective Sweeps in Thousands of Genomes." *Molecular Biology and Evolution* 30: 2224–2234. https://doi.org/10.1093/molbev/mst112.

Plaumann, P.-L., J. Schmidpeter, M. Dahl, L. Taher, and C. Koch. 2018. "A Dispensable Chromosome Is Required for Virulence in the Hemibiotrophic Plant Pathogen *Colletotrichum higginsianum*." *Frontiers in Microbiology* 9: 1005. https://doi.org/10.3389/fmicb.2018.01005.

Plett, J. M., and F. Martin. 2015. "Reconsidering Mutualistic Plant–Fungal Interactions Through the Lens of Effector Biology." *Current Opinion in Plant Biology* 26: 45–50. https://doi.org/10.1016/j.pbi.2015.06.001.

Porquier, A., C. Tisserant, F. Salinas, et al. 2021. "Retrotransposons as Pathogenicity Factors of the Plant-Pathogenic Fungus *Botrytis cinerea*." *Genome Biology* 22: 225. https://doi.org/10.1186/s13059-021-02446-4.

Qian, H., L. Song, L. Wang, B. Wang, and W. Liang. 2022. "The Secreted FoAPY1 Peptidase Promotes *Fusarium oxysporum* Invasion." *Frontiers in Microbiology* 13: 1040302. https://doi.org/10.3389/fmicb.2022.1040302.

Quinlan, A. R., and I. M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics* 26: 841–842. https://doi.org/10.1093/bioinformatics/btq033.

R Core Team. 2020. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rao, S., S. Sharda, V. Oddi, and M. R. Nandineni. 2018. "The Landscape of Repetitive Elements in the Refined Genome of Chilli Anthracnose Fungus *Colletotrichum truncatum*." *Frontiers in Microbiology* 9: 390001. https://doi.org/10.3389/fmicb.2018.02367.

Redkar, A., M. Sabale, C. Schudoma, et al. 2022. "Conserved Secreted Effectors Contribute to Endophytic Growth and Multihost Plant Compatibility in a Vascular Wilt Fungus." *Plant Cell* 34: 3214–3232. https://doi.org/10.1093/plcell/koac174.

Riegel, R., D. Veliz, I. von Baer, Y. Quitral, and M. Munoz. 2010. "Genetic Diversity and Virulence of *Colletotrichum lupini* Isolates Collected in Chile." *Tropical Plant Pathology* 35: 144–152.

Rogério, F., R. Baroncelli, F. B. Cuevas-Fernández, et al. 2022. "Population Genomics Provide Insights Into the Global Genetic Structure of *Colletotrichum graminicola*, the Causal Agent of Maize Anthracnose." *mBio* 14: e0287822. https://doi.org/10.1128/mbio.02878-22.

Sánchez-Vallet, A., S. Fouché, I. Fudal, et al. 2018. "The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens." *Annual Review of Phytopathology* 56: 21–40. https://doi.org/10.1146/annurev-phyto-080516-035303.

Seidl, M. F., and B. P. Thomma. 2017. "Transposable Elements Direct the Coevolution Between Plants and Microbes." *Trends in Genetics* 33: 842–851. https://doi.org/10.1016/j.tig.2017.07.003.

Snelders, N. C., H. Rovenich, and B. P. Thomma. 2022. "Microbiota Manipulation Through the Secretion of Effector Proteins Is Fundamental to the Wealth of Lifestyles in the Fungal Kingdom." *FEMS Microbiology Reviews* 46: fuac022. https://doi.org/10.1093/femsre/fuac022.

Sperschneider, J., and P. N. Dodds. 2022. "EffectorP 3.0: Prediction of Apoplastic and Cytoplasmic Effectors in Fungi and Oomycetes." *Molecular Plant–Microbe Interactions* 35: 146–156. https://doi.org/10.1094/MPMI-08-21-0201-R.

Stalder, L., U. Oggenfuss, N. Mohd-Assaad, and D. Croll. 2023. "The Population Genetics of Adaptation Through Copy Number Variation in a Fungal Plant Pathogen." *Molecular Ecology* 32: 2443–2460. https://doi.org/10.1111/mec.16435.

Storer, J., R. Hubley, J. Rosen, T. J. Wheeler, and A. F. Smit. 2021. "The Dfam Community Resource of Transposable Element Families, Sequence Models, and Genome Annotations." *Mobile DNA* 12: 1–14. https://doi.org/10.1186/s13100-020-00230-y.

Suyama, M., D. Torrents, and P. Bork. 2006. "PAL2NAL: Robust Conversion of Protein Sequence Alignments Into the Corresponding Codon Alignments." *Nucleic Acids Research* 34: W609–W612. https://doi.org/10.1093/nar/gkl315.

Talhinhas, P., and R. Baroncelli. 2021. "*Colletotrichum* Species and Complexes: Geographic Distribution, Host Range and Conservation Status." *Fungal Diversity* 110: 109–198. https://doi.org/10.1007/s13225-021-00491-9.

Taylor, J., D. Jacobson, and M. Fisher. 1999. "The Evolution of Asexual Fungi: Reproduction, Speciation and Classification." *Annual Review of Phytopathology* 37: 197–246. https://doi.org/10.1146/annurev.phyto.37.1.197.

Tellier, A., S. Moreno-Gámez, and W. Stephan. 2014. "Speed of Adaptation and Genomic Footprints of Host–Parasite Coevolution Under Arms Race and Trench Warfare Dynamics." *Evolution* 68: 2211–2224. https://doi.org/10.1111/evo.12427.

Teufel, F., J. J. Almagro Armenteros, A. R. Johansen, et al. 2022. "SignalP 6.0 Predicts all Five Types of Signal Peptides Using Protein Language Models." *Nature Biotechnology* 40: 1–3. https://doi.org/10.1038/s41587-021-01156-3.

Todd, R. T., A. Forche, and A. Selmecki. 2017. "Ploidy Variation in Fungi: Polyploidy, Aneuploidy, and Genome Evolution." *Microbiology Spectrum* 5: 1128. https://doi.org/10.1128/microbiolspec.funk-0051-2016.

Torres, D. E., U. Oggenfuss, D. Croll, and M. F. Seidl. 2020. "Genome Evolution in Fungal Plant Pathogens: Looking Beyond the Two-Speed Genome Model." *Fungal Biology Reviews* 34: 136–143. https://doi.org/10.1016/j.fbr.2020.07.001.

Torres, D. E., B. P. Thomma, and M. F. Seidl. 2021. "Transposable Elements Contribute to Genome Dynamics and Gene Expression Variation in the Fungal Plant Pathogen *Verticillium dahliae*." *Genome Biology and Evolution* 13: evab135. https://doi.org/10.1093/gbe/evab135.

Tsushima, A., P. Gan, N. Kumakura, et al. 2019. "Genomic Plasticity Mediated by Transposable Elements in the Plant-Pathogenic Fungus *Colletotrichum higginsianum*." *Genome Biology and Evolution* 11: 1487–1500. https://doi.org/10.1093/gbe/evz087.

van Dam, P., L. Fokkens, S. M. Schmidt, et al. 2016. "Effector Profiles Distinguish Formae Speciales of *Fusarium oxysporum*." *Environmental Microbiology* 18: 4087–4102. https://doi.org/10.1111/1462-2920.13445.

Van Kan, J. A., J. H. Stassen, A. Mosbach, et al. 2017. "A Gapless Genome Sequence of the Fungus *Botrytis cinerea*." *Molecular Plant Pathology* 18: 75–89. https://doi.org/10.1111/mpp.12384.

van Rossum, B.-J., W. Kruijer, F. van Eeuwijk, et al. 2020. Package 'statgenGWAS'. https://CRAN.R-project.org/package=statgenGWAS.

van Westerhoven, A. C., C. Aguilera-Galvez, G. Nakasato-Tagami, et al. 2024. "Segmental Duplications Drive the Evolution of Accessory Regions in a Major Crop Pathogen." *New Phytologist* 242: 610–625. https://doi.org/10.1111/nph.19604.

van Westerhoven, A. C., H. J. Meijer, M. F. Seidl, and G. H. Kema. 2022. "Uncontained Spread of *Fusarium* Wilt of Banana Threatens African Food Security." *PLoS Pathogens* 18: e1010769. https://doi.org/10.1094/PDIS-07-22-1576-SC.

Voigt, C. A., W. Schäfer, and S. Salomon. 2005. "A Secreted Lipase of *Fusarium graminearum* Is a Virulence Factor Required for Infection of Cereals." *Plant Journal* 42: 364–375. https://doi.org/10.1111/j.1365-313X.2005.02377.x.

Wang, H., R. Huang, J. Ren, et al. 2023. "The Evolution of Mini-Chromosomes in the Fungal Genus *Colletotrichum*." *mBio* 14: e00629-23. https://doi.org/10.1128/mbio.00629-23.

Wells, J. N., and C. Feschotte. 2020. "A Field Guide to Eukaryotic Transposable Elements." *Annual Review of Genetics* 54: 539–561. https://doi.org/10.1146/annurev-genet-040620-022145.

Wilson, A. M., R. V. Lelwala, P. W. Taylor, M. J. Wingfield, and B. D. Wingfield. 2021. "Unique Patterns of Mating Pheromone Presence and Absence Could Result in the Ambiguous Sexual Behaviors of *Colletotrichum* Species." *G3 Genes Genomes Genetics* 11: jkab187. https://doi.org/10.1093/g3journal/jkab187.

Yang, Z. 2007. "PAML 4: Phylogenetic Analysis by Maximum Likelihood." *Molecular Biology and Evolution* 24: 1586–1591. https://doi.org/10.1093/molbev/msm088.

Zheng, J., Q. Ge, Y. Yan, X. Zhang, L. Huang, and Y. Yin. 2023. "dbCAN3: Automated Carbohydrate-Active Enzyme and Substrate Annotation." *Nucleic Acids Research* 51: W115–W121. https://doi.org/10.1093/nar/gkad328.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.