# UC Davis
## UC Davis Previously Published Works

**Title**

A chromosome-scale genome assembly of the grape powdery mildew pathogen Erysiphe necator reveals its genomic architecture and previously unknown features of its biology.

**Permalink**

https://escholarship.org/uc/item/49n8m65q

**Journal**

mBio, 14(4)

**Authors**

Zaccaron, Alex
Neill, Tara
Corcoran, Jacob
et al.

**Publication Date**

2023-08-31

**DOI**

10.1128/mbio.00645-23

Peer reviewed

# A chromosome-scale genome assembly of the grape powdery mildew pathogen *Erysiphe necator* reveals its genomic architecture and previously unknown features of its biology

Alex Z. Zaccaron,[1] Tara Neill,[2] Jacob Corcoran,[2] Walter F. Mahaffee,[2] Ioannis Stergiopoulos[1]

**AUTHOR AFFILIATIONS** See affiliation list on p. 15.

**ABSTRACT** *Erysiphe necator* is an obligate fungal pathogen that causes grape powdery mildew, globally the most important disease on grapevines. Previous attempts to obtain a quality genome assembly for this pathogen were hindered by its high repetitive DNA content. Here, chromatin conformation capture (Hi-C) with long-read PacBio sequencing was combined to obtain a chromosome-scale assembly and a high-quality annotation for *E. necator* isolate EnFRAME01. The resulting 81.1 Mb genome assembly is 98% complete and consists of 34 scaffolds, 11 of which represent complete chromosomes. All chromosomes contain large centromeric-like regions and lack synteny to the 11 chromosomes of the cereal PM pathogen *Blumeria graminis*. Further analysis of their composition showed that repeats and transposable elements (TEs) occupy 62.7% of their content. TEs were almost evenly interspersed outside centromeric and telomeric regions and massively overlapped with regions of annotated genes, suggesting that they could have a significant functional impact. Abundant gene duplicates were observed as well, particularly in genes encoding candidate secreted effector proteins. Moreover, younger in age gene duplicates exhibited more relaxed selection pressure and were more likely to be located physically close in the genome than older duplicates. A total of 122 genes with copy number variations among six isolates of *E. necator* were also identified and were enriched in genes that were duplicated in EnFRAME01, indicating they may reflect an adaptive variation. Taken together, our study illuminates higher-order genomic architectural features of *E. necator* and provides a valuable resource for studying genomic structural variations in this pathogen.

**IMPORTANCE** Grape powdery mildew caused by the ascomycete fungus *Erysiphe necator* is economically the most important and recurrent disease in vineyards across the world. The obligate biotrophic nature of *E. necator* hinders the use of typical genetic methods to elucidate its pathogenicity and adaptation to adverse conditions, and thus comparative genomics has been a major method to study its genome biology. However, the current reference genome of *E. necator* isolate C-strain is highly fragmented with many non-coding regions left unassembled. This incompleteness prohibits in-depth comparative genomic analyses and the study of genomic structural variations (SVs) that are known to affect several aspects of microbial life, including fitness, virulence, and host adaptation. By obtaining a chromosome-scale genome assembly and a high-quality gene annotation for *E. necator*, we reveal the organization of its chromosomal content, unearth previously unknown features of its biology, and provide a reference for studying genomic SVs in this pathogen.

**KEYWORDS** Erysiphales, genome architecture, biotrophic lifestyle, transposable elements, gene duplications, copy number variation

*E* *rysiphe necator* (Ascomycetes; Leotiomycetes, Erysiphaceae) is an obligate biotrophic fungal pathogen that causes grapevine powdery mildew (GPM), one of the most common and economically important fungal diseases in vineyards around the globe (1). The pathogen can significantly reduce grape yield and quality, and most cultivated varieties of grapevine (*Vitis vinifera*) are susceptible to it (1–3). As a consequence, GPM is commonly managed by fungicides, which dramatically increase the overall production costs and the risk of resistance development (4, 5).

The obligate nature of powdery mildews (PMs) prohibits the functional analysis of their genes by means of standard genetic manipulations. Instead, comparative and population genomics have been used as alternatives for studying the molecular mechanisms underlying obligate biotrophy, pathogenicity, and other aspects of the biology of these pathogens (6–12). To date, the genomes of at least 16 species or formae speciales of PMs have been obtained, including 3 monocotyledonous-infecting and 13 dicotyledonous-infecting species. However, the highly repetitive nature of these genomes has posed major challenges to the construction of high-quality genome assemblies based on short sequencing reads alone and has further hindered in-depth comparative genome analyses. As a result, chromosome-scale genome assemblies have, so far, only been obtained from just two monocot-infecting species of PM, namely the wheat pathogen *Blumeria graminis* f.sp. *tritici* and the triticale pathogen *B. graminis* f.sp. *triticale* (9, 13) but none from dicot-infecting PMs.

Despite challenges in obtaining high-quality genome assemblies and annotations for PM fungi, analysis of their genomic content has shown that they possess some of the largest genomes among filamentous ascomycetes, with sizes typically ranging from 120 to 180 Mb (6, 9, 14, 15). The increase is due to the extensive proliferation of transposable elements (TEs) in their genomes, which typically comprise up to 85% of their genomic content (6, 7, 9, 11). However, contrary to their enlarged genomes, PMs have a reduced number of circa (ca.) 7,000 genes (11), which are considerably smaller compared with the ca. 11,000 genes typically present in non-obligate fungal plant pathogens (16). The reduction is due to marked losses in genes encoding key enzymes in primary and secondary metabolism as well as in hydrolytic enzymes that cause damage to host cells during infection, a hallmark of their obligate biotrophic lifestyle (10, 17). PMs also lack a repeat-induced point mutation (RIP) defense mechanism against the deleterious effects caused by TE replication in their genomes (18, 19). As a consequence, their genomes experience higher rates of TE and gene duplication and retention, as these are more prone to pseudogenization in species with an active RIP mechanism. An examination of their genome architecture has further shown that PM genomes are generally deprived of large-scale compartmentalization, AT-rich isochores, and accessory chromosomes, which constitute signatures of "plastic" or "two-speed" genomes (6, 9). Instead, they adhere mostly to the "one-speed" model of genome evolution, in which gene duplication is an important mechanism of evolution and adaptation (20).

Previous efforts to sequence the genome of *E. necator* were constrained by its highly repetitive nature and the limitations of the short-read sequencing technologies used at the time. Consequently, the current reference genome of *E. necator* isolate C-strain (C-strain) is estimated to be 36.5%–48.6% complete and is assembled into 5,935 scaffolds, which forbid a rigorous analysis of its architecture (7). In this study, we present a chromosome-scale genome assembly and gene annotation for *E. necator* isolate FRAME01 (EnFRAME01). The new reference genome of *E. necator* presented herein is the first chromosome-scale assembly obtained for a dicot-infecting PM species and elucidates major aspects of their biology.

## MATERIALS AND METHODS

A detailed version of materials and methods is provided in Supplementary Results at https://doi.org/10.5281/zenodo.7738565.

## Fungal isolate, nucleic acid extraction, and sequencing

*E. necator* isolate EnFRAME01 was isolated from greenhouse-grown grapes in Corvallis, Oregon, USA, in 2018. EnFRAME01 was propagated by dusting from detached leaves (21) on *V. vinifera* L., cv. "Chardonnay" seedlings grown hydroponically in half-strength Hoagland's solution (22). High-molecular weight DNA was obtained from conidia as in Feehan et al. (23) with modifications. PacBio library construction and sequencing were outsourced to the DNA Technologies and Expression Analysis Core Laboratory at the UC Davis Genome Center. The constructed library was sequenced using two SMRT (Single-Molecule Real Time sequencing) Cells 1M v2 on a Sequel Chemistry v2 platform (Pacific Biosciences, Menlo Park, CA, USA). Extracted DNA was also used to generate an Illumina WGS library and a Hi-C library using the Proximo Hi-C Kit (microbial) (Phase Genomics), according to the manufacturer's instructions. Both Illumina libraries were sequenced on a NovaSeq 6000 instrument (PE150 format). To assist gene prediction, total RNA was extracted from conidia of *E. necator* isolate EnFRAME01 held at the USDA-ARS Horticultural Crops Disease and Pest Management Research Unit in Corvallis, Oregon, using Trizol reagent (ThermoFisher) according to the manufacturer's instructions. Sample integrity analysis, cDNA library preparation, and sequencing on the Illumina NovaSeq 6000 platform using the paired-end (PE150) format were carried out at Novogene, Inc. (Sacramento, CA, USA).

## Genome assembly and annotation of repetitive DNA

PacBio reads were assembled with Canu v1.8 (24) and then used to polish the contigs with pbmm2 and Arrow from the GenomicConsensus package v2.3.3. The assembly was further polished with Pilon v1.23 (25) after mapping the Illumina reads with BWA-MEM v0.7.17 (26). Hi-C reads were mapped with BWA-MEM v0.7.17, and chromatin interaction frequencies were estimated with the 3D-DNA package (27). They were then visualized with Juicebox v1.11.08 (28), which allowed the grouping of contigs into putative chromosomes. Repetitive regions were identified with RepeatModeler v2.0.2a (29) and masked with RepeatMasker v4.1.2-p1. The repeat divergence landscape was estimated with the script *parseRM.pl* v5.8.2.

## Gene prediction

RNA-seq reads were mapped to the genome assembly with HISAT2 v2.2.0 (30), and transcripts were reconstructed with Stringtie v2.1.1 (31) and Trinity v2.9.1 (32). Genes were predicted with Maker v2.31.10 (33) by integrating (i) the trained *ab initio* predictors GeneMark-ES v4.57 (34), SNAP v2013-11-29 (35), and Augustus v3.2.3 (36), (ii) gene models generated with GeMoMa (37), (iii) assembled transcripts, and (iv) protein sequences from close relative species.

## Homology-based functional annotations

Conserved PFAM domains were identified with InterProScan v5.32-71.0 (38) or the NCBI CDD database (39). Carbohydrate-active enzymes (CAZymes) were predicted with dbCAN2 (40). Proteases and transporters were classified based on the top BLASTp hit (E-value <1E-10) against the MEROPS database v12.1 (41) and the TCDB database version of 2020-07-12 (42), respectively. Secreted proteins (SPs) were predicted with SignalP v5.0 (43). Membrane-bound proteins were predicted with PredGPI (44) and TMHMM v2.0 (45). Secreted proteins were classified into CSEPs based on three lines of evidence (Fig. S21): (i) EffectorP v2.0 (46); (ii) proteins shorter than 250 aa with at least 2% cysteines; and (iii) proteins with no homologs in Leotiomycetes, except Erysiphales, based on a BLASTp search (E-value <1E-3) against 93 Leotiomycetes genomes.

## Identification of core genes missing in EnFRAME01

Protein sequences from *E. necator*, *B. graminis* f.sp. *hordei*, *B. cinerea*, *Zymoseptoria tritici*, *Aspergillus niger*, *Neurospora crassa*, and *Saccharomyces cerevisiae* were organized into

orthogroups with OrthoFinder v2.5.4 (47). Orthogroups containing proteins from all non-PM species but not from *E. necator* were considered core genes missing in *E. necator*.

## Classification and enrichment of duplicated genes

Duplicated genes were identified based on an *all-vs-all* BLASTp (E-value < 1E-5) search, with minimum identity of 40% and minimum coverage of 50%. The script *duplicate_gene_classifier* from MCScanX (48) was used to classify gene duplications into dispersed, proximal, or tandem. Enrichment of gene categories within duplicated genes was performed with hypergeometric tests using the *phyper* function within R v4.1.2. Pairwise $K_A/K_S$ ratios were estimated with $K_A/K_S$_calculator v3 (49). Conserved domain enrichment was performed with the *enricher* function from the R package clusterProfiler v4.2.2 (50) within R v4.1.2 with adjusted *P*-value <0.01.

## Identification of CNVs

Whole-genome sequencing (WGS) reads of five *E. necator* isolates (7) were mapped to the genome with BWA-MEM v0.7.17 (26). PCR duplicates were marked with samblaster v0.1.24 (51) and removed with SAMtools v1.9 (52). Copy number variation (CNV) regions were identified with CNVnator (53). Genes with at least 80% overlapping with CNV regions were considered CNV genes.

## Comparative analysis of carboxylesterases

The predicted carboxylesterase (CE) HI914_00624 was queried with BLASTp against the NCBI nr database (2022, 08-13), and proteins from EnFRAME01 and the 400 most similar sequences (E-value <1E-50) were obtained. The acetylcholinesterase DmAChE from *Drosophila melanogaster* (1QO9) (54) was included as an outgroup and also used as reference to identify conserved residues. The 401 amino acid sequences were aligned with MAFFT v7.490 (55), and sites composed of more than 50% gaps were removed with trimAl v1.4 (56). The phylogenetic tree was inferred with IQ-TREE v1.6.12 (57) using the built-in ModelFinder (58) and 1,000 rapid bootstrap replicates (59). The tree was visualized and edited with iTOL (60). Quantitative PCR (qPCR) and quantitative reverse transcription PCR (RT-qPCR) were used to determine the copy number and gene expression of the *HI914_00624* gene, respectively, in six isolates of *E. necator*. qPCR reactions were run in triplicate on an Applied Biosystems QuantStudio5 qPCR machine using PerfeCTa qPCR ToughMix Low ROX (Quantabio) and the primers and probes listed in Table S26.

## RESULTS

### The genome of *E. necator* consists of 11 chromosomes with large centromeric-like regions

The genome of EnFRAME01 was assembled using a combination of PacBio reads and Hi-C data into 34 scaffolds, totaling 81.1 Mb in size. This is a considerable improvement over the previous reference genome of C-strain that was fragmented into 5,935 scaffolds (Table S1 and Supplementary Results). Of the 34 assembled scaffolds, 11 embodied distinct chromosomes (Chr1-to-Chr11) (Fig. 1; Fig. S1; Table 1), 22 were unplaced scaffolds, and 1 scaffold represented the complete mitochondrial genome (Fig. S2). The size of the 11 chromosomes ranged from 11.3 Mb (Chr1) to 3.3 Mb (Chr11), and all were putatively assembled telomere-to-telomere, containing only five collapsed regions (Fig. S3). All chromosomes had 22–31 copies of the canonical telomeric repeat 5′-TTAGGG-3′ at their ends and predicted centromeric regions with high inter-chromosomal Hi-C contact frequency (Fig. S1A), as previously observed in other fungi (61, 62). However, in contrast to other ascomycetes (61, 63–66), the predicted centromeres of *E. necator* were large segments that accounted for 15.8% of the genome (Table 1 and Table S2). Centromeric regions of similar sizes have been reported before for the wheat PM *B. graminis* f.sp. *tritici* (9), but whole-genome alignment showed that the predicted
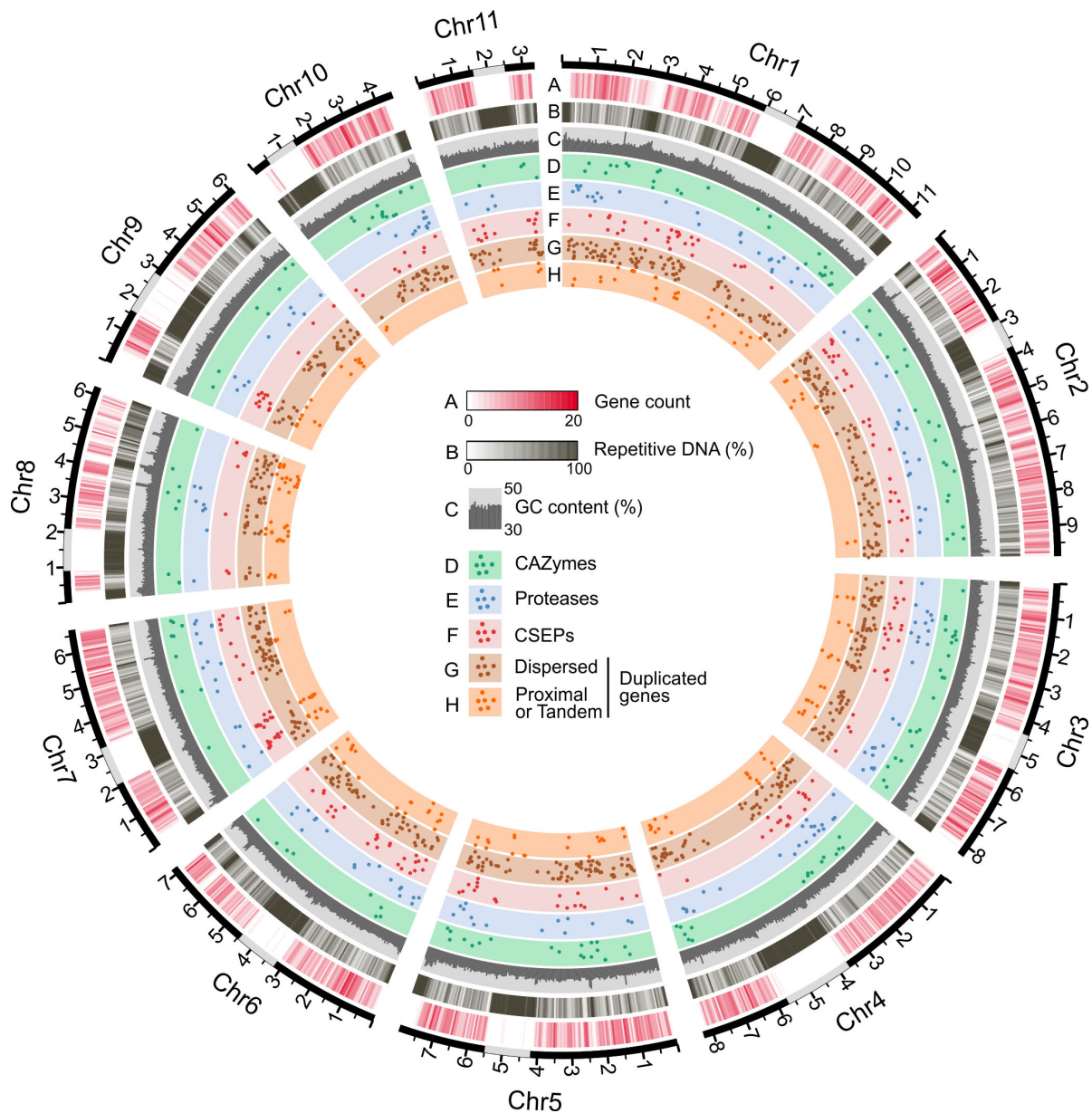
**FIG 1** Schematic representation of the 11 chromosomes of *E. necator* isolate EnFRAME01. The Circos plot shows the assembled chromosomes as solid black lines with major tick marks representing Mb. Predicted location of centromeric regions is indicated with gray rectangles. The outermost-to-innermost tracks represent (A) density of protein-coding genes, (B) repetitive DNA content, (C) GC content from 30% to 50%, (D) location of genes encoding carbohydrate-active enzymes, (E) location of genes encoding proteases, (F) location of genes encoding CSEPs, (G) location of dispersed gene duplicates (i.e., gene copies located in different chromosomes or separated by more than 10 genes), and (H) location of proximal or tandem gene duplicates (i.e., gene copies located less than 10 genes apart or next to each other). Gene locations are represented by bullet points on the perpendicular axis. Gene count, repetitive DNA, and GC content were determined using a sliding window of 50 kb. The figure shows that the chromosomes of EnFRAME01 contain long centromeric regions, which are abundant in repeats and nearly devoid of protein-coding genes.

centromeric regions of *E. necator* are poorly conserved in *B. graminis* f.sp. *tritici*. Moreover, although both PM species have 11 chromosomes, they exhibited an overall low synteny as no one-to-one chromosome match was observed between them (Fig. S4). These results indicate poor conservation of centromeric regions and extensive inter-chromosomal rearrangements between *E. necator* and *B. graminis* f.sp. *tritici*.

## A reduced gene complement underlies the obligate biotrophic lifestyle of

**TABLE 1** Size and content of the 11 chromosomes of *Erysiphe necator* isolate EnFRAME01

| Chromosome | Size (Mb) | GC (%) | Centromere size (Mb) | Predicted genes | Genes per Mb | Median intergenic size (bp) | Repeats (%) |
|---|---|---|---|---|---|---|---|
| Chr1 | 11.30 | 39.8 | 1.00 | 1,000 | 88 | 4,333 | 61.7 |
| Chr2 | 9.87 | 39.6 | 0.85 | 1,060 | 107 | 3,391 | 55.6 |
| Chr3 | 8.28 | 39.5 | 1.05 | 840 | 101 | 3,502 | 58.4 |
| Chr4 | 8.27 | 39.6 | 2.00 | 684 | 83 | 4,000 | 64.4 |
| Chr5 | 7.98 | 39.8 | 1.30 | 758 | 95 | 3,286 | 61.8 |
| Chr6 | 7.15 | 39.8 | 1.25 | 608 | 85 | 4,402 | 64.9 |
| Chr7 | 6.66 | 39.4 | 1.12 | 641 | 96 | 3,639 | 60.5 |
| Chr8 | 6.21 | 40.2 | 1.20 | 393 | 63 | 5,218 | 70.0 |
| Chr9 | 6.07 | 39.9 | 1.25 | 474 | 78 | 3,249 | 69.4 |
| Chr10 | 4.49 | 39.5 | 0.90 | 387 | 86 | 3,536 | 63.6 |
| Chr11 | 3.34 | 39.4 | 0.90 | 261 | 78 | 3,666 | 68.3 |

## *E. necator*

A total of 7,146 protein-coding genes were predicted in the genome of EnFRAME01, with a BUSCO completeness of 98.2%. This gene number is comparable with the 6,046–8,470 genes reported in other PM fungi (6, 8, 9, 67), and a notable improvement over the gene annotation of C-strain, for which 6,484 genes were predicted with an estimated completeness of 90.1% (Table S3). Functional gene annotations showed that the genome of EnFRAME01 contained 174 proteases (Table S4), 8 key enzymes for secondary metabolism (Table S5), 11 cytochrome P450s (Fig. S5; Table S6), 1,238 putative transporters (Fig. S6; Table S7), 160 CAZymes (Table S8), and 527 SPs (Table S9), of which 234 were candidate secreted effector proteins (CSEPs) (Table S10) (Supplementary Results). The number of genes in these functional categories is low compared with other plant pathogenic ascomycete fungi (8, 68) but similar to PMs (8, 15). Consistent also with an obligate biotrophic lifestyle, 181 core genes were identified that are typically present in *S. cerevisiae* and non-obligate biotrophic fungi but were missing in EnFRAME01 (Table S11). Included in these genes were 95 of the 99 so-called "missing ascomycete pathogen core genes," generally reported as absent in PMs (10) (Table S12). Based on KEGG orthology (KO) identifiers (69), the 181 genes are predicted to partake in 47 conserved pathways (Table S13), of which 23 were significantly enriched in genes missing in EnFRAME01 (Supplementary Results). The two pathways most affected by gene losses were thiamine and sulfur metabolism, in accordance to other obligate biotrophic fungi (10, 17, 68). Absence of a sterol O-acyltransferase (EC:2.3.1.26) gene, and of the *ERG5* (C-22 sterol desaturase; EC:1.14.19.41) and *ERG4* (EC:1.3.1.71) genes whose products catalyze the last two steps of ergosterol biosynthesis in yeast (70), was also observed. Collectively, these results indicate that the obligate lifestyle of *E. necator* is driven by losses in genes involved in several biochemical pathways, in accordance to what has been observed in other PMs and obligate biotrophs (10, 17).

## *E. necator* harbors a reduced arsenal of CSEPs

The small number of 234 CSEP-encoding genes identified in the genome of EnFRAME01 is in line with reports from dicot-infecting PMs but in contrast to monocot-infecting PMs such as different *B. graminis* formae speciales (6, 9). Of the 234 CSEPs, 49 (20.9%) were species specific, and 185 (79.1%) had homologs in PMs (*n* = 183) and/or non-PM fungi (*n* = 86) (Fig. S7). Moreover, 86 (36.7%) contained the Y/F/WxC sequence motif that is typically found in CSEPs of *B. graminis* and other PMs (Fig. S8 and Table S10). PM fungi are also known to harbor many ribonuclease-like effectors that belong to a large family of catalytically inactive RNases, known as RALPHs (RNase-like proteins associated with haustoria) (71–73). A genome-wide search in EnFRAME01 identified 38 genes encoding RALPH-like proteins, 24 of which could also be classified as CSEPs (Table S14). A phylogenetic analysis grouped the 38 RALPH-like proteins into two major clades, whose members differed in average protein size and the location in the genome of

their encoding genes (Fig. S9). An Egh16-like virulence factor domain (PF11327) was also commonly found in the *E. necator* CSEPs. CSEPs with an Egh16 domain are members of a multigene family in fungi (74, 75) and often play a role during the early stages of host infection. A total of 11 genes encoding Egh16-like proteins that could be further clustered into two clades (Fig. S10) were identified in EnFRAME01, but only four of these were classified as CSEPs (Table S15). Finally, an analysis of the localization of the 234 CSEP-encoding genes of *E. necator* on the 11 chromosomes of the fungus showed there was no enrichment of genes encoding CSEPs in subtelomeric regions, as has been observed in other fungi (76, 77) (Fig. 1).

## TE bursts have drastically shaped the genome of *E. necator*

The genome of EnFRAME01 is highly repetitive, with repeats accounting for 62.7% (50.8 Mb) of its DNA content. Class I retrotransposons, such as long terminal repeat (LTR) retrotransposons (26.5%, 21.5 Mb) and non-LTR retrotransposons (16.8%, 13.6 Mb), were more abundant than class II DNA transposons (6.3%, 5.1 Mb) and unclassified interspersed repeats (13.1%, 10.6 Mb) (Table S16). This is consistent with most fungi (78, 79) but in contrast to cereal PMs, whose genomes are mainly dominated by non-LTR retrotransposons (6, 9). TEs were fairly evenly dispersed outside centromeric and subtelomeric regions, which generally contained smaller amounts of non-LTR elements and exhibited an overall lower TE divergence. A similar pattern was also observed in genomic islands rich in rolling-circle (RC) elements (Fig. 2A). Collectively, these observations indicate that younger TEs accumulated preferentially in centromeric and subtelomeric regions and that RC elements are younger than other TEs (Fig. 2A). Interestingly, an examination of the nucleotide divergence among TE copies revealed a bimodal distribution with two peaks of contrasting TE composition. This suggests the presence of two TE burst events in the evolutionary history of *E. necator* that involved different TE classes (Fig. 2B and Supplementary Results). A similar pattern was also observed in the genomes of different *B. graminis* formae speciales, although the bimodal peaks were less pronounced and lacked RC elements (Fig. 2B; Table S16). In addition, highly divergent TEs were in all genomes enriched in non-LTR rather than LTR elements, whereas the opposite was observed for TEs with low divergence. By using the *E. necator* repeat library to mask the genomes of the cereal PMs, and vice versa, nearly all low-divergence TEs were left unmasked (Fig. S11). These observations suggest that *E. necator* and *B. graminis* underwent a similar burst of non-LTR TEs, possibly prior to their divergence, followed by clade-specific proliferation of LTR-retrotransposons and, in the case of *E. necator*, of RC elements as well.

## The genome of *E. necator* exhibits small-scale compartmentalization

An examination of the distribution of repeats and of protein coding genes on the 11 chromosomes of EnFRAME01 revealed large differences in gene density among the chromosomes and an inverse correlation between density of protein coding genes and repetitive DNA content (Fig. 1; Fig. S12 and Supplementary Results). An assessment of whether certain gene categories were associated with specific TE superfamilies showed no major differences in TE content within the flanking regions of genes encoding CAZymes, proteases, CSEPs, and non-CSEP secreted proteins (Fig. S13). However, the intergenic regions of CSEP genes were significantly longer, richer in repetitive DNA, and had a different TE composition as compared with other functional gene categories (Fig. 3A and B; Table S17). Collectively, these observations indicate the absence of large-scale compartmentalization in gene-dense and gene-sparse regions in the genome of *E. necator* (Fig. 3C and D), consistent with the "one-speed" genome hypothesis suggested for PM species (6, 20). Instead, small-scale compartmentalization of CSEP-encoding genes was seen, which was preferentially located in somewhat gene-sparse and repeat-rich regions, as commonly reported in other fungal pathogens (80, 81).
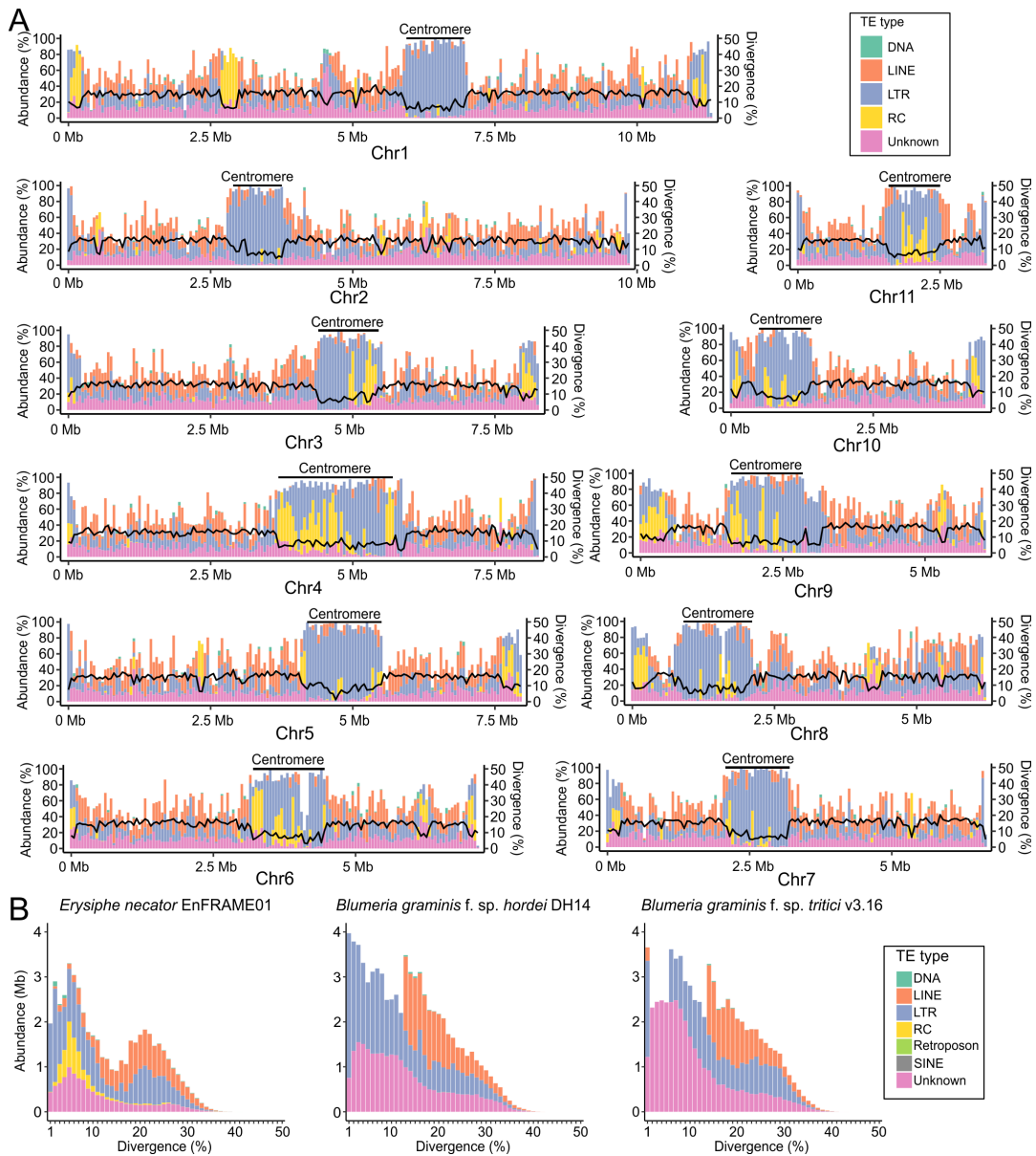
**FIG 2** The transposable element composition of *E. necator* differs from that of the cereal powdery mildew pathogens *B. graminis* f.sp. *hordei* and *B. graminis* f.sp. *tritici*. (A) Distribution of TEs in the 11 chromosomes of *E. necator* isolate EnFRAME01. The figure shows the abundance of the different TE classes, represented as stacked bar plots along the chromosomes. Overall divergence of TE families is indicated by solid black lines along the chromosomes. Predicted centromeric regions are indicated as well. The figure shows high abundance of repeats near chromosome ends and at centromeres. Predicted centromeric regions are enriched mainly in long terminal repeat retrotransposons with overall low sequence divergence compared with the rest of the genome. Rolling-circle elements are also abundantly found in centromeres and have an overall low sequence divergence. TE abundance and divergence were calculated using a sliding window of 50 kb. (B) Repetitive DNA landscape represented as bar plots showing the number of bases covered by predicted TEs from different (sub)classes. The predicted divergence of the TEs is shown on the *x*-axis. The figure shows a bimodal repeat divergence landscape with peaks for *E. necator* at approximately 5% and 21% divergence. The two peaks differ in their composition, with the peak at 5% divergence being dominated by LTRs, RCs, and unknown elements, and the peak at 21% divergence being dominated by LTR and long interspersed nuclear elements (LINE). The landscape of TE divergence of the cereal PM pathogens also follows a bimodal distribution, but it is less pronounced as compared with *E. necator,* and the peaks are void of RCs.

## Gene duplication asymmetrically affects different functional gene categories in *E. necator* and their genomic organization

A self-BLASTp search revealed a total of 941 genes (13.1%) duplicated in the genome of EnFRAME01, with CSEP-encoding genes experiencing significantly (*P*-value = 1.8E-35)
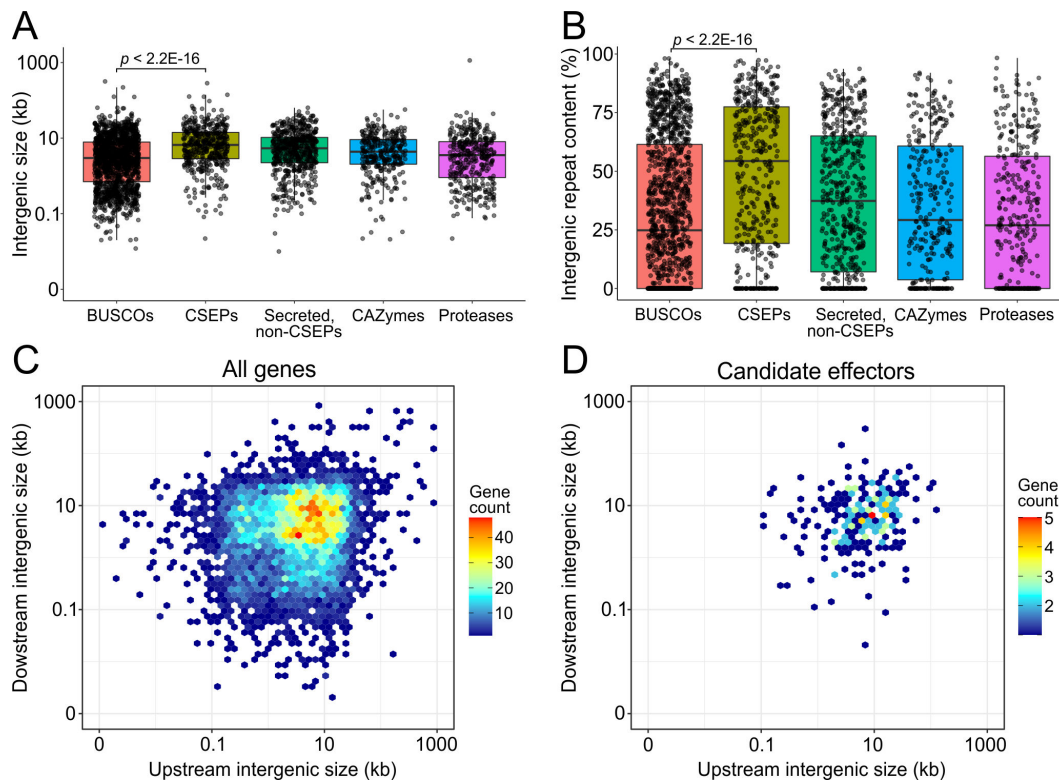
**FIG 3** The genome of *E. necator* isolate EnFRAME01 exhibits small-scale compartmentalization of genes encoding candidate secreted effector proteins in repeat-rich genomic regions. (A and B) Boxplots showing the size distribution and repetitive DNA content of upstream and downstream intergenic regions flanking BUSCO genes, genes encoding CSEPs, genes encoding secreted proteins not classified as CSEPs, genes encoding carbohydrate-active enzymes, and genes encoding proteases. The figure shows that intergenic regions of CSEP-encoding genes typically have higher repetitive DNA content compared with genes in the other categories. The *P*-values shown in panels A and B were obtained with the Wilcoxon rank sum test. (C and D) Heatmaps of the number of protein coding genes (panel C) and CSEPs (panel D) with certain sizes of upstream (*y*-axis) and downstream (*x*-axis) intergenic regions. The figure shows that the genome of EnFRAME01 does not exhibit large-scale compartmentalization of CSEP-encoding genes in gene-sparse regions.

higher rates of gene duplications, as compared with genes in other functional categories (Fig. 4A; Table S18; Supplementary Results). A conserved domain enrichment analysis further identified 30 domains that were significantly enriched among duplicated genes (adjusted *P*-value <0.01), with the two most significantly enriched being the microbial ribonuclease (cl00212) and the Egh16-like virulence factor (PF11327) domains that are associated with CSEPs as well (Fig. 4B; Table S19). When considering the arrangement of the 941 gene duplicates in the genome of *E. necator*, the majority were dispersed gene duplicates (DGDs; *n* = 712; 75.6%), as opposed to being proximal gene duplicates (PGDs; *n* = 139; 14.8%) or tandem gene duplications (TGDs; *n* = 90; 9.5%) (Fig. 4A). However, genes encoding CSEPs significantly deviated from this pattern as they exhibited almost equal frequencies of dispersed (*n* = 38; 35.5%), proximal (*n* = 34; 31.8%), and tandem (*n* = 35; 32.7%) duplications (Fig. 4A). Indeed, several of the multicopy CSEP-encoding genes, including the RALPH-like and Egh16-like CSEPs, were found to be tandemly arranged in clusters (Fig. S14), suggesting that CSEPs families expand by frequent local duplications in *E. necator*. A prominent example of this trend was the discovery of a 350-kb region on Chr1 that harbored 20 copies of a CSEP-encoding gene (i.e., *HI914_00480*), which were tandemly arranged in the same orientation on the same DNA strand (Fig. S15) and with 15 consecutive copies encoding identical proteins. Our analyses also suggested that local gene duplicates (i.e., PGDs and TGDs) are more conserved and thus more likely to contribute to genetic redundancy than DGDs, which due to their higher divergence, are likely to contribute more to functional diversification (Fig. 4C). Similarly, when examining the rate of synonymous ($K_S$) and non-synonymous substitutions ($K_A$), local gene
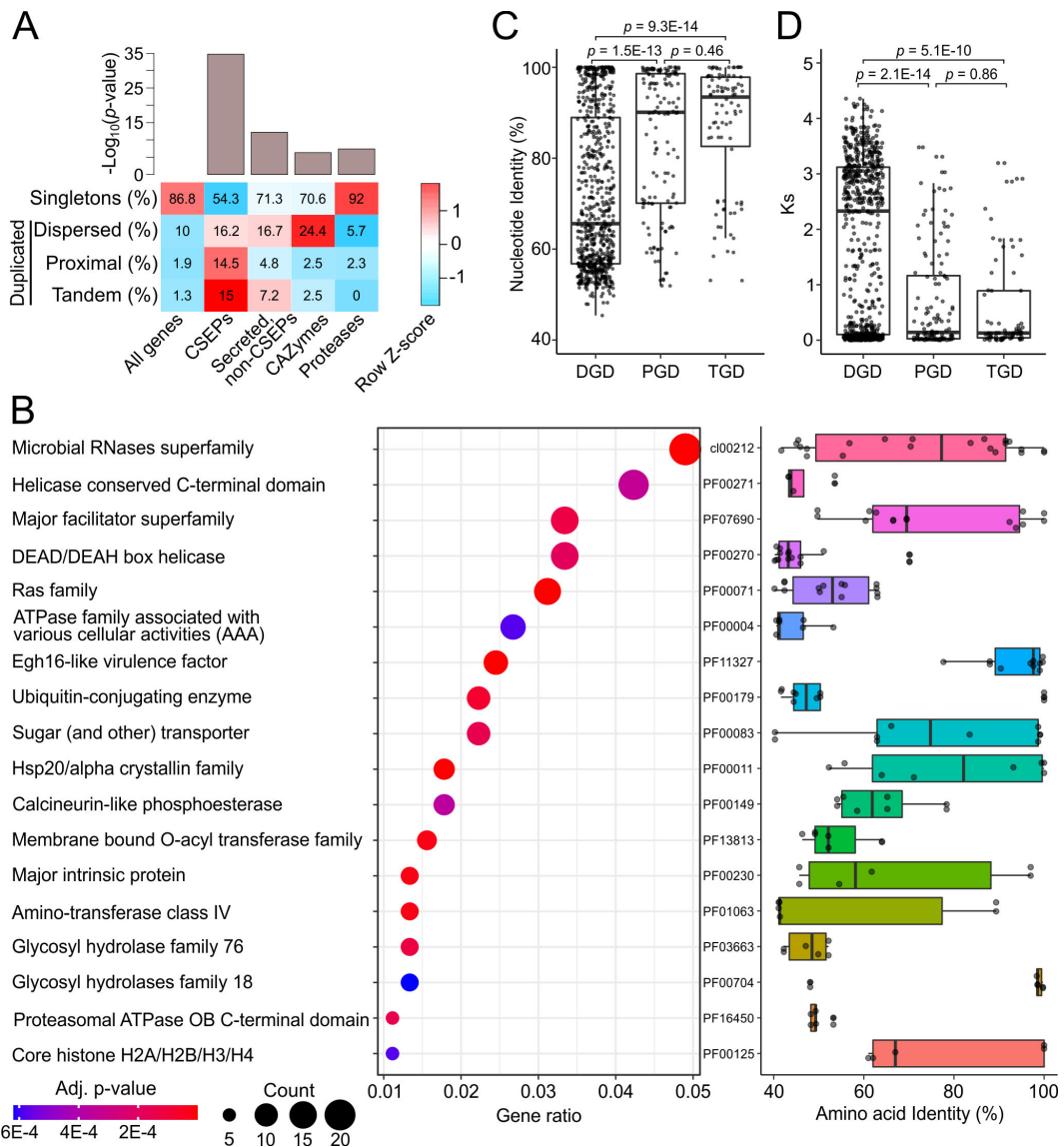
**FIG 4** Landscape of gene duplications in the genome of *E. necator* isolate EnFRAME01. (A) Heatmap showing the percentage of genes in different functional categories that are singletons, dispersed duplications, proximal duplications, and tandem duplications. The bar chart shows *P*-values for enrichment of duplicated genes based on hypergeometric tests. The figure shows that ~13% of the genes are duplicated and that the percentages of duplicated genes encoding candidate secreted effector proteins are significantly higher than genes encoding secreted proteins not classified as CSEPs, carbohydrate-active enzymes, and proteases. (B) The dot plot shows conserved domains significantly enriched within duplicated genes. The size of the dots corresponds to the number of duplicated genes containing the respective domain. The *x*-axis shows the proportion of the duplicated genes containing the respective domain that contributes to all duplicated genes containing a conserved domain. Dots are color coded based on enrichment *P*-values adjusted using the Benjamini–Hochberg method. Distributions of pairwise identity values of duplicated copies are shown on the right-hand side based on top BLASTp hit. The *P*-values shown in (C) and (D) were obtained with the Wilcoxon rank sum test. (C) Boxplots showing the distribution of pairwise nucleotide identity values of dispersed gene duplicates, proximal gene duplicates, and tandem gene duplicates. Each point represents a duplicated gene with the percent identity of its top BLASTn hit shown in the *y*-axis. The figure shows that copies of DGDs share significantly less nucleotide identity (median = 65.6%) than copies of PGDs (median = 90.1%) and TGDs (median = 93.4%). (D) Boxplots showing the distribution of $K_A/K_S$ values for DGDs, PGDs, and TGDs. Each point represents a duplicated gene with the $K_A/K_S$ value of its top BLASTp hit shown in the *y*-axis. The figure shows that copies of PGDs and TGDs share higher conservation of $K_A/K_S$ values than copies of DGDs.

duplicates exhibited overall lower $K_S$ and higher $K_A/K_S$ values as compared with DGDs (Fig. 4D; Fig. S16). This indicated that PGDs and TGDs were likely more recent duplicates and were under more relaxed selection pressure as compared with DGDs. Collectively, the above results indicate that gene duplication is a driver of genome evolution in

*E. necator* that has differentially affected different gene categories, thereby leading to differences in their mode of evolution and organization of their paralogs in the genome.

## Duplicated genes in EnFRAME01 frequently vary in copy number among *E. necator* isolates

The WGS data of five *E. necator* isolates (7) were used to identify genomic regions in EnFRAME01 with CNV (i.e., deleted or duplicated). A total of 1,760 distinct CNV regions were identified, of which 1,589 (90.3% with an average size of 2.9 kb) were deletions, and only 171 (9.7% with an average size of 5.6 kb) were duplications (Fig. S17A; Table S20). CNV regions were dispersed throughout the 11 chromosomes of EnFRAME01 (Fig. S18) and were more frequently located near chromosome ends rather than gene-rich and repeat-rich regions (Fig. S17B). However, despite the lack of enrichment of CNV regions in repeat-rich regions, 80.5% ($n = 1,279$) of the deleted and 70.1% ($n = 120$) of the duplicated regions overlapped with predicted TEs (Fig. S17C). Moreover, 122 of the CNV regions overlapped with protein-coding genes and could, therefore, be considered as CNV genes (Supplementary Results). Of these, 53 genes were duplicated (average of 0.3 duplicated gene per all duplicated regions), and 69 genes were deleted (average of 1.5E-5 deleted gene per all deleted regions) among the *E. necator* isolates (Table S21), indicating that genes with CNV were most likely to be affected by duplications rather than deletions. Most CNV regions also typically affected single genes rather than groups of genes, and no significant over- or under-representation of CSEPs, CAZymes, and proteases was observed among CNV genes. Instead, genes with CNV were significantly (*P*-value = 1.7E-27) enriched with the 941 genes predicted to be duplicated in EnFRAME01 (Table S22). A notable example is the CSEP-encoding gene *HI914_00480*, which is present in 20 copies in EnFRAME01 and 8–12 copies in other strains (Table S23). This indicated that rates of gain, retention, and loss of duplicated genes were asymmetric among different isolates of *E. necator*.

### *E. necator* exhibits extensive CNV of a novel and PM-specific CE

An inspection of the 122 genes with CNVs among the isolates of *E. necator* showed that gene *HI914_00624,* encoding a predicted secreted CE, exhibited the most dynamic changes in copy numbers, ranging from 1 in isolate EnFRAME01 to 31 in isolate Lodi (Supplementary Results). Moreover, RT-qPCR assays showed that the relative expression of this gene in six 2-week-old isolates of *E. necator* was strongly correlated (Pearson's linear correlation coefficient $r^2 = 0.998$, $P < 0.001$) to its copy number (Fig S19 and Table S24), indicating that increases in copy number resulted in a gene dosage effect. In all isolates, the duplication affected the same 9.5 kb fragment that contained only the *HI914_00624* gene and was flanked by short direct repeats (Fig. 5A). A blast search within the NCBI nr database indicated that homologs of HI914_00624 are abundantly present both within PM and non-PM fungal species (Fig. 5B). However, a phylogenetic tree constructed using the top 400 best BLASTp hits, representing at least 195 distinct fungal species, showed that HI914_00624 belonged to a distinct clade that included 22 CEs, all from PM species (Fig. 5C). Moreover, a multiple sequence alignment showed that the catalytic triad Ser-Asp/Glu-His, that is indispensable to the function of CEs (82, 83), is poorly conserved in these 22 PM-specific CEs (Fig. S20; Table S25). This suggests that HI914_00624 is a member of new clade of potentially non-catalytically active CEs or CEs with a modified enzymatic activity (84).

## DISCUSSION

In this study, we used deep WGS sequencing to obtain a chromosome-scale genome for the grape PM *E. necator*, the first Erysiphales genome for a dicot-infecting PM. The availability of a high-quality assembly enabled us to uncover prominent genomic and biological features of *E. necator* that were previously left unexplored due to its fragmented assembly. This highlights the importance of obtaining high-quality
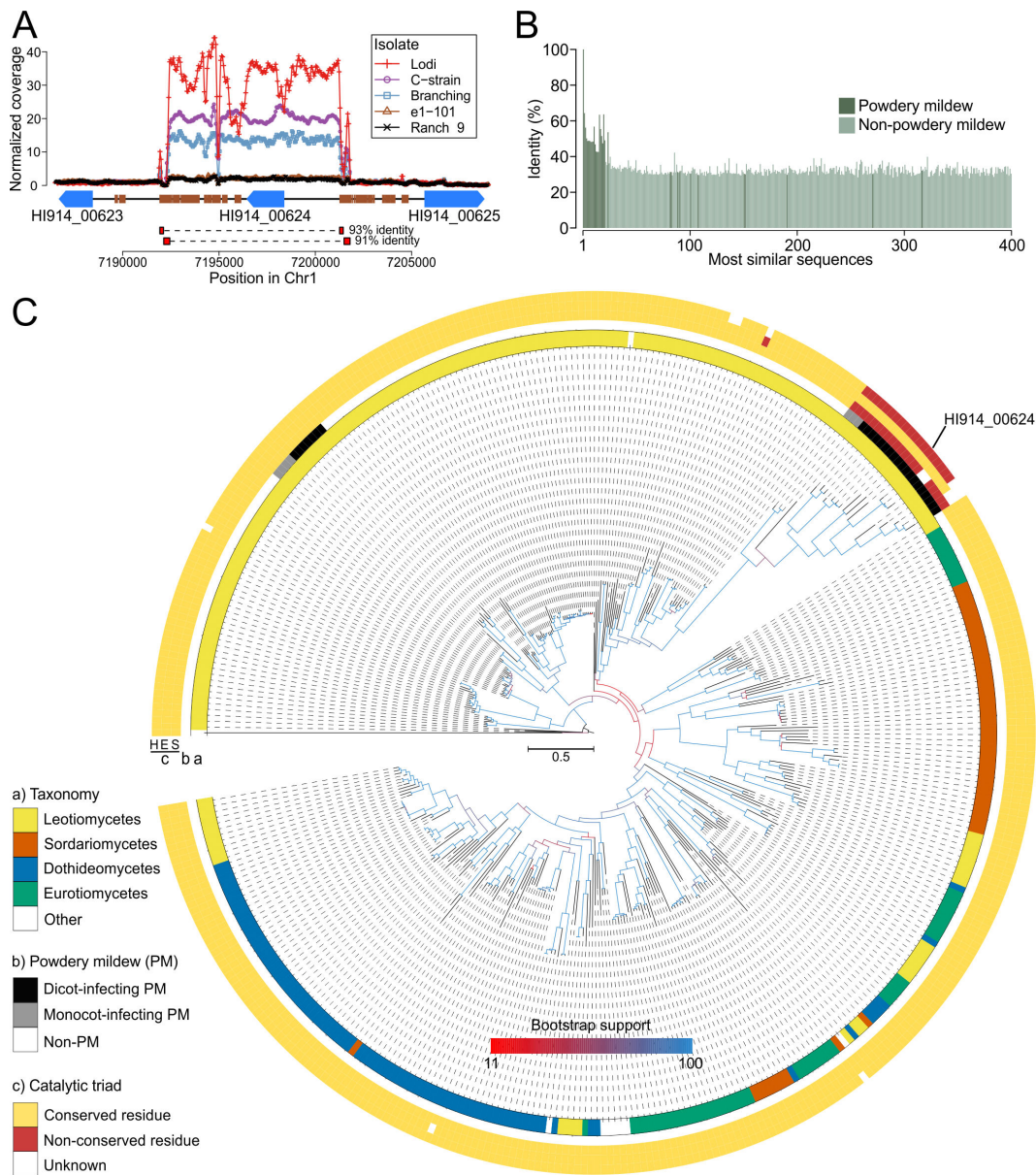
**FIG 5** *E. necator* shows extensive copy number variation of a putative secreted carboxylesterase that is poorly conserved in non-powdery mildew fungi. (A) Region of chromosome 1 (Chr1) in the genome of *E. necator* isolate EnFRAME01 containing the gene *HI914_00624* encoding a putative secreted CE. Genes are represented as blue arrows and repetitive DNA as small brown rectangles. Lines above the genes indicate estimated copy numbers of the region in five different isolates. The figure shows that isolates Lodi, C-strain, and Branching have more than 10 predicted copies of *HI914_00624*. The duplicated segment is flanked by short direct repeats of more than 90% identity. The figure also shows that genes flanking *HI914_00624* are not duplicated in the isolates analyzed. (B) Percent identity values of most similar sequences to the HI914_00624 protein sequence based on BLASTp searches. The figure shows that nearly all sequences from non-PM species have less than 40% amino acid identity. (C) Maximum likelihood phylogenetic tree of HI914_00624 and its most similar protein sequences from GenBank (2022, 08-13) and EnFRAME01. The protein sequence of the acetylcholinesterase DmAChE from *D. melanogaster* (54) was included as an outgroup. Tree branches are color coded based on their support of 1,000 bootstrap replicates. The tree was rooted at DmAChE. Track (A) shows the distribution of taxonomy classes of the sequences. Track (B) indicates sequences from monocot-infecting and dicot-infecting PMs. Track (C) shows the conservation of the Ser, Asp/Glu, and His residues that comprise the catalytic triad conserved in CEs. The figure shows that homologs of HI914_00624 are conserved in other PMs but poorly conserved in non-PM fungal species. The figure also shows that homologs of HI914_00624 in PMs lack the Ser and His residues of the catalytic triad, which are largely conserved in predicted CEs from other fungal species.

chromosome-level assemblies and further sets a solid basis for a detailed structural genomics studies in this pathogen.

The 81.1 Mb genome of *E. necator* is organized into 11 chromosomes, which are broadly characterized by the presence of large centromeric-like regions rich in repetitive DNA, a high content of retrotransposons and unclassified repeats that are mostly evenly dispersed outside their centromeric and telomeric regions, and the lack of compartmentalization in repeat-rich/gene-sparse regions, in agreement with the "one-speed genome" model of evolution. Moreover, *E. necator* had a reduced complement of genes encoding lytic enzymes (e.g., CAZymes and proteases) and those involved in carbohydrate metabolism, amino acid and purine metabolism, thiamine biosynthesis, and assimilation of inorganic nitrogen and sulfur. Loss of genes affecting these pathways is a characteristic feature of obligate biotrophs (8, 17, 68, 85, 86) and potentially a strategy used by such pathogens for conserving their resources, if the end products of the impaired pathways are available through leaky metabolic processes in the host (17, 87–89). This model of reductive genome evolution in which organisms abolish genes needed to synthesize metabolites that can be obtained directly through the host environment is frequently observed in nature, including obligate biotrophic fungi (68, 87, 88, 90).

The search for genes that are missing in *E. necator* but are commonly present in other ascomycete fungi also revealed that the fungus lacks *ERG5* and *ERG4*, whose products catalyze the two last steps, respectively, of ergosterol biosynthesis in fungi (70). This corroborates previous reports that ergosterol is essentially absent in this fungus as well as in PMs, in general (91, 92). The deletion of *ERG5* and/or *ERG4* is typically not lethal to fungi but leads to an altered ergosterol biosynthesis that may affect their physiology, increase their sensitivity to multiple chemicals, and generally decrease their fitness under stress conditions (93–97). For instance, both genes are required for the conidiation of *A. fumigatus* (98, 99), *ERG4* is crucial for vegetative differentiation and virulence in *Fusarium graminearum* (100), and deletion of *ERG4* increased the production of extracellular pigments in *Monascus purpureus* (101). Likewise, the deletion of *ERG5* increased the susceptibility of *Candida albicans, N. crassa,* and *F. verticillioides* to azole antifungals (102, 103). Loss of *ERG4* and *ERG5* is unlikely to have a fitness effect on *E. necator* or impact major features of its physiology, and it is thus intriguing to speculate possible biological explanations for the loss of ergosterol biosynthesis. As ergosterol is an inducer of innate immunity in plants (104), its absence may help *E. necator* avoid sterol-induced immunity in grapes. This might be possible as the activation in *V. vinifera* of the type I lipid transfer protein VvLTP1 by ergosterol treatment (105, 106) leads to the induction of the stilbene synthase gene *Vst1* (105). *Vst1*, in turn, regulates the biosynthesis of the phytoalexin resveratrol that enhances resistance against *Botrytis cinerea* (107) and *E. necator* (108).

The overall genomic characteristics of *E. necator* conform to those reported for other PM pathogens, including different formae speciales of the cereal PM pathogen *B. graminis* (8, 10, 14). However, the 11 chromosomes of *E. necator* are not syntenic to the 11 chromosomes of *B. graminis*, indicating rapid diversification of their genomic architecture following speciation. Moreover, the genome of *E. necator* has a different TE complement compared with *B. graminis*, as it contains mostly LTR retrotransposons rather than the non-LTR retrotransposons. TEs are a major force of evolution and adaptation to stressful environments, as their bursts and mobilization provoke chromosomal reorganization and phylogenetic divergence (109, 110). Our analysis showed that both species have experienced not one, as previously reported (6), but at least two bursts of TEs in their evolutionary history. The first burst possibly preceded their divergence and involved mostly non-LTR retrotransposons, and the second burst likely took place after their speciation and involved LTR-retrotransposons but also RCs (i.e., Helitrons) in *E. necator*. Such differences in TE bursts in *E. necator* and *B. graminis* are likely to have restructured their genomes, accelerated their speciation, and influenced their adaptation on different hosts by, among others, affecting virulence-associated genes such as CSEPs. Indeed, despite the lack of large-scale compartmentalization in *E. necator*, its CSEP-encoding genes exhibited significantly higher duplication rates as compared with other functional gene categories and were embedded in larger intergenic regions

that were richer in TEs. The increase in duplication rates could have been prompted by the presence of TEs, as the repetitive nature of transposons provides a substrate for non-allelic homologous recombination that would typically generate tandemly arranged gene copies in their flanking regions (111–113). TEs have also been hypothesized to mediate the duplication and proliferation of CSEPs in *B. graminis* (9, 73), as CSEP-encoding genes in this species are frequently duplicated and present in tandem in physical proximity to similar repetitive DNA (9, 73). Thus, next to promoting chromosomal reorganization, TEs seem to have had a major role in shaping the evolution of *E. necator* and *B. graminis* as plant pathogens by providing a favorable environment for CSEP duplication.

The inflation of the *E. necator* genome by TEs was further accompanied by high rates of gene duplication, which likely contributed further to its genomic plasticity and genetic diversity. Gene duplication is a major force of evolution as it provides material for functional, regulatory, and transcriptional divergence through the generation of new genes and their subsequent neo-, sub-, or hypo-functionalization (114, 115). A variety of mechanisms can trigger gene duplications, with different mechanisms creating suites of duplicated genes in different configurations within a genome, which in turn contribute differentially to functional innovation and redundancy (116, 117). Our analysis indicated that in *E. necator,* genes from different functional categories exhibited different rates and modes of duplications and that different modes of gene duplication were under different strengths of selection pressure. These features were again more prominent with CSEP-encoding genes, as CSEP gene duplicates were more likely to be in close (i.e., tandem or proximal) physical location in the genome of EnFRAME01 than the copies of other gene classes and had on average higher $K_A/K_S$ values, indicating that duplicated CSEP genes are potentially subject to higher rates of evolution. This is consistent with the role of effectors on host adaptation and overcoming of the host immune system and indicates an ongoing arms race between *E. necator* and its grapevine host (9).

Next to gene duplications, CNVs within a species population can significantly affect its fitness (118). It has been shown, for example, that an increase in *CYP51* (*ERG11*) copy numbers, the gene encoding a key enzyme for ergosterol biosynthesis, creates a gene dosage effect that reduces the sensitivity of *E. necator* to demethylase inhibitor fungicides (7). In *B. graminis* f.sp. *hordei* (73, 119) and *B. graminis* f.sp. *tritici* (9), high levels of CNV in genes encoding CSEPs are thought to be major drivers of virulence and rapid adaptation to host genotypes. Our CNV analysis revealed that the CE-encoding gene *HI914_00624* exhibited the most dynamic changes in copy numbers, suggesting that it is a target of natural selection. Moreover, we found that *HI914_00624* is a member of a novel family of CE-encoding genes with multiple duplications in PM species. CEs are a large superfamily of structurally diverse, multifunctional enzymes that hydrolyze carboxylesters in natural and synthetic molecules (83), including pharmaceutical drugs, pesticides, environmental pollutants, and toxins. Due to their catalytic flexibility, they may have crucial roles not only in detoxifying cells from harmful compounds and metabolites but also in physiological processes such as lipid metabolism and energy homeostasis (120). We speculate that the putative CE encoded by *HI914_00624*, and its homologs in PM species, represents a new family of non-catalytic CEs, as they were poorly conserved in non-PM fungi and lack the conserved Ser-Asp/Glu-His amino-acid triad required for their proper function (83). Catalytic competence is thought to be the ancestral state of CEs, but several non-catalytic clades that have acquired new functions are present in higher eukaryotes (83, 121). Among fungi, *vdtD* from the opportunistic human pathogen *Paecilomyces variotii* encodes a putative non-catalytic CE that is part of a gene cluster mediating the biosynthesis of the antibacterial viriditoxin (122). It has been suggested that instead of acting as a hydrolase, vdtD could bind to the compound to protect the methyl ester from being hydrolyzed by endogenous hydrolases (123). These examples highlight the capability of CEs to evolve new functions, and it is, therefore, possible that the putative CE encoded by *HI914_00624* and its homologs in PMs has evolved new functions compared with ancestral CEs.

## Supplementary materials

All supplementary materials are available through Zenodo at https://doi.org/10.5281/zenodo.7738565 and include Supplementary Results, Supplementary Materials and Methods, Fig. S1 to S21, and Supplementary Tables S1 to S26.

## AUTHOR AFFILIATIONS

[1]Department of Plant Pathology, University of California Davis, Davis, California, USA
[2]USDA-ARS, Horticultural Crops Disease and Pest Management Research Unit, Corvallis, Oregon, USA

## PRESENT ADDRESS

Jacob Corcoran, USDA-ARS, Biological Control of Insects Research Laboratory, Columbia, Missouri, USA

## AUTHOR ORCIDs

Ioannis Stergiopoulos  http://orcid.org/0000-0002-2368-6119

## AUTHOR CONTRIBUTIONS

Alex Z. Zaccaron, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Tara Neill, Conceptualization, Formal analysis, Methodology, Resources, Writing – review and editing | Jacob Corcoran, Resources, Writing – review and editing | Walter F. Mahaffee, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing | Ioannis Stergiopoulos,

Conceptualization, Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review and editing

## DATA AVAILABILITY STATEMENT

Raw sequencing reads generated in this study were deposited at NCBI SRA under accessions SRR18712274 through SRR18712279 (BioProject PRJNA627990). The annotated genome of EnFRAME01 was deposited at NCBI under accession JABETL000000000.1.

## REFERENCES

1. Gadoury DM, Cadle-Davidson L, Wilcox WF, Dry IB, Seem RC, Milgroom MG. 2012. Grapevine powdery mildew (*Erysiphe necator*): a fascinating system for the study of the biology, ecology and epidemiology of an obligate biotroph. Mol Plant Pathol 13:1–16. https://doi.org/10.1111/j.1364-3703.2011.00728.x
2. Gaforio L, Garcia-Munoz S, Cabello F, Munoz-Organero G. 2011. Evaluation of susceptibility to powdery mildew (*Erysiphe necator*) in *Vitis vinifera* varieties. Vitis 50:123–126. https://doi.org/10.5073/vitis.2011.50.123-126
3. Qiu W, Feechan A, Dry I. 2015. Current understanding of grapevine defense mechanisms against the biotrophic fungus (*Erysiphe necator*), the causal agent of powdery mildew disease. Hortic Res 2:15020. https://doi.org/10.1038/hortres.2015.20
4. Fuller KB, Alston JM, Sambucci OS. 2014. The value of powdery mildew resistance in grapes: evidence from California. Wine Econ Policy 3:90–107. https://doi.org/10.1016/j.wep.2014.09.001
5. Kunova A, Pizzatti C, Saracchi M, Pasquali M, Cortesi P. 2021. Grapevine powdery mildew: fungicides for its management and advances in molecular detection of markers associated with resistance. Microorganisms 9:1541. https://doi.org/10.3390/microorganisms9071541
6. Frantzeskakis L, Kracher B, Kusch S, Yoshikawa-Maekawa M, Bauer S, Pedersen C, Spanu PD, Maekawa T, Schulze-Lefert P, Panstruga R. 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 19:381. https://doi.org/10.1186/s12864-018-4750-6
7. Jones L, Riaz S, Morales-Cruz A, Amrine KCH, McGuire B, Gubler WD, Walker MA, Cantu D. 2014. Adaptive genomic structural variation in the grape powdery mildew pathogen, *Erysiphe necator*. BMC Genomics 15:1081. https://doi.org/10.1186/1471-2164-15-1081
8. Liang P, Liu S, Xu F, Jiang S, Yan J, He Q, Liu W, Lin C, Zheng F, Wang X, Miao W. 2018. Powdery mildews are characterized by contracted carbohydrate metabolism and diverse effectors to adapt to obligate biotrophic lifestyle. Front. Microbiol 9:3160. https://doi.org/10.3389/fmicb.2018.03160
9. Müller MC, Praz CR, Sotiropoulos AG, Menardo F, Kunz L, Schudel S, Oberhänsli S, Poretti M, Wehrli A, Bourras S, Keller B, Wicker T. 2019. A chromosome-scale genome assembly reveals a highly dynamic effector repertoire of wheat powdery mildew. New Phytol 221:2176–2189. https://doi.org/10.1111/nph.15529
10. Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, Loren van Themaat EV, Brown JKM, Butcher SA, Gurr SJ, Lebrun M-H, Ridout CJ, Schulze-Lefert P, Talbot NJ, Ahmadinejad N, Ametz C, Barton GR, Benjdia M, Bidzinski P, Bindschedler LV, Both M, Brewer MT, Cadle-Davidson L, Cadle-Davidson MM, Collemare J, Cramer R, Frenkel O, Godfrey D, Harriman J, Hoede C, King BC, Klages S, Kleemann J, Knoll D, Koti PS, Kreplak J, López-Ruiz FJ, Lu X, Maekawa T, Mahanil S, Micali C, Milgroom MG, Montana G, Noir S, O'Connell RJ, Oberhaensli S, Parlange F, Pedersen C, Quesneville H, Reinhardt R, Rott M, Sacristán S, Schmidt SM, Schön M, Skamnioti P, Sommer H, Stephens A, Takahara H, Thordal-Christensen H, Vigouroux M, Weßling R, Wicker T, Panstruga R. 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330:1543–1546. https://doi.org/10.1126/science.1194573
11. Wu Y, Ma X, Pan Z, Kale SD, Song Y, King H, Zhang Q, Presley C, Deng X, Wei C-I, Xiao S. 2018. Comparative genome analyses reveal sequence features reflecting distinct modes of host-adaptation between dicot and monocot powdery mildew. BMC Genomics 19:705. https://doi.org/10.1186/s12864-018-5069-z
12. Zaccaron AZ, Stergiopoulos I. 2021. Characterization of the mitochondrial genomes of three powdery mildew pathogens reveals remarkable variation in size and nucleotide composition. Microb Genom 7:000720. https://doi.org/10.1099/mgen.0.000720
13. Müller MC, Kunz L, Graf J, Schudel S, Keller B. 2021. Host adaptation through hybridization: genome analysis of triticale powdery mildew reveals unique combination of lineage-specific effectors. Mol Plant Microbe Interact 34:1350–1357. https://doi.org/10.1094/MPMI-05-21-0111-SC
14. Bindschedler LV, Panstruga R, Spanu PD. 2016. Mildew-omics: how global analyses aid the understanding of life and evolution of powdery mildews. Front Plant Sci 7:123. https://doi.org/10.3389/fpls.2016.00123
15. Hacquard S. 2014. The genomics of powdery mildew fungi: past achievements, present status and future prospects. Adv Bot Res 70:109–142. https://doi.org/10.1016/B978-0-12-397940-7.00004-5
16. Aylward J, Steenkamp ET, Dreyer LL, Roets F, Wingfield BD, Wingfield MJ. 2017. A plant pathology perspective of fungal genome sequencing. IMA Fungus 8:1–15. https://doi.org/10.5598/imafungus.2017.08.01.01
17. Spanu PD. 2012. The genomics of obligate (and nonobligate) biotrophs. Annu Rev Phytopathol 50:91–109. https://doi.org/10.1146/annurev-phyto-081211-173024
18. Irelan JT, Selker EU. 1996. Gene silencing in filamentous fungi: RIP, MIP and quelling. J Genet 75:313–324. https://doi.org/10.1007/BF02966311
19. Selker EU. 1990. Premeiotic instability of repeated sequences in *Neurospora crassa*. Annu Rev Genet 24:579–613. https://doi.org/10.1146/annurev.ge.24.120190.003051
20. Frantzeskakis L, Kusch S, Panstruga R. 2019. The need for speed: compartmentalized genome evolution in filamentous phytopathogens. Mol Plant Pathol 20:3–7. https://doi.org/10.1111/mpp.12738
21. Miles TD, Neill TM, Colle M, Warneke B, Robinson G, Stergiopoulos I, Mahaffee WF. 2021. Allele-Specific detection methods for qoi fungicide-resistant *Erysiphe necator* in vineyards. Plant Dis 105:175–182. https://doi.org/10.1094/PDIS-11-19-2395-RE
22. Hoagland DR, Arnon DI. 1950. In The water-culture method for growing plants without soil, 2nd ed. circular 347. University of California, Berkeley, CA.
23. Feehan JM, Scheibel KE, Bourras S, Underwood W, Keller B, Somerville SC. 2017. Purification of high molecular weight genomic DNA from powdery mildew for long-read sequencing. J Vis Exp. https://doi.org/10.3791/55463
24. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116
25. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963
26. Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324
27. Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP, Aiden EL. 2017. De novo

assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science 356:92–95. https://doi.org/10.1126/science.aal3327

28. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst 3:99–101. https://doi.org/10.1016/j.cels.2015.07.012

29. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A 117:9451–9457. https://doi.org/10.1073/pnas.1921046117

30. Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12:357–360. https://doi.org/10.1038/nmeth.3317

31. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33:290–295. https://doi.org/10.1038/nbt.3122

32. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q others. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-seq data. Nat Biotechnol 29:644. https://doi.org/10.1038/nbt.1883

33. Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. 2008. Maker: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res 18:188–196. https://doi.org/10.1101/gr.6743907

34. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res 26:1107–1115. https://doi.org/10.1093/nar/26.4.1107

35. Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics 5:59. https://doi.org/10.1186/1471-2105-5-59

36. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: *ab initio* prediction of alternative transcripts. Nucleic Acids Res 34:W435–W439. https://doi.org/10.1093/nar/gkl200

37. Keilwagen J, Hartung F, Grau J. 2019. GeMoMa: homology-based gene prediction utilizing intron position conservation and RNA-Seq data. Methods Mol Biol Clifton NJ 1962:161–177. https://doi.org/10.1007/978-1-4939-9173-0

38. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240. https://doi.org/10.1093/bioinformatics/btu031

39. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res 45:D200–D203. https://doi.org/10.1093/nar/gkw1129

40. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. DbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 46:W95–W101. https://doi.org/10.1093/nar/gky418

41. Rawlings ND, Waller M, Barrett AJ, Bateman A. 2014. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res 42:D503–D509. https://doi.org/10.1093/nar/gkt953

42. Saier MH, Reddy VS, Tamang DG, Västermark A. 2014. The transporter classification database. Nucleic Acids Res 42:D251–D258. https://doi.org/10.1093/nar/gkt1097

43. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37:420–423. https://doi.org/10.1038/s41587-019-0036-z

44. Pierleoni A, Martelli PL, Casadio R. 2008. PredGPI: a GPI-anchor predictor. BMC Bioinformatics 9:392. https://doi.org/10.1186/1471-2105-9-392

45. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580. https://doi.org/10.1006/jmbi.2000.4315

46. Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB, Manners JM, Taylor JM. 2016. EffectorP: predicting fungal effector proteins from secretomes using machine learning. New Phytol 210:743–761. https://doi.org/10.1111/nph.13794

47. Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol 16:157. https://doi.org/10.1186/s13059-015-0721-2

48. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee T, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40:e49. https://doi.org/10.1093/nar/gkr1293

49. Zhang Z. 2022. KaKs_Calculator 3.0: calculating selective pressure on coding and non-coding sequences. Genomics Proteomics Bioinformatics 20:536–540. https://doi.org/10.1016/j.gpb.2021.12.002

50. Yu G, Wang L-G, Han Y, He Q-Y. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics J Integr Biol 16:284–287. https://doi.org/10.1089/omi.2011.0118

51. Faust GG, Hall IM. 2014. SAMBLASTER: fast duplicate marking and structural variant read extraction. Bioinformatics 30:2503–2505. https://doi.org/10.1093/bioinformatics/btu314

52. 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. https://doi.org/10.1093/bioinformatics/btp352

53. Abyzov A, Urban AE, Snyder M, Gerstein M. 2011. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res 21:974–984. https://doi.org/10.1101/gr.114876.110

54. Harel M, Kryger G, Rosenberry TL, Mallender WD, Lewis T, Fletcher RJ, Guss JM, Silman I, Sussman JL. 2000. Three-Dimensional structures of *Drosophila melanogaster* acetylcholinesterase and of its complexes with two potent inhibitors. Protein Sci 9:1063–1072. https://doi.org/10.1110/ps.9.6.1063

55. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066. https://doi.org/10.1093/nar/gkf436

56. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973. https://doi.org/10.1093/bioinformatics/btp348

57. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300

58. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https://doi.org/10.1038/nmeth.4285

59. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518–522. https://doi.org/10.1093/molbev/msx281

60. Letunic I, Bork P. 2021. Interactive tree of life (itol) V5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:W293–W296. https://doi.org/10.1093/nar/gkab301

61. Seidl MF, Kramer HM, Cook DE, Fiorin GL, van den Berg GCM, Faino L, Thomma BPHJ. 2020. Repetitive elements contribute to the diversity and evolution of centromeres in the fungal genus *Verticillium*. mBio 11:e01714-20. https://doi.org/10.1128/mBio.01714-20

62. Winter DJ, Ganley ARD, Young CA, Liachko I, Schardl CL, Dupont P-Y, Berry D, Ram A, Scott B, Cox MP. 2018. Repeat elements Organise 3D genome structure and mediate transcription in the filamentous fungus *Epichloë festucae*. PLoS Genet 14:e1007467. https://doi.org/10.1371/journal.pgen.1007467

63. Smith KM, Phatale PA, Sullivan CM, Pomraning KR, Freitag M. 2011. Heterochromatin is required for normal distribution of *Neurospora crassa* CenH3. Mol Cell Biol 31:2528–2542. https://doi.org/10.1128/MCB.01285-10

64. Schotanus K, Soyer JL, Connolly LR, Grandaubert J, Happel P, Smith KM, Freitag M, Stukenbrock EH. 2015. Histone modifications rather than the novel regional centromeres of *zymoseptoria tritici* distinguish core and

accessory chromosomes. Epigenetics Chromatin 8:41. https://doi.org/10.1186/s13072-015-0033-5

65. Yadav V, Yang F, Reza MdH, Liu S, Valent B, Sanyal K, Naqvi NI, Idnurm A. 2019. Cellular dynamics and genomic identity of centromeres in cereal blast fungus. mBio 10:e01581–19. https://doi.org/10.1128/mBio.01581-19

66. King R, Urban M, Hammond-Kosack MCU, Hassani-Pak K, Hammond-Kosack KE. 2015. The completed genome sequence of the pathogenic ascomycete fungus *Fusarium graminearum*. BMC Genomics 16:544. https://doi.org/10.1186/s12864-015-1756-1

67. Frantzeskakis L, Németh MZ, Barsoum M, Kusch S, Kiss L, Takamatsu S, Panstruga R. 2019. The *Parauncinula polyspora* draft genome provides insights into patterns of gene erosion and genome expansion in powdery mildew fungi. mBio 10:e01692-19. https://doi.org/10.1128/mBio.01692-19

68. Cissé OH, Pagni M, Hauser PM. 2014. Comparative genomics suggests that the human pathogenic fungus *Pneumocystis jirovecii* acquired obligate biotrophy through gene loss. Genome Biol Evol 6:1938–1948. https://doi.org/10.1093/gbe/evu155

69. Kanehisa M, Sato Y, Kawashima M. 2022. KEGG mapping tools for uncovering hidden features in biological data. Protein Sci 31:47–53. https://doi.org/10.1002/pro.4172

70. Hu Z, He B, Ma L, Sun Y, Niu Y, Zeng B. 2017. Recent advances in ergosterol biosynthesis and regulation mechanisms in *Saccharomyces cerevisiae* Indian J Microbiol 57:270–277. https://doi.org/10.1007/s12088-017-0657-1

71. Spanu PD. 2017. Cereal immunity against powdery mildews targets RNase-like proteins associated with haustoria (RALPH) effectors evolved from a common ancestral gene. New Phytol 213:969–971. https://doi.org/10.1111/nph.14386

72. Pennington HG, Jones R, Kwon S, Bonciani G, Thieron H, Chandler T, Luong P, Morgan SN, Przydacz M, Bozkurt T, Bowden S, Craze M, Wallington EJ, Garnett J, Kwaaitaal M, Panstruga R, Cota E, Spanu PD. 2019. The fungal ribonuclease-like effector protein CSEP0064/BEC1054 represses plant immunity and interferes with degradation of host ribosomal RNA. PLoS Pathog 15:e1007620. https://doi.org/10.1371/journal.ppat.1007620

73. Pedersen C, Ver Loren van Themaat E, McGuffin LJ, Abbott JC, Burgis TA, Barton G, Bindschedler LV, Lu X, Maekawa T, Wessling R, Cramer R, Thordal-Christensen H, Panstruga R, Spanu PD. 2012. Structure and evolution of barley powdery mildew effector candidates. BMC Genomics 13:694. https://doi.org/10.1186/1471-2164-13-694

74. Xue C, Park G, Choi W, Zheng L, Dean RA, Xu J-R. 2002. Two novel fungal virulence genes specifically expressed in appressoria of the rice blast fungus. Plant Cell 14:2107–2119. https://doi.org/10.1105/tpc.003426

75. Grell MN, Mouritzen P, Giese H. 2003. A *Blumeria graminis* gene family encoding proteins with a C-terminal variable region with homologues in pathogenic fungi. Gene 311:181–192. https://doi.org/10.1016/s0378-1119(03)00610-3

76. Zaccaron AZ, Chen L-H, Samaras A, Stergiopoulos I. 2022. A chromosome-scale genome assembly of the tomato pathogen *Cladosporium fulvum* reveals a compartmentalized genome architecture and the presence of a dispensable chromosome. Microb Genom 8:000819. https://doi.org/10.1099/mgen.0.000819

77. Gan P, Hiroyama R, Tsushima A, Masuda S, Shibata A, Ueno A, Kumakura N, Narusaka M, Hoat TX, Narusaka Y, Takano Y, Shirasu K. 2020. Subtelomeric regions and a repeat-rich chromosome harbor multicopy effector gene clusters with variable conservation in multiple plant pathogenic *Colletotrichum* species. bioRxiv. https://doi.org/10.1101/2020.04.28.061093

78. Castanera R, López-Varas L, Borgognone A, LaButti K, Lapidus A, Schmutz J, Grimwood J, Pérez G, Pisabarro AG, Grigoriev IV, Stajich JE, Ramírez L, Feschotte C. 2016. Transposable elements versus the fungal genome: impact on whole-genome architecture and transcriptional profiles. PLoS Genet 12:e1006108. https://doi.org/10.1371/journal.pgen.1006108

79. Amselem J, Lebrun M-H, Quesneville H. 2015. Whole genome comparative analysis of transposable elements provides new insight into mechanisms of their inactivation in fungal genomes. BMC Genomics 16:141. https://doi.org/10.1186/s12864-015-1347-1

80. Dong S, Raffaele S, Kamoun S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Curr Opin Genet Dev 35:57–65. https://doi.org/10.1016/j.gde.2015.09.001

81. Raffaele S, Kamoun S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol 10:417–430. https://doi.org/10.1038/nrmicro2790

82. Sood S, Sharma A, Sharma N, Kanwar SS. 2018. Carboxylesterases: sources, characterization and broader applications. Insights Enzyme Res 01:2. https://doi.org/10.21767/2573-4466.100002

83. Oakeshott J, Claudianos C, Campbell P, Newcomb R, Russell R 2005. Biochemical genetics and genomics of insect esterases, p. 309–381. *In* Biochemical genetics and genomics of insect esterases. Elsevier, Oxford, Amsterdam.

84. Alam M, Vance DE, Lehner R. 2002. Structure-function analysis of human triacylglycerol hydrolase by site-directed mutagenesis: identification of the catalytic triad and a glycosylation site. Biochemistry 41:6679–6687. https://doi.org/10.1021/bi0255625

85. Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, Thines M, Ah-Fong A, Anderson R, Badejoko W, Bittner-Eddy P, Boore JL, Chibucos MC, Coates M, Dehal P, Delehaunty K, Dong S, Downton P, Dumas B, Fabro G, Fronick C, Fuerstenberg SI, Fulton L, Gaulin E, Govers F, Hughes L, Humphray S, Jiang RHY, Judelson H, Kamoun S, Kyung K, Meijer H, Minx P, Morris P, Nelson J, Phuntumart V, Qutob D, Rehmany A, Rougon-Cardoso A, Ryden P, Torto-Alalibo T, Studholme D, Wang Y, Win J, Wood J, Clifton SW, Rogers J, Van den Ackerveken G, Jones JDG, McDowell JM, Beynon J, Tyler BM. 2010. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome . Science 330:1549–1551. https://doi.org/10.1126/science.1195203

86. Duplessis S, Cuomo CA, Lin Y-C, Aerts A, Tisserant E, Veneault-Fourrey C, Joly DL, Hacquard S, Amselem J, Cantarel BL, Chiu R, Coutinho PM, Feau N, Field M, Frey P, Gelhaye E, Goldberg J, Grabherr MG, Kodira CD, Kohler A, Kües U, Lindquist EA, Lucas SM, Mago R, Mauceli E, Morin E, Murat C, Pangilinan JL, Park R, Pearson M, Quesneville H, Rouhier N, Sakthikumar S, Salamov AA, Schmutz J, Selles B, Shapiro H, Tanguay P, Tuskan GA, Henrissat B, Van de Peer Y, Rouzé P, Ellis JG, Dodds PN, Schein JE, Zhong S, Hamelin RC, Grigoriev IV, Szabo LJ, Martin F. 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc Natl Acad Sci U S A 108:9166–9171. https://doi.org/10.1073/pnas.1019315108

87. RoyChowdhury M, Sternhagen J, Xin Y, Lou B, Li X, Li C. 2022. Evolution of pathogenicity in obligate fungal pathogens and allied genera. PeerJ 10:e13794. https://doi.org/10.7717/peerj.13794

88. Morris JJ, Lenski RE, Zinser ER. 2012. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. mBio 3:e00036-12. https://doi.org/10.1128/mBio.00036-12

89. Hauser PM. 2014. Genomic insights into the fungal pathogens of the genus *Pneumocystis*: obligate biotrophs of humans and other mammals. PLoS Pathog 10:e1004425. https://doi.org/10.1371/journal.ppat.1004425

90. Morris JJ. 2015. Black Queen evolution: the role of leakiness in structuring microbial communities. Trends Genet 31:475–482. https://doi.org/10.1016/j.tig.2015.05.004

91. Loeffler RST, Butters JA, Hollomon DW. 1992. The sterol composition of powdery mildews. Phytochemistry 31:1561–1563. https://doi.org/10.1016/0031-9422(92)83106-9

92. Debieu D, Corio-Costet M-F, Steva H, Malosse C, Leroux P. 1995. Sterol composition of the Vine powdery mildew fungus, *Uncinula necator*: comparison of triadimenol-sensitive and resistant strains. Phytochemistry 39:293–300. https://doi.org/10.1016/0031-9422(95)00015-Y

93. Kodedová M, Sychrová H. 2015. Changes in the sterol composition of the plasma membrane affect membrane potential, salt tolerance and the activity of multidrug resistance pumps in *Saccharomyces cerevisiae*. PLoS One 10:e0139306. https://doi.org/10.1371/journal.pone.0139306

94. Aguilar PS, Heiman MG, Walther TC, Engel A, Schwudke D, Gushwa N, Kurzchalia T, Walter P. 2010. Structure of sterol aliphatic chains affects yeast cell shape and cell fusion during mating. Proc Natl Acad Sci U S A 107:4170–4175. https://doi.org/10.1073/pnas.0914094107

95. Hu C, Zhou M, Wang W, Sun X, Yarden O, Li S. 2018. Abnormal ergosterol biosynthesis activates transcriptional responses to antifungal azoles. Front Microbiol 9:fox063. https://doi.org/10.3389/fmicb.2018.00009

96. Liu G, Chen Y, Færgeman NJ, Nielsen J. 2017. Elimination of the last reactions in ergosterol biosynthesis alters the resistance of *Saccharomyces cerevisiae* to multiple stresses. FEMS Yeast Res 17:fox063. https://doi.org/10.1093/femsyr/fox063

97. Bhattacharya S, Esquivel BD, White TC, Lorenz M. 2018. Overexpression or deletion of ergosterol biosynthesis genes alters doubling time, response to stress agents, and drug susceptibility in Saccharomyces cerevisiae. mBio 9:e01291-18. https://doi.org/10.1128/mBio.01291-18

98. Long N, Xu X, Zeng Q, Sang H, Lu L. 2017. *Erg4A* and *Erg4B* are required for conidiation and azole resistance via regulation of ergosterol biosynthesis in *Aspergillus fumigatus*. Appl Environ Microbiol 83:e02924-16. https://doi.org/10.1128/AEM.02924-16

99. Long N, Zhong G. 2022. The C-22 sterol desaturase Erg5 is responsible for ergosterol biosynthesis and conidiation in *Aspergillus fumigatus*. J Microbiol:1–7. https://doi.org/10.1007/s12275-022-1564-7

100. Liu X, Jiang J, Yin Y, Ma Z. 2013. Involvement of *FgERG4* in ergosterol biosynthesis, vegetative differentiation and virulence in *Fusarium graminearum*. Mol Plant Pathol 14:71–83. https://doi.org/10.1111/j.1364-3703.2012.00829.x

101. Liu J, Chai X, Guo T, Wu J, Yang P, Luo Y, Zhao H, Zhao W, Nkechi O, Dong J, Bai J, Lin Q. 2019. Disruption of the ergosterol biosynthetic pathway results in increased membrane permeability, causing overproduction and secretion of extracellular *Monascus* pigments in submerged fermentation. J Agric Food Chem 67:13673–13683. https://doi.org/10.1021/acs.jafc.9b05872

102. Sun X, Wang W, Wang K, Yu X, Liu J, Zhou F, Xie B, Li S. 2013. Sterol C-22 desaturase ERG5 mediates the sensitivity to antifungal azoles in *Neurospora crassa and Fusarium verticillioides*. Front Microbiol 4:127. https://doi.org/10.3389/fmicb.2013.00127

103. Martel CM, Parker JE, Bader O, Weig M, Gross U, Warrilow AGS, Kelly DE, Kelly SL. 2010. A clinical isolate of *Candida albicans* with mutations in *ERG11* (encoding sterol 14alpha-demethylase) and *ERG5* (encoding C22 desaturase) is cross resistant to azoles and amphotericin B. Antimicrob Agents Chemother 54:3578–3583. https://doi.org/10.1128/AAC.00303-10

104. Klemptner RL, Sherwood JS, Tugizimana F, Dubery IA, Piater LA. 2014. Ergosterol, an orphan fungal microbe-associated molecular pattern (MAMP). Mol Plant Pathol 15:747–761. https://doi.org/10.1111/mpp.12127

105. Laquitaine L, Gomès E, François J, Marchive C, Pascal S, Hamdi S, Atanassova R, Delrot S, Coutos-Thévenot P. 2006. Molecular basis of ergosterol-induced protection of grape against *Botrytis cinerea*: induction of type I LTP promoter activity, WRKY, and stilbene synthase gene expression. Mol Plant Microbe Interact 19:1103–1112. https://doi.org/10.1094/MPMI-19-1103

106. Gomès E, Sagot E, Gaillard C, Laquitaine L, Poinssot B, Sanejouand Y-H, Delrot S, Coutos-Thévenot P. 2003. Nonspecific lipid-transfer protein genes expression in grape (*Vitis sp.*) cells in response to fungal elicitor treatments. Mol Plant Microbe Interact 16:456–464. https://doi.org/10.1094/MPMI.2003.16.5.456

107. Coutos-Thévenot P, Poinssot B, Bonomelli A, Yean H, Breda C, Buffard D, Esnault R, Hain R, Boulay M. 2001. In vitro tolerance to *Botrytis cinerea* of grapevine 41B rootstock in transgenic plants expressing the stilbene synthase *Vst1* gene under the control of a pathogen-inducible PR 10 promoter. J Exp Bot 52:901–910. https://doi.org/10.1093/jexbot/52.358.901

108. Schnee S, Viret O, Gindro K. 2008. Role of stilbenes in the resistance of grapevine to powdery mildew. Physiol Mol Plant Pathol 72:128–133. https://doi.org/10.1016/j.pmpp.2008.07.002

109. Muszewska A, Steczkiewicz K, Stepniewska-Dziubinska M, Ginalski K. 2019. Transposable elements contribute to fungal genes and impact fungal lifestyle. Sci Rep 9:4307. https://doi.org/10.1038/s41598-019-40965-0

110. Belyayev A. 2014. Bursts of transposable elements as an evolutionary driving force. J Evol Biol 27:2573–2584. https://doi.org/10.1111/jeb.12513

111. Kuzmin E, Taylor JS, Boone C. 2022. Retention of duplicated genes in evolution. Trends Genet 38:59–72. https://doi.org/10.1016/j.tig.2021.06.016

112. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. 2009. Mechanisms of change in gene copy number. Nat Rev Genet 10:551–564. https://doi.org/10.1038/nrg2593

113. Cerbin S, Jiang N. 2018. Duplication of host genes by transposable elements. Curr Opin Genet Dev 49:63–69. https://doi.org/10.1016/j.gde.2018.03.005

114. Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. 2013. Gene duplication as a major force in evolution. J Genet 92:155–161. https://doi.org/10.1007/s12041-013-0212-8

115. Birchler JA, Yang H. 2022. The multiple fates of gene duplications: deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. Plant Cell 34:2466–2474. https://doi.org/10.1093/plcell/koac076

116. Qiao X, Yin H, Li L, Wang R, Wu J, Wu J, Zhang S. 2018. Different modes of gene duplication show divergent evolutionary patterns and contribute differently to the expansion of gene families involved in important fruit traits in pear (*Pyrus bretschneideri*). Front Plant Sci 9:161. https://doi.org/10.3389/fpls.2018.00161

117. Wang Y, Wang X, Tang H, Tan X, Ficklin SP, Feltus FA, Paterson AH. 2011. Modes of gene duplication contribute differently to genetic novelty and redundancy, but show parallels across divergent angiosperms. PLoS One 6:e28150. https://doi.org/10.1371/journal.pone.0028150

118. Katju V, Bergthorsson U. 2013. Copy-Number changes in evolution: rates, fitness effects and adaptive significance. Front Genet 4:273. https://doi.org/10.3389/fgene.2013.00273

119. Menardo F, Praz CR, Wicker T, Keller B. 2017. Rapid turnover of effectors in grass powdery mildew (*Blumeria graminis*). BMC Evol Biol 17:223. https://doi.org/10.1186/s12862-017-1064-2

120. Ross MK, Streit TM, Herring KL, Xie S. 2010. Carboxylesterases: dual roles in lipid and pesticide metabolism. J Pestic Sci 35:257–264. https://doi.org/10.1584/jpestics.R10-07

121. Oakeshott JG, Claudianos C, Russell RJ, Robin GC. 1999. Carboxyl/cholinesterases: a case study of the evolution of a successful multigene family. Bioessays 21:1031–1042. https://doi.org/10.1002/(SICI)1521-1878(199912)22:1<1031::AID-BIES7>3.0.CO;2-J

122. Urquhart AS, Hu J, Chooi Y-H, Idnurm A. 2019. The fungal gene cluster for biosynthesis of the antibacterial agent viriditoxin. Fungal Biol Biotechnol 6:2. https://doi.org/10.1186/s40694-019-0072-y

123. Hu J, Li H, Chooi Y-H. 2019. Fungal dirigent protein controls the stereoselectivity of multicopper oxidase-catalyzed phenol coupling in viriditoxin biosynthesis. J Am Chem Soc 141:8068–8072. https://doi.org/10.1021/jacs.9b03354