Tarallo *et al. Phytopathology Research* (2024) 6:37
https://doi.org/10.1186/s42483-024-00255-8

Phytopathology Research

RESEARCH

Open Access

Genomic and culture-based analysis of *Cyclaneusma minus* in New Zealand provides evidence for multiple morphotypes

Mariana Tarallo[1,2*], Kiryn Barbara Dobbie[3], Luciano Nunes Leite[3,4], Tammy Leigh Waters[3], Kristin Nikki Tasmin Gillard[3], Diya Sen[3], Carl Hayden Mesarich[2,5], Rosie Ellen Bradshaw[1,2] and Rebecca Louise McDougal[3]

Abstract

Cyclaneusma needle cast, caused by *Cyclaneusma minus*, affects *Pinus* species world wide. Previous studies suggested the presence of two distinct morphotypes in New Zealand, 'verum' and 'simile'. Traditional mycological analyses revealed a third morphotype with clear differences in colony morphology and cardinal growth rates at varying temperatures. Genome sequencing of eight *C. minus* isolates provided further evidence of the existence of a third morphotype, named as 'novus' in this study. To further analyse these morphotypes, we predicted candidate effector proteins for all eight isolates, and also characterized a cell-death eliciting effector family, Ecp32, which is present in other pine phytopathogens. In concordance with their distinct classification into three different morphotypes, the number of Ecp32 family members differed, with patterns of pseudogenization in the 'simile' morphotype, and some members being found exclusively either in the 'simile' or 'verum' morphotypes. We also showed that the Ecp32 family proteins trigger cell death in non-host *Nicotiana* species, and, as previously demonstrated in other plant pathogens, the Ecp32 family proteins in *C. minus* adopt a β-trefoil fold. These analyses provide further evidence that the three morphotypes might be distinct species that need formal descriptions. Understanding the geographical range of different *Cyclaneusma* species and variations in virulence and pathogenicity will provide a better understanding of pine needle diseases and enable the development of more durable methods to control this disease.

Keywords *Cyclaneusma minus*, Morphotypes, Cyclaneusma needle cast, Genome sequencing, Ecp32 candidate effector family

Background

Cyclaneusma needle cast (CNC) is a disease of *Pinus* species in many parts of the world (Bednářová et al. 2013). CNC affects more than 30 pine species, including *Pinus radiata* in New Zealand and Australia, while in forests of Europe and North America, Scots pine (*Pinus sylvestris*) is prone to this disease (Merrill and Wenner 1996; Müller & Kurkela 2012). CNC results in premature needle loss, and therefore losses in growth and wood production. For New Zealand's plantation forestry, the losses caused by CNC were estimated to be around $38 million per year (Bulman 2009). Fortunately, this impact has likely now

*Correspondence:
Mariana Tarallo
M.Tarallo@massey.ac.nz
[1] School of Natural Sciences, Massey University, Palmerston North 4410, New Zealand
[2] Bioprotection Aotearoa, Massey University, Palmerston North 4410, New Zealand
[3] Scion, Titokorangi Drive, Private Bag 3020, Rotorua 3046, New Zealand
[4] New Zealand Ministry of Primary Industries, PO Box 2095, Auckland 1140, New Zealand
[5] School of Agriculture and Environment, Massey University, Palmerston North 4410, New Zealand

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

been reduced through the planting of less susceptible genotypes. Regardless, the disease persists in some locations, and there are no alternative control measures (Bulman 1993).

Cyclaneusma needle cast is caused by the Leotiomycete fungus *Cyclaneusma minus* (Butin) DiCosmo, Peredo and Minter. *C. minus* was initially reported as *Naemacyclus niveus* (Gilmour 1959), which was later separated into two species, *N. niveus* and *N. minor* (Von Butin 1973). This separation was based on morphological characters on needles, the formation of the sexual and asexual stages in culture, and the relationship with different host species. *Naemacyclus niveus* and *N. minor* were later reassigned to the genus *Cyclaneusma* as *Cyclaneusma niveum* and *C. minus*, respectively (Dicosmo et al. 1983). The morphological features distinguishing *C. niveum* and *C. minus* were described in Von Butin (1973) and again in Dick et al. (2001), where *C. niveum* isolates had sickle-shaped pycnidiospores and much shorter apothecia, and *C. minus* had bacilliform-shaped pycnidiospores and longer apothecia. Further investigation of the *C. minus* isolates in New Zealand found that, based on apothecial length variation as well as fungal morphology in culture, *C. minus* should be separated into two morphotypes: tagged as *C. minus* 'verum' and *C. minus* 'simile' (Dick et al. 2001). A relationship was shown between apothecium dimensions from needles and these morphotypes. A multigene phylogenetic analysis of New Zealand and Australian isolates of *Cyclaneusma* demonstrated that the separation of isolates by morphotype was consistent with genetic analysis and suggested the need for species descriptions to formally re-classify, as well as pathogenicity studies to understand phenotypic differences between the morphotypes (Prihatini et al. 2014).

Hunter et al. (2016) developed molecular tools for differentiating the 'verum' and 'simile' morphotypes and looked at their distribution throughout New Zealand. The study found that 'simile' was the most common morphotype in New Zealand, and that molecular classification was much more accurate than morphological assessments. Amongst the isolates studied, one (NZFS3325) showed DNA sequence variability in the Internal Transcribed Spacer (ITS) region of its genome, when compared to the corresponding region in both the 'simile' and 'verum', which indicated that additional morphotypes might exist (Hunter et al. 2016). Rightfully identifying these different morphotypes would help to determine their current distribution as well as to predict future distribution under different climatic scenarios, in New Zealand and globally. Comparative studies between morphotypes can also determine whether a morphotype is more virulent than others and can have implications for disease diagnostics and forest management.

To further characterize the morphotypes of *C. minus*, we carried out genome sequencing for eight isolates: two belonging to morphotype 'simile' (Cms; NZFS759 and NZFS809), four belonging to morphotype 'verum' (Cmv; NZFS110, NZFS725, NZFS1800, and NZFS3276), and two isolates (Cmn; NZFS3305 and NZFS3325) that could not be differentiated into 'verum' or 'simile' using previously developed molecular tools. This genomic information will provide additional data to assist taxonomic classification. An additional benefit of genomic resources is that they can provide insights into the virulence on their hosts, through the prediction of candidate effector (CE) proteins. Effectors are usually small proteins secreted by pathogens into and around host cells to facilitate invasion, for example by suppressing host defence responses (Lo Presti et al. 2015; Rocafort et al. 2020). If the host has the ability to recognize such effectors, through its extracellular and intracellular resistance (R) proteins, immune responses are activated in order to stop pathogen infection (Jones and Dangl 2006; Cook et al. 2015). This recognition can lead to a hypersensitive response, which is generally characterized by a localized plant cell death (Win et al. 2012). Notably, pathogen effectors can be heterologously expressed in host and non-host plants, with those triggering a plant cell death response being candidates for recognition by a plant R protein.

In this study, genome sequencing was performed for eight *C. minus* isolates with the main objective of confirming if a third morphotype is present in New Zealand, and its relationship to other morphotypes. Moreover, the genome sequences of the different *C. minus* morphotypes were examined for genes encoding CE proteins. Here, we focused on homologues of the cell death-eliciting Ecp32 protein family, firstly identified in the tomato pathogen *Fulvia fulva* and subsequently in the pine pathogen *Dothistroma septosporum* (Mesarich et al. 2018; Tarallo et al. 2022). We compared their sequences, predicted tertiary structures, and screened these homologues for their ability to trigger cell death in *Nicotiana benthamiana* and *N. tabacum* using *Agrobacterium tumefaciens*-mediated transient transformation assays (ATTAs). By doing so, we aimed to provide novel insights into whether the *C. minus* morphotypes differ with respect to the conserved Ecp32 family, and to highlight potential differences in the molecular basis for how the different morphotypes interact with their hosts.

Results

C. minus isolates can be differentiated into three morphotypes based on molecular and morphological analyses

Eight *C. minus* isolates were collected from different parts of New Zealand and classified into two different morphotypes, 'verum' (Cmv) and 'simile' (Cms), according to molecular analysis (Table 1 and Additional file 1: Figure S1) (Hunter et al. 2016). However, two of these isolates, NZFS3305 and NZFS3325 (NZFS3325 was previously identified in Hunter et al. (2016)), could not be differentiated based on this molecular analysis and were, therefore, classified as a new *C. minus* morphotype, 'novus' (Cmn). The phylogeny suggested that morphotype Cmn is more closely related to Cms than to Cmv (Additional file 1: Figure S1).

This classification was also supported by morphological analyses. Results for the temperature study are shown in Fig. 1. The widest temperature range for growth, as well as the fastest growth rate, was seen in the Cmv morphotype. It had an optimal growth at 25 °C, with a mean radial growth rate of 2.0 ± 0.7 mm/day at 25 °C. However, growth at our lowest (4°C) and highest (30°C) temperatures were also detected for this

Table 1 Isolates of *Cyclaneusma minus* used in this study

Isolate number[a]	New Zealand location[b]	Host	*C. minus* 'simile' qPCR[c]	*C. minus* 'verum' qPCR	Morphotype[d]
NZFS110	South Canterbury	*Pinus radiata*	-	+	verum
NZFS725	Coromandel	*P. radiata*	-	+	verum
NZFS1800	Dunedin	*P. radiata*	-	+	verum
NZFS3276	Bay of Plenty	*P. radiata*	-	+	verum
NZFS759	Northland	*P. radiata*	+	-	simile
NZFS809	Southland	*P. radiata*	+	-	simile
NZFS3305	Bay of Plenty	*P. radiata*	-	-	novus[e]
NZFS3325	Bay of Plenty	*P. radiata*	-	-	novus[e]

[a] Scion Forest Health Reference Laboratory Culture Collection number (NZFS)

[b] Location in New Zealand where the isolate was collected, according to Crosby et al. (1998)

[c] qPCR, quantitative polymerase chain reaction (McDougal et al. 2016)

[d] Morphotype based on molecular analysis according to Hunter et al. (2016)

[e] 'novus' is a new morphotype name proposed in this study

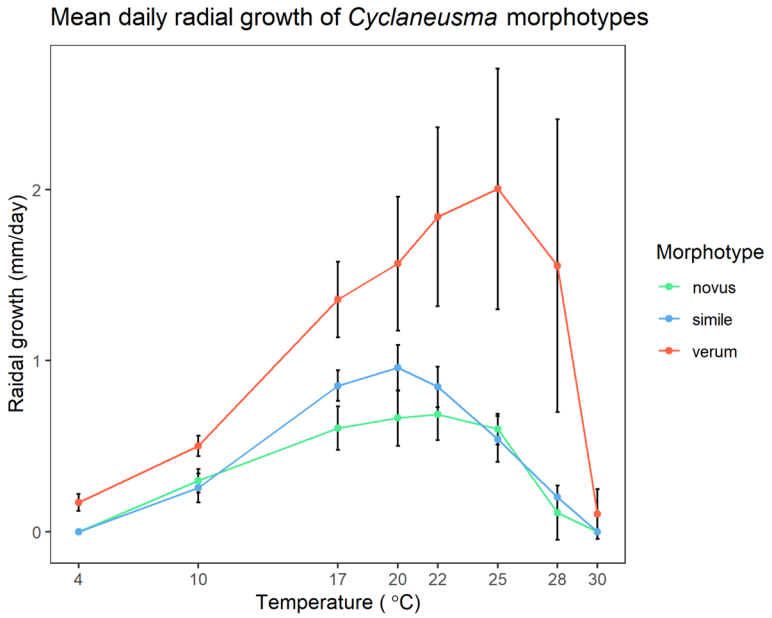

Fig. 1 Effect of temperature on the growth of *C. minus* morphotypes. The 'simile' (blue), 'verum' (red), and 'novus' (green) morphotypes were assessed using mean radial growth, measured in millimetres per day with standard deviations

morphotype. The Cms and Cmn morphotypes shared similar optimal growth temperatures at 20°C and 22°C, respectively, while growth ceased at 4°C and 30°C for both. The Cms morphotype had a slightly faster mean radial growth rate of 0.96 ± 0.13 mm/day at its optimal growth temperature of 20°C, compared to that of the Cmn morphotype being 0.69 ± 0.15 mm/day at 22°C. Analysis of temperature study data using two-way ANOVA with Tukey's post hoc showed that the radial growth per day of Cmv was statistically different from Cms and Cmn ($p = 1.8 \times 10^6$ and 2.9×10^5 respectively), while there was no significant difference in the radial growth between Cms and Cmv ($p = 0.65$). Isolates and their corresponding morphotypes grown on 2% and 9% MEA at 22°C in the dark after 8 weeks are shown in Fig. 2. The colony morphology of each morphotype on 2% and 9% MEA at 22°C in the dark after 8 weeks was displayed in Additional file 2: Table S1.

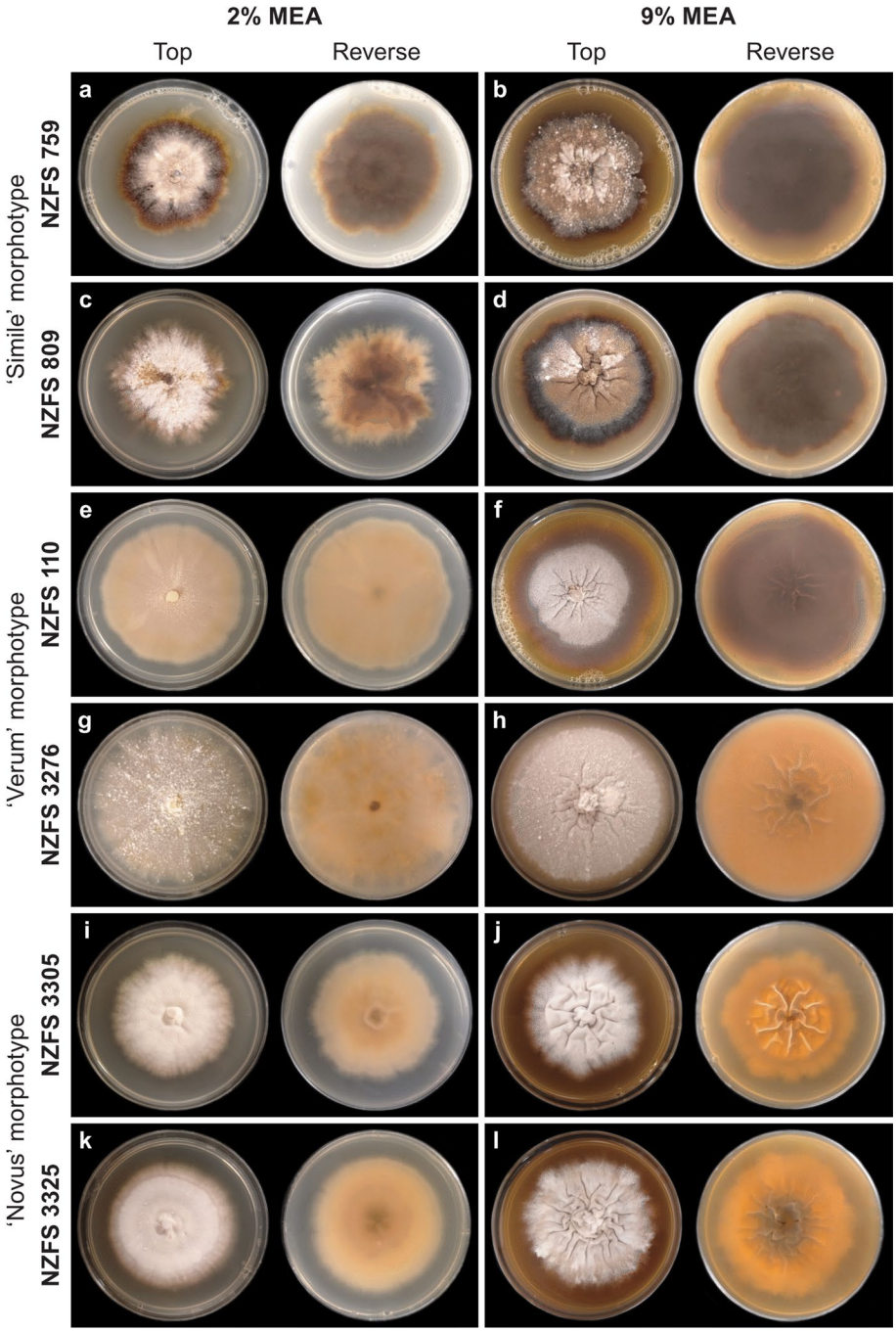

Fig. 2 Colony morphology of *C. minus* isolates. All isolates were grown at 22°C in the dark for 8 weeks

Differences in *C. minus* morphotype are supported by differences in genome size

The final genome assemblies for the eight *C. minus* isolates ranged in length from 31 to 43 Mb (Table 2 and Additional file 2: Table S2). The final set of gene models for each genome showed a completeness of > 96% with the retrieval of at least 728 complete BUSCO hits from a total of 758 fungal BUSCOs (Table 2 and Additional file 2: Table S2). The Cmv morphotype was predicted to have a larger genome size (based on assembly length) and higher number of gene models compared to the Cms and Cmn morphotypes (Table 2).

Prediction of candidate effector proteins from *C. minus* morphotypes

Effectors can have virulence functions during host infection by supressing plant defence responses or can be avirulence factors, in which case they are recognized by R proteins from the hosts. For this reason, following genome sequencing and assembly, secreted CE proteins were identified from the predicted protein set of all eight *C. minus* isolates to determine whether their CE protein repertoires were morphotype-specific. Strikingly, genomes from the same morphotype were predicted to have similar numbers of proteins and effectors, with Cmv having more predicted CEs than Cms or Cmn (Table 3). In all cases, a higher number of apoplastic effectors, when compared to cytoplasmic effectors, were predicted in all isolates analysed. Additional file 2: Table S3 shows all amino acid sequences of the CEs predicted from the isolates by using EffectorP.

The number of Ecp32 effector family members differs between the three morphotypes of *C. minus*, further supporting their differentiation

To investigate potential differences in the effector protein repertoires of *C. minus* morphotypes, we focused on the Ecp32 effector family that had been characterised in another fungal pine pathogen (Tarallo et al. 2022). An apoplastic protein from isolate NZFS809, Cms835, was firstly identified as an orthologue of DsEcp32-1 from *D. septosporum* and shown to be part of the Ecp32 effector family. With the Cms835 (NZFS809) amino acid sequence as reference, other Ecp32 family members in this and other isolates/morphotypes were identified using BLASTp and tBLASTn, respectively. Based on this analysis, it was determined that all isolates of all morphotypes had at least three members of the Ecp32 family of proteins (Fig. 3 and Additional file 2: Table S4).

Although all isolates and morphotypes appeared to have proteins belonging to the Ecp32 family, the number of family members between each morphotype varied. A BLASTp search indicated that both the NZFS759 and NZFS809 genomes (both Cms) encode eight Ecp32 family proteins. Each of the Cmv isolates (NZFS110, NZFS725, NZFS1800, and NZFS3276) have six paralogues in their Ecp32 family, while Cmn isolates (NZFS3305 and NZFS3325) have only three. The *Ecp32* genes from one representative isolate of each morphotype are illustrated in Fig. 3. Unlike what was observed for orthologues between *D. septosporum* and *F. fulva*, proteins from the Ecp32 family from each *C. minus* isolate did not group with the previously identified members of the Ds/FfEcp32 family (Additional file 1: Figure S2).

Table 2 *Cyclaneusma minus* genomes statistics

Isolate[a]	Morphotype[b]	NCBI accession	Scaffolds[c]	Gene models	N50[d] (bp)	Assembly length (bp)	Complete BUSCOs (%)	GC (%)
NZFS110	verum	JAKZHM000000000	2616	10,445	570.2	42,876,071	96.6	45.6
NZFS725	verum	JAKZHN000000000	1804	10,409	1103.7	41,971,199	96.9	45.8
NZFS1800	verum	JAKZHQ000000000	1644	10,455	438.3	42,298,080	97.1	45.8
NZFS3276	verum	JAKZHT000000000	2004	10,397	528.9	42,485,107	97.3	45.7
NZFS759	simile	JAKZHO000000000	2588	9659	833.0	36,285,465	96.3	47.9
NZFS809	simile	JAKZHP000000000	1548	9684	666.1	35,864,553	97.3	48.0
NZFS3305	novus*	JAKZHR000000000	610	9266	126.2	31,933,912	97.6	51.2
NZFS3325	novus*	JAKZHS000000000	732	9269	119.0	32,988,501	97.7	51.1

[a] Scion Forest Health Reference Laboratory Culture Collection number (NZFS)

[b] 'verum' and 'simile' morphotypes according to Hunter et al. (2016). * 'novus' is a new morphotype name proposed in this study

[c] Number of scaffolds assembled using Redundans v0.13c (Pryszcz and Gabaldón 2016)

[d] The minimum contig length among contigs required to cover 50% of the assembled genome sequence

Table 3 Effector protein prediction statistics across the eight *C. minus* isolates investigated in this study

Isolate[a]	Morpho-type[b]	Predicted proteins[c]	Secreted proteins[d]	No membrane affiliation[e]	SSCP[f]	Effectors[g]	Cyto[g] effectors	Apo[g] effectors
NZFS110	verum	10,445	729	582	117	92	16	76
NZFS725	verum	10,409	714	568	109	90	17	73
NZFS1800	verum	10,455	721	578	120	100	18	82
NZFS3276	verum	10,397	715	568	111	94	17	77
NZFS759	simile	9659	642	504	95	72	17	55
NZFS809	simile	9684	640	497	91	69	13	56
NZFS3305	novus[h]	9266	639	509	91	70	17	53
NZFS3325	novus[h]	9269	629	499	87	68	15	53

[a] Scion Forest Health Reference Laboratory Culture Collection number (NZFS)

[b] Morphotype according to Hunter et al. (2016)

[c] Number of proteins predicted from each of the assembled *Cyclaneusma* genomes

[d] Number of predicted secreted proteins after being queried with SignalP v3.0 (Bendtsen et al. 2004)

[e] Number of predicted secreted proteins with no putative membrane affiliation after being queried with TMHMM v2.0 (Krogh et al. 2001) and BIG-PI (Eisenhaber et al. 2004)

[f] Number of predicted small, secreted cysteine-rich proteins (SSCPs), of ≤ 300 amino acids in length with ≥ 4 cysteine residues

[g] Number of effector proteins predicted using EffectorP v3.0 (Sperschneider and Dodds 2021). Effector proteins were classified as cytoplasmic (cyto) or apoplastic (apo)

[h] 'novus' is a new morphotype name proposed in this study

As was also found for the *D. septosporum* and *F. fulva* Ecp32 proteins, all Ecp32 family members from *C. minus* have predicted signal peptides but no known functional domains.

Some Ecp32 family members are found exclusively in one morphotype. This is the case for a set of orthologous proteins (Cmv2795, Cmv1388, Cmv116, and Cmv3467) from the four Cmv isolates. These proteins have longer amino acid sequences than the rest of the proteins in this family. Normally, they range from 197 to 226 amino acids in length, however these four proteins each contain 268 amino acids. The additional amino acids form a predicted intrinsically disordered region (IDR) at the N-terminus, immediately following the putative signal peptide (Additional file 1: Figure S3).

The morphotypes also showed differences in pseudogenisation of *Ecp32* genes. For Cms isolates NZFS809 and NZFS759, three of the eight *Ecp32* genes were pseudogenes that were predicted to encode truncated proteins, with identical mutations in the homologues from these two isolates (NZFS809 is shown in Fig. 3). Two of the genes (*Cms834* and *Cms4060*) had a premature stop codon at the

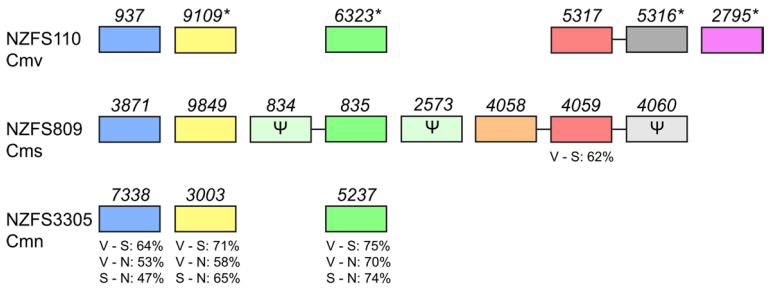

Fig. 3 Schematic diagrams of the *Ecp32* family from three different morphotypes of *C. minus*. One representative isolate from each morphotype was chosen: NZFS110 for morphotype 'verum' (Cmv), NZFS809 for morphotype 'simile' (Cms), and NZFS3305 for morphotype 'novus' (Cmn). Each rectangle represents a gene and orthologous genes between morphotypes are colour-coded in the same way as Fig. 4. Ψ refers to pseudogenes. The horizontal lines connecting the rectangles indicate the genes are on the same scaffold and adjacent in the genome. The percentages refer to full-length pairwise amino acid identity between orthologous proteins encoded by each gene from each representative morphotype (V = 'verum', S = 'simile', N = 'novus'). The pink rectangle indicates the gene encoding the protein with a predicted intrinsically disordered region. An asterisk (*) next to the gene number indicates amino acid variants in the respective encoded proteins that occurred in one or more of the other Cmv isolates. NZFS3276 orthologues of 9109, 6323, and 2795 have variants I34T, V188I, and E40D, respectively. Orthologues of the other polymorphic protein showed that 5316 has a Q82H substitution in all three other Cmv isolates and a S215G substitution in isolates NZFS1800 and 3276

position encoding amino acids 54 and 149, respectively, whilst *Cms2573* had a cytosine insertion at nt position 213 in the coding sequence, causing a frameshift that also led to a premature stop codon. Interestingly, in both Cms isolates, *Cms834* (and the NZFS759 orthologue) is on the same scaffold and adjacent to *Cms835* (and the NZFS759 orthologue) (Fig. 3). Similarly, another gene cluster in this morphotype includes the *Cms4060* pseudogene, alongside *Cms4058* and *Cms4059*, which are predicted to encode functional proteins (Fig. 3). This is also true for the respective orthologues from NZFS759.

Given that the Cms and Cmn morphotypes are the most closely-related of the three morphotypes based on the multigene phylogeny generated in this study (Additional file 1: Figure S1), the difference in the number of *Ecp32* gene family members between them was striking, although three of the eight Cms genes were pseudogenised (Fig. 3). Furthermore, the three genes that shared orthologues between Cms and Cmn showed sequence diversity, with predicted amino acid sequence identities of only 47, 65, and 74% between these morphotypes (Fig. 3). Together, these results support the identification of Cmn as a distinct morphotype from Cms and Cmv, and also suggest that several gene duplication events and possibly gene loss may have occurred in this effector family, either in *C. minus* or in an ancestral species.

The Ecp32 family from different *C. minus* morphotypes are sequence-divergent and some members are found exclusively in certain morphotypes

To further investigate differences between the sets of Ecp32 family members across the morphotypes and isolates of *C. minus*, we performed a protein sequence alignment of all full-length members from the Ecp32 family (Additional File 1: Figure S3). Family members that had identical amino acid sequences were excluded from this alignment, as were sequences encoded by pseudogenes. Figure 4 shows a phylogenetic tree constructed based on these alignments, including all family members identified, even those with identical amino acid sequences, but still excluding those encoded by pseudogenes.

The results affirm that proteins from the Ecp32 family are more expanded in some morphotypes than in others, and they cluster into seven separate groups (colour-coded to match Fig. 3). The clustered groups coloured in blue, yellow, and green are made up of eight protein members, with one paralogue belonging to each of the eight genomes analysed. The clustered group colour-coded in red has six members, with paralogues only belonging to Cms and Cmv, and no members belonging to Cmn. The pink group corresponds to Cmv proteins with a predicted disordered region of 73 additional amino acids, which are most closely related to green group proteins, despite their larger size.

A high level of sequence divergence was seen between the Ecp32 proteins from the three morphotypes. Between all orthologous Cms and Cmv Ecp32 protein family members, the pairwise amino acid identity varied between 25.1% and 71.4%. Likewise, between Cms and Cmn, the pairwise identity was between 32.1% and 68.9%, and between Cmv and Cmn it was 28.8% and 68.6%. Differences were also observed between isolates of the same morphotype, as the amino acid sequence identity among the Ecp32 proteins ranged from 25.3–50.6% (Cmv), 30.9–42.3% (Cms), and 34.4–46.8% (Cmn).

We next compared the locations of clustered *Ecp32* genes with their positions in the phylogeny. All genes encoding Ecp32 proteins of the red and grey groups in the Cmv morphotypes are adjacent to each other in their respective genomes, as shown for *Cmv5316* and *Cmv5317* of isolate NZFS110 in Fig. 3; the corresponding proteins are more similar in sequence to each other than to any other Ecp32 protein from that isolate, with identities ranging from 26.8 to 50.6% (between *Cmv5316* and *Cmv5317*) (Additional file 1: Figure S3). Together, these results suggest that these 'red and grey' family members may have arisen by an ancestral tandem duplication event. A similar linkage of these two family members was seen in the Cms morphotype (genes *Cms4059* and *Cms4060*), although in that case the 'grey' *Cm*s4060 gene was pseudogenised (Fig. 3). Although not shown in the tree, Cms4060 showed greatest similarity with the proteins colour-coded in grey (Additional file 1: Figure S3), which suggests a common ancestral origin for the red and grey-coloured gene pairs of the Cmv and Cms morphotypes. A third *Ecp32* gene was also present in the red-grey gene cluster in the Cms morphotype: *Cms4058* (orange) (Fig. 3). However, despite its close proximity to *Cms4059* in the genome, Cms4058 was more similar in sequence to Cms835 (green) than to Cms4059 (red).

Together, these results highlight the expansion of the *Ecp32* family in the three different morphotypes analysed, with possible gene duplication and the generation of new family members. In addition, sequence diversification was observed, not only between morphotypes, but also between family members from the same isolate.

Proteins from the Ecp32 family of different *C. minus* morphotypes trigger cell death in *Nicotiana* species

The next questions we addressed were whether the Ecp32 family proteins identified from the *C. minus* morphotypes could trigger cell death in *Nicotiana* species, similar to what was observed in *D. septosporum* and *F. fulva* (Tarallo et al. 2022), and whether there were differences in cell death-eliciting ability between morphotypes. The selected proteins tested by ATTAs are indicated with a black dot in Fig. 4. One protein from one representative

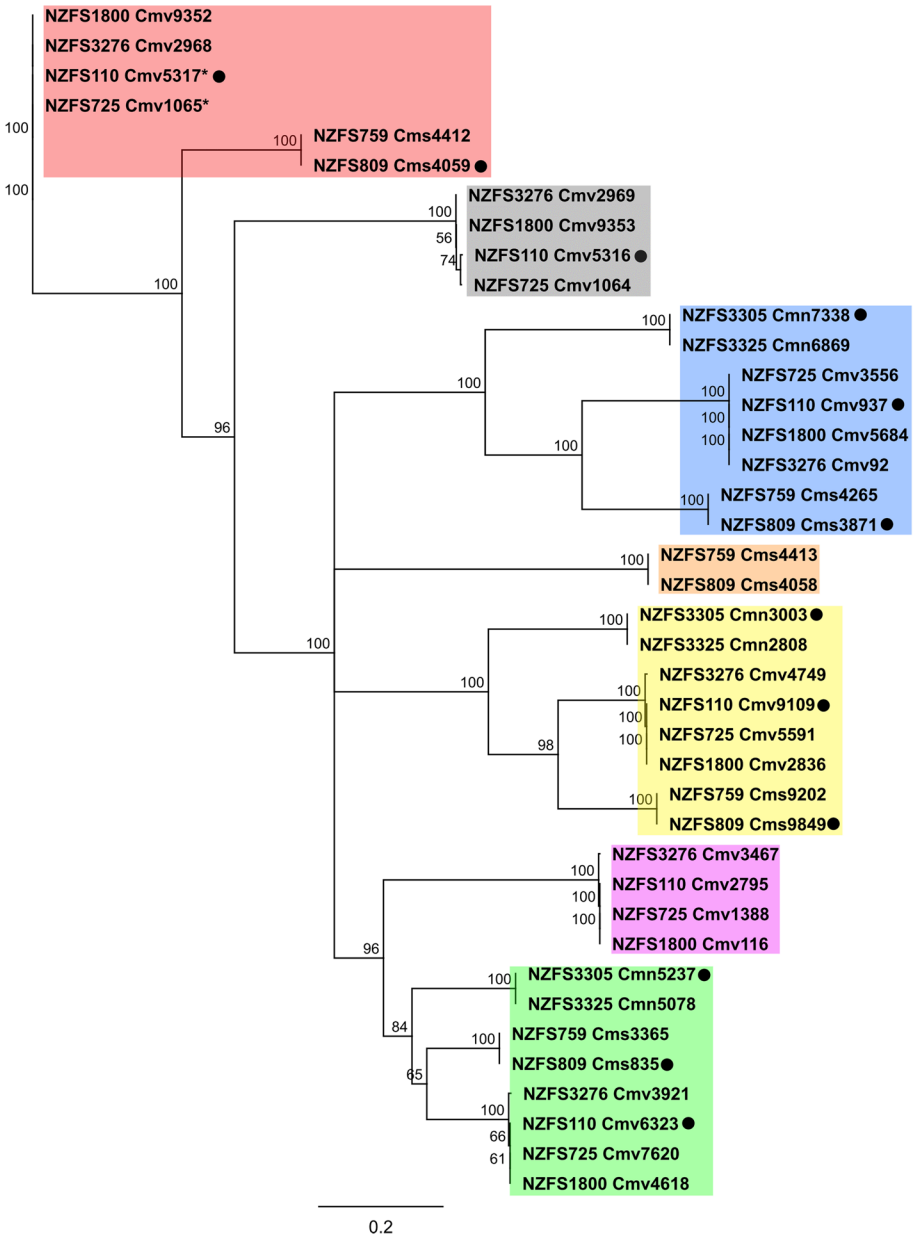

Fig. 4 Phylogenetic tree of *C. minus* Ecp32 family members across eight isolates and three morphotyes of *C. minus*. Neighbour-joining tree obtained from the amino acid sequence alignment of Ecp32 proteins using Geneious Software v9.1.8. Bootstrap values are shown at the nodes as percentages. The scale bar represents 0.2 substitutions per site. The number following each morphotype ID (NZFS) refers to the protein number from the annotated genomes. The asterisk (*) indicates that the protein was incorrectly predicted, and the corrected protein sequence was used. The sequences encoded by predicted pseudogenes were excluded from the analysis. The black dot refers to the proteins that were expressed in *N. benthamiana* and *N. tabacum* using an *Agrobacterium tumefaciens*-mediated transient expression assay to assess their ability to elicit cell death. Cms: *C. minus* morphotype 'simile'; Cmv: *C. minus* morphotype 'verum'; Cmn: *C. minus* morphotype 'novus'. Proteins were colour-coded according to their clustering in the phylogeny

isolate of each morphotype was selected from the clusters that had a higher number of family members.

In *N. benthamiana*, all proteins tested triggered certain degrees of cell death, except for Cmv937 (the Cmv representative from the blue cluster in Fig. 4) (Fig. 5). Among those tested proteins, proteins that consistently triggered cell death were those from the green cluster shown in Fig. 4 (Cmv6323, Cms835, and Cmn5237), and the yellow cluster (Cmv9109, Cms9849, and Cmn3003), with Cms9849 triggering a strong cell death response in all infiltration zones. Across all proteins tested, differences were seen between morphotypes

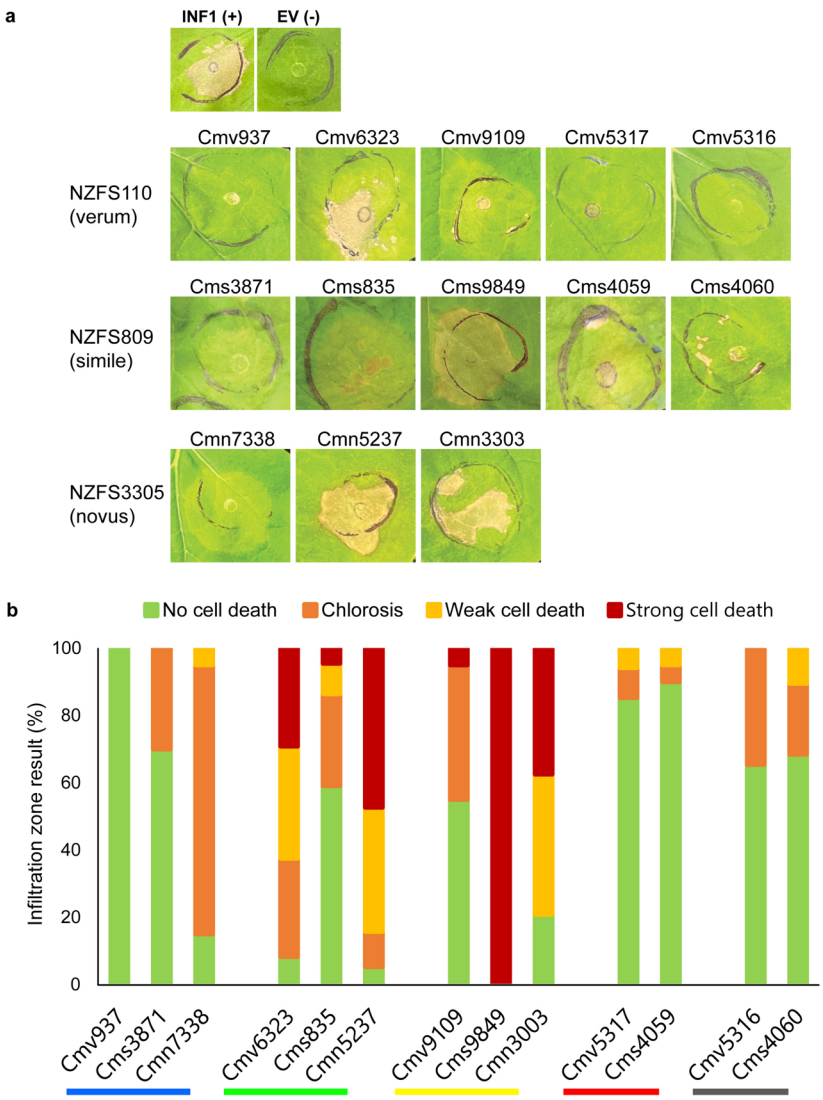

Fig. 5 Ecp32 proteins from different morphotypes of *C. minus* triggered cell death in *Nicotiana benthamiana*. **a** Proteins were expressed in *N. benthamiana* using an *Agrobacterium tumefaciens*-mediated transient expression assay to assess their ability to elicit cell death. Representative images are shown ($n = 12–30$ infiltration zones) from at least three independent experiments. INF1, *Phytophthora infestans* elicitin positive cell death control; EV, empty vector, a negative no-cell death control; Cms: *C. minus* morphotype 'simile'; Cmv: *C. minus* morphotype 'verum'; Cmn: *C. minus* morphotype 'novus'. Photos were taken 7 days after infiltration. **b** Graphs display the percentages of infiltration zones across the entire experiment that showed cell death in response to each different protein, divided into four categories: strong cell death, weak cell death, chlorosis, and no cell death. The proteins are shown as orthologous groups, with the different morphotypes represented for each. The coloured lines under the protein IDs refer to the different gene and protein groups in Figs. 3 and 4

(Fig. 5). The truncated protein Cms4060 (NZFS809) was also tested, which triggered a weak cell death response. Although this protein was not included in the phylogenetic tree shown in Fig. 4, it showed greatest similarity with the proteins colour-coded in grey, and an alignment between these proteins (Additional file 1: Figure S3) indicated conservation of the N-terminus (post signal peptide), which suggests that this region of the protein might be responsible for cell death elicitation, since Cmv5316 also triggered a cell death (Fig. 5).

In *N. tabacum*, yellow cluster proteins Cmv9109, Cms9849, and Cmn3003 (Fig. 4) again consistently triggered strong cell death, while all other proteins inconsistently triggered either weak cell death or chlorosis (Additional file 1: Figure S4). Interestingly, the truncated protein Cms4060 (NZFS809), as well as its paralogue Cmv5316, did not trigger cell death (Additional File 1: Figure S4).

Taken together, these results suggest that different Ecp32 family members from the same isolate differ in

their ability to trigger a cell death response in non-host plants. There were also differences in the responses observed for orthologues from different morphotypes (e.g., proteins from the blue cluster).

Predicted tertiary structures of the *C. minus* Ecp32 family members suggest possible roles in virulence

The tertiary structure of one *C. minus* Ecp32 family member from each of the clusters shown in Fig. 4 was predicted using AlphaFold2 (Additional file 1: Figure S5). The amino acid sequences of proteins from NZFS110 (Cmv) were used for these structural predictions since one Ecp32 family member from this isolate was present in all clusters, except for the orange one (Fig. 4). All of the structures had high pLDDT scores and TM-scores, indicative of highly confident predictions (Additional file 2: Table S5). As previously demonstrated for the Ecp32 proteins of *D. septosporum* and *F. fulva*, the Ecp32 family members from *C. minus* were predicted to be structurally similar to β-trefoil proteins, even though they had some differences in topology. The structures of Cmv937, Cmv6323, Cmv9109, Cmv5317, and Cmv2795 have two predicted disulphide bonds, while Cmv5316 has three.

Pairwise structural alignments between all of the predicted structures suggested that Cmv6323 (green group) and Cmv9109 (yellow group) are more similar to one another than any other structural pairwise alignment from the predicted structures for Cmv isolate NZFS110. Interestingly, Cmv6323 and Cmv9109 are part of the two groups that triggered the most consistent and strong cell death responses in *N. benthamiana*. This similarity, however, was not observed in the amino acid pairwise alignment of these proteins, for which the pair Cmv5316 and Cmv5317 have a higher percentage of identity than the Cmv6323 and Cmv9109 pair. The two predicted disulphide bonds shared across all predicted structures are conserved with the two predicted disulphide bonds present in all members of the Ecp32 family from *D. septosporum* and *F. fulva* (except Ds/FfEcp32-5, which does not have any cysteine residues), suggesting that this is a conserved feature in proteins belonging to this family (Additional file 1: Figure S5h) (Tarallo et al. 2022).

The structure of the truncated Ecp32 candidate Cms4060 was also predicted since the protein triggered weak cell death in the same manner as Cmv5316 (Fig. 5). Cms4060 has 77 amino acids less than the 226 expected for Cmv5316. The alignment between the predicted structures of Cms4060 and Cmv5316 (Additional file 1: Figure S5i) highlights the N-terminal conservation that was previously observed in the sequence alignment (Additional file 1: Figure S3). The predicted structure of Cms4060 (coloured green in the structural alignment) aligns with the region of Cmv5316 coloured in black, including one conserved disulphide bond (S-S1). The grey colouring represents the C-terminal region of Cmv5316 that is not present in Cms4060. Similarities in the structures of truncated protein Cms4060 and full-length Cmv5316, with conservation of cell death inducing activity, might provide insights into the regions of the proteins from the Ecp32 family that are responsible for plant receptor-mediated recognition.

Discussion

Cyclaneusma minus is an important pathogen of pines worldwide. However, very little is known about the biology of this pathogen. At first, differences in ascospore length and characteristics of the fungus grown on agar plates indicated that there were two morphological types of *C. minus* in New Zealand, termed *C. minus* 'simile' and *C. minus* 'verum' (Dick et al. 2001). However, with the development of molecular tools to differentiate these morphotypes, another morphotype was found to be present in the New Zealand *C. minus* population (Hunter et al. 2016). Of the 120 isolates tested in that study, only one isolate (NZFS3325) was found to differ from the simile and verum morphotypes. An additional isolate (NZFS3305) was subsequently identified (McDougal et al. 2016). Based on the symptoms observed on *Pinus radiata* needles, these morphotypes do not appear to differ in the symptoms that they cause, however this has not yet been studied in detail. Host genotype is known to strongly influence symptom development of CNC (Suontama et al. 2019; Ismael et al. 2020). With climate change and seasonal variation shifting, it becomes more important to understand the biology of these morphotypes, in order to enable mapping of likely distributions and elucidating their virulence and pathogenicity under future climatic conditions.

Through morphological and molecular analysis, we have confirmed the presence of a third *C. minus* morphotype in New Zealand, here classified as 'novus'. Growth temperature data suggested that the optimum is different for 'verum' compared to 'simile' and 'novus'. This in turn indicated that sporulation and germination may occur at different temperatures, which could influence the geographical distribution of these morphotypes, highlighting the importance of comparative studies between them. It has been proposed previously that 'verum' and 'simile' are likely different species, and that these should be formally described (Prihatini et al. 2014). The discovery of this new morphotype novus, that is different again to the other morphotypes, provides further impetus for the need of formal descriptions.

To help shed light on how genetic differences in morphotypes might affect *C. minus* virulence or pathogenicity, the genomes of isolates from three morphotypes of *C. minus* were sequenced. These are the first genomic resources

generated for *Cyclaneusma* and have enabled the prediction and comparison of CEs. The effector prediction pipeline generated a first list of possible effector proteins for *C. minus*. Because effectors are often described as SSCPs (Lo Presti et al. 2015), these features were used to identify *C. minus* CEs.

Some pathogens are known to secrete effectors that occur in families (Lo Presti et al. 2015; Denton-Giles et al. 2020; Rocafort et al. 2022). Following the observed similarity between DsEcp32-1 (part of the Ecp32 family from *D. septosporum*) and Cms835 (from isolate NZFS809), it was assessed to what extent the Ecp32 protein family was also present in this and the other seven isolates of *C. minus*. The results indicated that different morphotypes have different numbers of family members, with the morphotypes also showing differences in pseudogenisation of *Ecp32* genes. This was also observed in the Ecp32 family from *D. septosporum* and *D. pini* (both pine pathogens) and *F. fulva*, as they have four, three, and five family members respectively, with the presence of pseudogenes in some isolate genomes (Tarallo et al. 2022).

The diversification of gene families can occur due to tandem and/or segmental duplications (Leister 2004), which both seem to be present in the expansion of this particular effector family from *C. minus*. For example, the genes encoding proteins Cms4058, Cms4059, and Cms4060 (truncated) from NZFS809 are on the same scaffold and adjacent to one another. This is different from what was observed in *D. septosporum* and *F. fulva*, where the locations of the *Ecp32* genes are conserved in matching chromosomes between the two species but none of these family members are localized on the same chromosome (Mesarich et al. 2023). The differences in the number of family members and pseudogenes in each morphotype might reflect pathogen virulence and host range, as the selection pressure imposed by host *R* genes on plant pathogenic fungi can lead to an expansion and sequence divergence of an effector gene family (Pendleton et al. 2014; Franceschetti et al. 2017; Denton-Giles et al. 2020).

Also consistent with the predicted structures of Ecp32 proteins from *D. septosporum* and *F. fulva*, the predicted structures generated for representative members of the *C. minus* Ecp32 family were shown to adopt a β-trefoil fold. This is a fold present throughout different kingdoms that is involved in many different processes and interactions between organisms (Žurga et al. 2015). The most common types of proteins that share this fold are trypsin inhibitors and lectins, many being virulence factors in plant-pathogenic fungi (Schubert et al. 2012; Varrot et al. 2013; Juillot et al. 2016). It remains to be determined whether these β-trefoil proteins from *C. minus* have any virulence functions during infection. It would be interesting to determine the structures of other CEs from *C. minus* and identify to what extent this fold is present in the structures of any of those proteins. For example, the Ecp20 family, another cell death-eliciting family present in both *D. septosporum* and *F. fulva* (Tarallo et al. 2022), could not be detected in *C. minus* through primary sequence similarity searches. It remains possible, however, that searches in the predicted structures of all *C. minus* CEs might identify proteins that adopt the characteristic of Ecp20 family proteins, thereby expanding the repertoire of common effector types between these foliar fungal pathogens.

Proteins from the Ecp32 family in *C. minus* were found to have members that differed in their ability to trigger cell death in the non-host species *N. benthamiana* and *N. tabacum*. This difference was also observed in members of the Ecp32 family from *D. septosporum* and *F. fulva* (Tarallo et al. 2022). Whether the identified sequence variation contributes to the virulence functions of these proteins, as well as cell death elicitation, remains to be determined. Of note, other experiments, such as reactive oxygen species measurements and the analysis of defence marker gene expression in *Nicotiana* species, could be performed in the future to determine if these cell death responses are actually due to activation of the plant immune system. Similarly, the CE proteins should be transiently expressed in *BAK1*/*SOBIR1*-silenced or knock-out plants of *Nicotiana* species, to determine whether BAK1 and SOBIR1 (extracellular co-receptors involved in transducing defence response signals following apoplastic effector recognition) (Wang et al. 2018) are required for the elicitation of cell death by Ecp32 family members. The differences in cell death-triggering activity observed between different family members might be due to differences in the composition and positioning of amino acids on the surface on the proteins, for example as a result of insertions, deletions and/or substitutions of residues, which occur more often in loops than in α-helices and β-sheets (Fiser et al. 2000). In some cases, loops in tertiary structures of proteins are known to be involved in protein-protein interactions, recognition sites, signalling cascades, and ligand binding (Fetrow 1995), all of which could explain why some proteins from the *C. minus* Ecp32 family trigger more consistent cell death than others. Also, sequence and structural diversification observed for the proteins in this family could have a role in evasion of host immunity, another possible explanation for why this family has been expanded in *C. minus.*

Cms4060 (truncated) and Cmv5316 (full-length) triggered the same degree of cell death in *N. benthamiana*, but no cell death was observed in *N. tabacum*. This suggests that the N-terminal part of the protein (that is also conserved between Cms4060 and the other Ecp32

proteins) might contain the region, or conserved epitope, that is either recognized by a plant extracellular immune receptor in *N. benthamiana* or it interacts with plasma membrane proteins to trigger cell death. Previous studies have demonstrated the possibility of swapping loop regions from non-functional proteins for other loops from functional proteins (Wolfson et al. 1991; Pardon et al. 1995). Considering the consistent cell death triggered by Cmv9109, Cms9849, and Cmn3003, it would be informative to try to determine differences in the loop regions of these proteins that might be responsible for cell death, in comparison with other proteins that trigger inconsistent or no cell death. It would then be possible to determine if there are differences between the *C. minus* morphotypes in how these proteins function.

Conclusions

The generation of new genome sequences of eight *C. minus* isolates supported taxonomic classification of *C. minus* 'novus' as a third morphotype in the New Zealand population of this pathogen. It is important to understand the geographical range and variations in virulence and/or pathogenicity of these morphotypes, as well as the relationships between host genotype and fungal morphotypes, all of which have implications for diagnostics and biosecurity for the forest industry. The formal description of these morphotypes as separated species will be necessary to support forest biosecurity and disease management. It was also possible to identify CE proteins that might be the candidates for future investigation in the *C. minus*-pine interaction. We also showed that, just as in *D. septosporum* (another pine pathogen) and *F. fulva*, the Ecp32 family is present in *C. minus*, and the evidence suggests that some members from this family might be recognized by plant immune receptors. This, combined with the fact that these proteins are conserved at the primary sequence and tertiary structure levels, suggest that they have important and/or conserved roles in fungal virulence. Amino acid sequence searches revealed that the Ecp32 protein family is widespread across other fungal species (Tarallo et al. 2022) and, therefore, might be an important core effector of plant-pathogenic fungi. Furthermore, effector identification and analysis from the *Cyclaneusma* genomes has provided novel means to compare these morphotypes and demonstrated genomic-based differences that corroborate the hypothesis that these are in fact separate species. Disease resistance based on core effectors that are vital for a pathogen's ability to cause disease is more likely to be durable. Knowledge of core effectors from different *C. minus* morphotypes can enable the development of molecular tools that can both assist in better diagnostics and understanding of the pathogen population as well as complement breeding programmes.

Methods

Microorganisms and plants

Cyclaneusma minus isolates NZFS110, NZFS725, NZFS759, NZFS809, NZFS1800, NZFS3305, NZFS3276, and NZFS3325 (Table 1) were used in this study. Each was grown on a sterile cellophane sheet overlaying 2% (w/v) Difco™ Malt Extract Agar (Becton Dickinson & Company, New Jersey, USA) and incubated at 22°C in the dark. Once the mycelium filled half of the plate, it was scraped into a sterile 15 mL Falcon tube and stored at −80°C in preparation for freeze-drying and genomic DNA (gDNA) extraction.

Escherichia coli DH5α was used for CE gene cloning, and *A. tumefaciens* GV3101 was used for ATTAs. *Nicotiana tabacum* Wisconsin 38 and *N. benthamiana* were used as model non-host plants for ATTAs. Plants were grown in a temperature-controlled room at 22°C and 12/12 light/dark photoperiod with 80–85 μmol/m^2/s.

Colony morphology analysis

For colony morphology assessments, three 5 mm diameter plugs from the edge of an actively growing culture of each *C. minus* isolate were placed in the middle of 2% and 9% MEA plates. Plates were incubated as above and assessed for colony morphology at eight weeks post-inoculation using a standard colour chart used to describe colours in colonies from Rayner (1970). For the temperature study, a 5 mm diameter plug from each isolate was placed onto an independent 2% MEA plate as before, each with three replicates, and incubated at 4, 10, 17, 20, 22, 25, 28, or 30 °C in the dark. Two measurements at right angles for each replicate were taken at 14 days post-inoculation. The mean and standard deviation of the radial growth rate per day for each isolate were then calculated for each incubation temperature. Differences between radial growth per day at different temperatures were analysed using a two way ANOVA, containing temperature (as a character) and morphotype. Significant terms were followed up by applying a multiple-comparison procedure based on Tukey-adjusted contrast. The cardinal temperatures for each morphotype were determined from these data.

Genome sequencing, assembly, and annotation

Total gDNA for genome sequencing was extracted from freeze-dried mycelia of *C. minus* isolates using a DNeasy Plant Maxi Kit (Qiagen, USA), according to the manufacturer's instructions. For each isolate, 200 mg of tissue was added to Lysing Matrix C tubes containing 1.0 mm silica spheres (MP Biomedical, California, USA) and homogenised to a fine powder using a FastPrep-24™ Instrument (MP Biomedical) with two cycles of 20 s at 5 m/sec. The extracted gDNA was

concentrated by ethanol precipitation and resuspended in TE buffer (10 mM Tris and 1 mM EDTA, pH 8.0). The gDNA concentrations were quantified by spectrophotometry using both the Denovix DS-11 FX (DeNovix, Delaware, USA) and the DeNovix dsDNA Broad Range Assay (DeNovix, Delaware, USA).

The gDNA from eight isolates of *C. minus* (Table 1) was used for whole-genome sequencing using Illumina short-read sequencing technology (Illumina, California, USA). Paired-end libraries with insert sizes of 327–353 bp were sequenced on a NovaSeq 6000 system (Illumina, California, USA).

De novo genome assemblies of filtered reads were generated using SPAdes v3.14.0 (Nurk et al. 2013) with default settings. Contigs were assembled into scaffolds with redundans v0.13c (Pryszcz and Gabaldón 2016). Genome assembly quality was measured in QUAST (Gurevich et al. 2013). Assembly completeness was estimated with Benchmarking Universal Single-Copy Orthologs (BUSCO) comparisons to the fungal dataset (Simão et al. 2015).

Gene annotations were predicted in each masked genome assembly using the Funannotate v1.8.9 pipeline (Palmer and Stajich 2020). Gene prediction was carried out using predicted protein sequences of *Elytroderma deformans* CBS183.68 (available at: https://mycocosm.jgi.doe.gov/Elyde1/Elyde1.home.html) as a guide. At first, protein sequences of *E. deformans* were mapped to the *C. minus* assembly file and the orthologs from *Cyclaneusma* were extracted. These orthologous sequences were used to train Augustus v3.3.3 (Stanke et al. 2004), followed by three rounds of iterative gene prediction. Genes were predicted by Augustus v3.3.3, GeneMark ES v4.68 (Ter-Hovhannisyan et al. 2008), snap 2006-07-28 (Korf 2004), and GlimmerHMM v3.0.4 (Majoros et al. 2004), while transfer RNAs were predicted by tRNAscan-SE v2.0.9 (Chan et al. 2021). EVidenceModeler v1.1.1 (Haas et al. 2008) was used to generate consensus gene models. Putative functions of gene products were determined by searches to the MEROPS v12.0 (Rawlings et al. 2018), UniProt DB 2021_03 (The UniProt 2021), dbCAN v10.0 (Yin et al. 2012), PFAM v34.0 (Mistry et al. 2021), and BUSCO v2.0 (Simão et al. 2015) databases.

Phylogenetic trees

The phylogenetic positions of the *C. minus* strains were assessed by constructing phylogenetic trees using concatenated nucleotide sequences from the ITS of ribosomal DNA, nuclear LSU (nLSU), and translation elongation factor 1α (tef-1). An initial alignment of these barcode genes from *Lophodermium conigenum* NZFS790 to the assembled genomes was done using LASTZ (Harris 2007) and Geneious R10 (Kearse et al. 2012) to find the *C. minus* orthologs. Orthologs were then extracted and aligned to DNA sequences from other *Cyclaneusma* species, as well as *Lophodermium* and *Strasseria* species (Additional file 2: Table S6). The alignment for the tree was generated using a cost-matrix of 51% with pairwise alignment or multiple alignment gap opening penalty of 12, plus a gap extension penalty of 0.3. The DNA sequences were aligned in Geneious R10. Nucleotide sequences not aligned at the 5' and 3' extremities of alignments were trimmed, and internal gaps maintained. The multi-gene phylogenetic trees were created by a maximum likelihood algorithm with 1000 replicates and distances obtained using the General Time Reversible model using Mega X (Kumar et al. 2018).

Effector prediction

Following genome assembly and annotation of the eight *C. minus* isolates, a previously described pipeline (Hunziker et al. 2021) was used to identify CE proteins from all genomes, based on their full sets of predicted proteins. Briefly, signal peptides were predicted using SignalP v3.0 (Bendtsen et al. 2004). To exclude membrane-bound proteins, TMHMM v2.0 (Krogh et al. 2001) and BIG-PI (Eisenhaber et al. 2004) were used. Predicted proteins of ≤ 300 amino acids in length with ≥ 4 cysteine residues were then selected and classified as small, secreted cysteine-rich proteins (SSCPs). EffectorP v3.0 (Sperschneider and Dodds 2021) was then used to select proteins classified as apoplastic or cytoplasmic effectors, with candidates that were classified as non-effectors being discarded.

The amino acid sequence of DsEcp32-1, a member of the Ecp32 effector family from *D. septosporum* (Tarallo et al. 2022), was firstly used to identify the likely orthologue in *C. minus* isolate NZFS809. This *C. minus* sequence was then used to identify possible Ecp32 family homologues in the other seven isolates of *C. minus*, through reciprocal BLASTp and tBLASTn searches in conjunction with an E-value cut-off of $< 10^{-5}$.

Protein sequence alignments were performed using Clustal Ω (Sievers et al. 2011) in Jalview (Waterhouse et al. 2009). Phylogenetic trees were constructed with the neighbour-joining method based on 1000 bootstrap replicates using Geneious Software v9.1.8 (Kearse et al. 2012) in conjunction with CE protein sequence alignments. IDRs were predicted by querying protein sequences in the Predictor of Natural Disordered Regions (PONDR® VLXT) server (Romero et al. 2001). A prediction score ≥ 0.8 was considered a significant IDR prediction. Protein tertiary structure predictions of Ecp32 family members were performed using AlphaFold2 in conjunction with ColabFold (Jumper et al. 2021; Mirdita et al. 2022). Two values were used to assess the confidence of the predictions: the predicted local-distance difference test (pLDDT) score, which ranges from 0 to 100,

with values closer to 100 indicating a high confidence in prediction, and the global superposition metric template modelling score (TM-score), ranging from 0 to 1, with a value of ≥0.5 indicating a similar fold between structures. The Dali server (Holm 2020) was used to identify proteins with structural similarity in the Research Collaboratory for Structural Bioinformatics Protein Data Bank. Here, a Dali Z-score of ≥2 was used to infer structural similarity. Protein tertiary structures were visualized and rendered in PyMOL v2.5, with alignments carried out using the CEalign tool (DeLano 2002).

Agrobacterium tumefaciens-mediated transient expression assays (ATTAs)

Cyclaneusma minus ATTA expression vectors were generated using the method described by Guo et al. (2020). The expression vector used was pICH86988, which contains a *N. tabacum* N-terminal PR1α signal sequence for apoplast secretion followed by an N-terminal 3xFLAG tag for western blotting detection (Weber et al. 2011). CE gene sequences of *C. minus*, excluding the nucleotide sequence encoding the predicted signal peptide, were synthesized by Twist Bioscience (California, USA), and included BsaI recognition sites and 4 bp overhangs specific for Golden Gate modular cloning (Engler et al. 2009). CE genes were also cloned into a version of the vector that lacked the sequence encoding the PR1α signal peptide, but that contained only a start codon. Constructs were transformed into competent cells of *E. coli* by electroporation and plasmids were extracted using an E.Z.N.A.® Plasmid DNA Mini Kit I (Omega Bio-tek, Georgia, USA). Correct assemblies were confirmed by sequencing in association with the Massey Genome Service (Palmerston North, New Zealand). Expression vectors were then transformed into competent cells of *A. tumefaciens* by electroporation (Guo et al. 2020).

To perform the ATTAs, transformed cells of *A. tumefaciens* containing the CE genes were inoculated into lysogeny broth supplemented with 50 µg/mL kanamycin, 10 µg/mL rifampicin, and 30 µg/mL gentamycin and grown overnight at 28°C in the dark until an OD_{600} of 0.4–2.0 was reached. Cells were collected by centrifugation at 2500×*g* for 5 min and resuspended in 1 mL of infiltration buffer (10 mM $MgCl_2$:$6H_2O$, 10 mM MES-KOH, 100 µM acetosyringone (Sigma-Aldrich, Missouri, USA). The OD_{600} was measured again, and sterile water added to reach a final OD_{600} of 0.5. These cultures were then incubated for 3 h at room temperature before infiltration. An *A. tumefaciens* transformant carrying an expression vector for the extracellular elicitin INF1, from *Phytophthora infestans* (Kamoun et al. 1997), was used as a positive control and the empty pICH86988 vector was used as a negative control. The abaxial surface of *N. benthamiana* and *N. tabacum* leaves was infiltrated with the appropriate *A. tumefaciens* solution using 1 mL needleless syringes. The screening was done with at least three biological replicates. Each biological replicate consisted of two plants, each with two leaves infiltrated, resulting in at least 12 infiltration zones. Plants were monitored for up to seven days, at which time photographs were taken with a Nikon D7000 camera. To assess the plant response caused by each expressed protein, infiltration zones were divided into a strong or weak cell death response, chlorosis, or no cell death. A strong cell death response was considered when it was indistinguishable from the response elicited by INF1 and covered the entire infiltration zone, while a weak response was considered when the protein elicited a response across approximately 50% of the infiltration zone when compared with INF1. Chlorosis was considered when leaf discoloration occurred, but no apparent cell death was observed. A no-cell death response was indicated when the infiltration zone had the same lack of a response as the negative empty pICH86988 vector control.

Abbreviations

ATTAs	*Agrobacterium tumefaciens*-mediated transient transformation assays
bp	Base pairs
CE	Candidate effector
Cmn	*Cyclaneusma minus* morphotype 'novus'
Cms	*Cyclaneusma minus* morphotype 'simile'
Cmv	*Cyclaneusma minus* morphotype 'verum'
CNC	Cyclaneusma needle cast
gDNA	Genomic DNA
IDR	Intrinsically disordered region
ITS	Internal transcribed spacer
R	Resistance
SSCPs	Small, secreted cysteine-rich proteins
tef-1α	Translation elongation factor 1α

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42483-024-00255-8.

Additional file 1: Figure S1. Multi-locus phylogeny of *Cyclaneusma* species and other fungal associates of *Pinus radiata* showing the placement of the three *Cyclaneusma* morphotypes. **Figure S2.** Protein phylogeny of the Ecp32 families from *D. septosporum*, *F. fulva*, and *C. minus* NZFS110. **Figure S3.** Protein sequence alignment of Ecp32 family members across eight isolates and three morphotypes of *C. minus*. **Figure S4.** Ecp32 proteins from different morphotypes of *C. minus* triggered cell death in *Nicotiana tabacum*. **Figure S5.** Predicted protein tertiary structures of Ecp32 family members from different isolates of *C. minus*.

Additional file 2: Table S1. Descriptions of colony morphology of 'verum', 'smile', and 'novus' morphotypes on 2% and 9% MEA at 22°C in the dark after 8 weeks. **Table S2.** *C. minus* isolates' genome sequencing results. **Table S3.** Protein sequence for the predicted candidate effectors from *C. minus* isolates NZFS110, 725, 1800, 3276 (morphotype 'verum'), NZFS759, 809 (morphotype 'simile'), and NZFS3305, 3325 (morphotype 'novus'). **Table S4.** Amino acid sequences of *C. minus* Ecp32 family members in eight isolates. **Table S5.** Predicted local-distance difference test (pLDDT) score and global superposition metric template modelling score (TM-score) for *Cyclaneusma minus* Ecp32 proteins structural predictions. **Table S6.** Accessions for phylogenetic tree used in Additional file 1: Figure S1.

Acknowledgements
The authors thank Rita Tetenburg and Sara Carey for technical assistance with culture collections, Dale Corbett for help with graphic design and Emily McLay for statistical advice (Scion, Rotorua, New Zealand). The authors acknowledge Beccy Ganley for useful discussions and strain collections from past projects. Isolates are maintained by the Forest Health Reference Laboratory Culture Collection (NZFS) at Scion, New Zealand.

Authors' contributions
MT completed the effector analysis and wrote the manuscript. KD performed the culture-based analysis (with assistance from KG) and contributed substantially to the manuscript. LL, TW, and DS performed the DNA extraction and genome sequencing analysis. RM, RB, and CM conceived and supervised the study, made substantial contributions to the interpretation of the results, and contributed to writing and revision of the manuscript. All authors contributed to the article and approved the submitted version.

Funding
This research was funded by Scion (New Zealand Forest Research Institute, Ltd., Rotorua, New Zealand) through the Resilient Forests Research Program via Strategic Science Investment Funding from the New Zealand Ministry of Business Innovation and Employment (MBIE) (CO4X1703), and New Zealand Forest Grower Levy Trust funding (QT-9108).

Availability of data and materials
The sequences were deposited in NCBI under the accession numbers outlined in Table 2.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 2 June 2023 Accepted: 20 May 2024
Published online: 30 July 2024

References

Bednářová M, Milon D, Joseph Janoušek, Jankovský L. Other foliar diseases of coniferous trees. In: Gonthier P, Nicolotti G, editors. Infectious forest diseases. CAB International; 2013. pp. 458–87. https://doi.org/10.1079/9781780640402.0458.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol. 2004;340(4):783–95. https://doi.org/10.1016/j.jmb.2004.05.028.

Bulman LS. Cyclaneusma needle-cast and Dothistroma needle blight in NZ pine plantations. N Z J for. 1993;38(2):21–4.

Bulman LS. Economic impact of Cyclaneusma needle cast in New Zealand. In: Client report FBRC and FRST. 2009. http://www.nzfoa.org.nz/images/stories/pdfs/content/needlecast/cyclaneusma_economic_loss_final_may_2009.pdf. Accessed 15 Oct 2022.

Chan PP, Lin BY, Mak AJ, Lowe TM. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. 2021;49(16):9077–96. https://doi.org/10.1093/nar/gkab688.

Cook DE, Mesarich CH, Thomma BPHJ. Understanding plant immunity as a surveillance system to detect invasion. Annu Rev Phytopathol. 2015;53(1):541–63. https://doi.org/10.1146/annurev-phyto-080614-120114.

Crosby T, Dugdale J, Watt J, Crosby TK, Dugdale JS, Watt JC. Area codes for recording specimen localities in the New Zealand subregion. N Z J Zool. 1998;25:175–83. https://doi.org/10.1080/03014223.1998.9518148.

DeLano WL. Pymol: an open-source molecular graphics tool. CCP4 Newsl Protein Crystallogr. 2002;40(1):82–92.

Denton-Giles M, McCarthy H, Sehrish T, Dijkwel Y, Mesarich CH, Bradshaw RE, et al. Conservation and expansion of a necrosis-inducing small secreted protein family from host-variable phytopathogens of the Sclerotiniaceae. Mol Plant Pathol. 2020;21(4):512–26. https://doi.org/10.1111/mpp.12913.

Dick MA, Somerville JG, Gadgil PD. Variability in the fungal population. In: Bulman LS, Gadgil PD, editors. Cyclaneusma needle-cast in New Zealand, Forest Research bulletin no 222. Rotorua Forest Research Institute Ltd.; 2001.

Dicosmo F, Peredo H, Minter DW. *Cyclaneusma* gen. nov., *Naemacyclus* and *Lasiostictis*, a nomenclatural problem resolved. Eur J Pathol. 1983;13(4):206–12. https://doi.org/10.1111/j.1439-0329.1983.tb00119.x.

Eisenhaber B, Schneider G, Wildpaner M, Eisenhaber F. A sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans*, *Candida albicans*, *Neurospora crassa*, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. J Mol Biol. 2004;337(2):243–53. https://doi.org/10.1016/j.jmb.2004.01.025.

Engler C, Gruetzner R, Kandzia R, Marillonnet S. Golden Gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS ONE. 2009;4(5). https://doi.org/10.1371/journal.pone.0005553.

Fetrow JS. Omega loops: nonregular secondary structures significant in protein function and stability. FASEB J. 1995;9(9):708–17. https://doi.org/10.1096/fasebj.9.9.7601335.

Fiser A, Do RKG, Šali A. Modeling of loops in protein structures. Protein Sci. 2000;9(9):1753–73. https://doi.org/10.1110/ps.9.9.1753.

Franceschetti M, Maqbool A, Jiménez-Dalmaroni MJ, Pennington HG, Kamoun S, Banfield MJ. Effectors of filamentous plant pathogens: commonalities amid diversity. Microbiol Mol Biol Rev. 2017;81(2). https://doi.org/10.1128/MMBR.00066-16.

Gilmour JW. Pathogenic needle-cast fungi. Interim Research Release No8. New Zealand Forest Service, Forest Research Institute; 1959.

Guo Y, Hunziker L, Mesarich CH, Chettri P, Dupont PY, Ganley RJ, et al. DsEcp2-1 is a polymorphic effector that restricts growth of *Dothistroma Septosporum* in pine. Fungal Genet Biol. 2020;135. https://doi.org/10.1016/j.fgb.2019.103300.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072–5. https://doi.org/10.1093/bioinformatics/btt086.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, et al. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. Genome Biol. 2008;9(1). https://doi.org/10.1186/gb-2008-9-1-r7.

Harris RS. Improved pairwise alignment of genomic DNA. The Pennsylvania State University; 2007.

Holm L. Using Dali for protein structure comparison. In: Gáspári Z, editor. Structural bioinformatics: methods and protocols. New York: Springer; 2020. pp. 29–42. https://doi.org/10.1007/978-1-0716-0270-6_3.

Hunter S, Glen M, McDougal R. Molecular tools for differentiating *Cyclaneusma minus* morphotypes and assessing their distribution in *Pinus radiata* forests in New Zealand. N Z J Sci. 2016;46(1):21. https://doi.org/10.1186/s40490-016-0080-0.

Hunziker L, Tarallo M, Gough K, Guo M, Hargreaves C, Loo TS, et al. Apoplastic effector candidates of a foliar forest pathogen trigger cell death in host and non-host plants. Sci Rep. 2021;11(1). https://doi.org/10.1038/s41598-021-99415-5.

Ismael A, Suontama M, Klápště J, Kennedy S, Graham N, Telfer E, et al. Indication of quantitative multiple disease resistance to foliar pathogens in *Pinus radiata* D.Don in New Zealand. Front Plant Sci. 2020;11:1044. https://doi.org/10.3389/fpls.2020.01044.

Jones JDG, Dangl JL. The plant immune system. Nat. 2006;444(7117):323–9. https://doi.org/10.1038/nature05286.

Juillot S, Cott C, Madl J, Claudinon J, van der Velden NSJ, Künzler M, et al. Uptake of *Marasmius oreades* agglutinin disrupts integrin-dependent cell adhesion. Biochim Biophys Acta. 2016;1860(2):392–01. https://doi.org/10.1016/j.bbagen.2015.11.002.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nat. 2021;596(7873):583–9. https://doi.org/10.1038/s41586-021-03819-2.

Kamoun S, van West P, de Jong AJ, de Groot KE, Vleeshouwers VGAA, Govers F. A gene encoding a protein elicitor of *Phytophthora infestans* is down-regulated during infection of potato. Mol Plant-Microbe Interact. 1997;10(1):13–20. https://doi.org/10.1094/MPMI.1997.10.1.13.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28(12):1647–9. https://doi.org/10.1093/bioinformatics/bts199.

Korf I. Gene finding in novel genomes. BMC Bioinf. 2004;5(1):59. https://doi.org/10.1186/1471-2105-5-59.

Krogh A, Larsson B, Von Heijne G, Sonnhammer E. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305(3):567–80. https://doi.org/10.1006/jmbi.2000.4315.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9. https://doi.org/10.1093/molbev/msy096.

Leister D. Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. Trends Genet. 2004;20(3):116–22. https://doi.org/10.1016/j.tig.2004.01.007.

Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, et al. Fungal effectors and plant susceptibility. Annu Rev Plant Biol. 2015;66:513–45. https://doi.org/10.1146/annurev-arplant-043014-114623.

Majoros WH, Pertea M, Salzberg SL. TigrScan and GlimmerHMM: two open source *ab initio* eukaryotic gene-finders. Bioinformatics. 2004;20(16):2878–9. https://doi.org/10.1093/bioinformatics/bth315.

McDougal R, Hunter S, Low C. Validation of qPCR assays for *Cyclaneusma minus* simile and verum and assessment with *Pinus radiata* of varying susceptibility. In Needle Disease Strategy Research Programme Report - Rotorua, New Zealand, Scion (New Zealand Forest Research Institute, Ltd.). 2016. https://fgr.nz/documents/download/1118. Accessed 29 April 2024.

Merrill W, Wenner N. Cyclaneusma needle-cast and needle retention in scots pine. Plant Dis. 1996;80(3):294–8.

Mesarich CH, Ökmen B, Rovenich H, Griffiths SA, Wang C, Karimi Jashni M, et al. Specific hypersensitive response–associated recognition of new apoplastic effectors from *Cladosporium Fulvum* in wild tomato. Mol Plant-Microbe Interact. 2018;31(1):145–62. https://doi.org/10.1094/MPMI-05-17-0114-FI.

Mesarich CH, Barnes I, Bradley EL, de la Rosa S, de Wit PJGM, Guo Y, et al. Beyond the genomes of *Fulvia Fulva* (syn. *Cladosporium Fulvum*) and *Dothistroma Septosporum*: new insights into how these fungal pathogens interact with their host plants. Mol Plant Pathol. 2023;24(5):474–94. https://doi.org/10.1111/mpp.13309.

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods. 2022;19:679–82. https://doi.org/10.1038/s41592-022-01488-1.

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar Gustavo A, Sonnhammer ELL, et al. Pfam: the protein families database in 2021. Nucleic Acids Res. 2021;49(1):412–9. https://doi.org/10.1093/nar/gkaa913.

Müller M, Kurkela T, editors. Do *Sphaeropsis sapinea* and *Naemacyclus minor* exist in Finland? In: Seminar on Forest Pathology. Finnish Forest Research Institute, Vantaa Research Unit. 2012. http://www.metla.fi/tapahtumat/2012/pathology/Muller-abstract.pdf. Acessed 15 Oct 2022.

Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, et al. Assembling genomes and mini-metagenomes from highly chimeric reads. In: Deng M, Jiang R, Sun F, Zhang X, editors. Research in Computational Molecular Biology. Berlin: Springer; 2013. pp. 158–70. https://doi.org/10.1007/978-3-642-37195-0_13.

Palmer JM, Stajich J. Funannotate v1.8.1: Eukaryotic genome annotation. 2020. https://doi.org/10.5281/zenodo.4054262.

Pardon E, Haezebrouck P, De Baetselier A, Hooke SD, Fancourt KT, Desmet J, et al. A Ca^{2+}-binding chimera of human lLysozyme and bovine α-Lactalbumin that can form a molten globule. J Biol Chem. 1995;270(18):10514–24. https://doi.org/10.1074/jbc.270.18.10514.

Pendleton AL, Smith KE, Feau N, Martin FM, Grigoriev IV, Hamelin R, et al. Duplications and losses in gene families of rust pathogens highlight putative effectors. Front Plant Sci. 2014;5(299). https://doi.org/10.3389/fpls.2014.00299.

Prihatini I, Glen M, Wardlaw TJ, Mohammed CL. Multigene phylogenetic study of *Cyclaneusma species*. Eur J Pathol. 2014;44(4):299–09. https://doi.org/10.1111/efp.12101.

Pryszcz LP, Gabaldón T. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Res. 2016;44(12). https://doi.org/10.1093/nar/gkw294.

Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res. 2018;46(1):624–32. https://doi.org/10.1093/nar/gkx1134.

Rayner RW. A mycological colour chart. Commonwealth Mycological Institute, Kew & British Mycological Society; 1970.

Rocafort M, Fudal I, Mesarich CH. Apoplastic effector proteins of plant-associated fungi and oomycetes. Curr Opin Plant Biol. 2020;56:9–19. https://doi.org/10.1016/j.pbi.2020.02.004.

Rocafort M, Bowen JK, Hassing B, Cox MP, McGreal B, de la Rosa S, et al. The *Venturia inaequalis* effector repertoire is dominated by expanded families with predicted structural similarity, but unrelated sequence, to avirulence proteins from other plant-pathogenic fungi. BMC Biol. 2022;20(1):246. https://doi.org/10.1186/s12915-022-01442-9.

Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. Proteins. 2001;42(1):38–48. https://doi.org/10.1002/1097-0134(20010101)42:1.

Schubert M, Bleuler-Martinez S, Butschi A, Wälti MA, Egloff P, Stutz K, et al. Plasticity of the β-trefoil protein fold in the recognition and control of invertebrate predators and parasites by a fungal defence system. PLoS Pathog. 2012;8(5). https://doi.org/10.1371/journal.ppat.1002706.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539. https://doi.org/10.1038/msb.2011.75.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210–2. https://doi.org/10.1093/bioinformatics/btv351.

Sperschneider J, Dodds PN. EffectorP 3.0: prediction of apoplastic and cytoplasmic effectors in fungi and oomycetes. Mol Plant-Microbe Interact. 2021;35(2):146–56. https://doi.org/10.1094/MPMI-08-21-0201-R.

Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. 2004;32:309–12. https://doi.org/10.1093/nar/gkh379.

Suontama M, Li Y, Low CB, Dungey HS. Genetic improvement of resistance to Cyclaneusma needle-cast in *Pinus radiata*. Can J Res. 2019;49(2):128–33. https://doi.org/10.1139/cjfr-2018-0075.

Tarallo M, McDougal RL, Chen Z, Wang Y, Bradshaw RE, Mesarich CH. Characterization of two conserved cell death elicitor families from the Dothideomycete fungal pathogens *Dothistroma septosporum* and *Fulvia fulva* (syn. *Cladosporium fulvum*). Front Microbiol. 2022;13. https://doi.org/10.3389/fmicb.2022.964851.

Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal genomes using *an ab* initio algorithm with unsupervised training. Genome Res. 2008;18(12):1979–90. https://doi.org/10.1101/gr.081612.108.

The UniProt C. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 2021;49(1):480–9. https://doi.org/10.1093/nar/gkaa1100.

Varrot A, Basheer SM, Imberty A. Fungal lectins: structure, function and potential applications. Curr Opin Struct Biol. 2013;23(5):678–85. https://doi.org/10.1016/j.sbi.2013.07.007.

Von Butin H. Morphological and taxonomic studies on *Naemacyclus Niveus* (Pers. Ex Fr.) Fuck. Ex Sacc. And related species. Eur J Pathol. 1973;3(3):146–63. https://doi.org/10.1111/j.1439-0329.1973.tb00389.x.

Wang Y, Xu Y, Sun Y, Wang H, Qi J, Wan B, et al. Leucine-rich repeat receptor-like gene screen reveals that *Nicotiana* RXEG1 regulates glycoside hydrolase 12 MAMP detection. Nat Commun. 2018;9. https://doi.org/10.1038/s41467-018-03010-8.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview version 2 - a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189–91. https://doi.org/10.1093/bioinformatics/btp033.

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS ONE. 2011;6(2). https://doi.org/10.1371/journal.pone.0016765.

Win J, Chaparro-Garcia A, Belhaj K, Saunders DG, Yoshida K, Dong S, et al. Effector biology of plant-associated organisms: concepts and perspectives. Cold Spring Harbor Symp Quant Biol. 2012;77:235–47. https://doi.org/10.1101/sqb.2012.77.015933.

Wolfson AJ, Kanaoka M, Lau FTK, Ringe D. Insertion of an elastase-binding loop into interleukin-1β. Protein Eng Des Sel. 1991;4(3):313–7. https://doi.org/10.1093/protein/4.3.313.

Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012;40(1):445–51. https://doi.org/10.1093/nar/gks479.

Žurga S, Pohleven J, Kos J, Sabotič J. β-Trefoil structure enables interactions between lectins and protease inhibitors that regulate their biological functions. J Biochem. 2015;158(1):83–90. https://doi.org/10.1093/jb/mvv025.