Effector identification in the pine pathogen *Dothistroma septosporum*

L. Hunziker¹, C.H. Mesarich², R.L. McDougal³ and R.E. Bradshaw¹

¹*Institute of Fundamental Sciences, Massey University, Palmerston North 4442, New Zealand*

²*Institute of Agriculture and Environment, Massey University, Palmerston North 4442, New Zealand*

³*Scion (New Zealand Forest Research Institute, Ltd.), 49 Sala Street, Rotorua 3046, New Zealand*

Corresponding author: r.e.bradshaw@massey.ac.nz

Abstract Effectors are molecules produced by plant-associated organisms, such as fungal pathogens, that facilitate infection of specific plant hosts. Some effectors directly or indirectly suppress host defence responses while others help to mask the foreign invader from recognition by the plant. Plants can also recognise and respond to specific effectors to mount a resistance response. Thus the study of effectors is important for an understanding of plant–pathogen interactions. *Dothistroma septosporum* causes Dothistroma needle blight (DNB), a conifer disease of worldwide importance. In this study, effector genes specific to the *D. septosporum*–*Pinus radiata* interaction were identified. A bioinformatics pipeline was employed to predict effector candidates on the basis of characteristics such as secretion from the fungus and expression *in planta*. Using this pipeline, 55 effector candidates for future analysis were identified.

Keywords bioinformatics pipeline, forest pathogen, *Pinus radiata*, secreted proteins, *in planta* expression.

INTRODUCTION

Dothistroma septosporum is a hemibiotrophic fungal foliar pathogen and the main cause of Dothistroma needle blight (DNB). DNB occurs in at least 82 Pinaceae hosts in 63 countries (Barnes et al. 2014) and can cause substantial losses in both plantation and native forests. An important recent trend in pathology research involves the identification and functional characterization of effectors, molecules secreted by the pathogen that promote

host colonization. Some effectors can be recognised by plant hosts with corresponding receptors and trigger host defence responses. Thus, effectors have been used successfully in effector-assisted breeding for resistance in various crops (Vleeshouwers & Oliver 2014). Comparatively little is known about effectors in forest pathogens, but the availability of the genome sequence of *D. septosporum* (de Wit et al. 2012) provides a way to identify predicted

Figure 1 Pipeline for the selection of candidate effectors. Numbers on the left refer to the steps outlined in the methods; numbers on the right indicate the number of genes after each step. SSCP, small secreted cysteine-rich proteins; SSP, small secreted proteins; LSP, large secreted proteins; CAZY, carbohydrate-active enzymes.

effector proteins in this pathogen. A preliminary analysis showed that *D. septosporum* has some effectors similar to those of its close relative, the tomato pathogen *Cladosporium fulvum*, and that these effectors, like their counterparts in *C. fulvum*, trigger defence responses in tomato plants carrying corresponding receptors (de Wit et al. 2012). This suggested they might have a similar function in their pine host. However, a gene expression study suggested that most of these effectors are produced at only low levels by *D. septosporum* in its host *Pinus radiata* (Bradshaw et al. 2016). Thus the aim of this study was to identify effector candidates in *D. septosporum* that are highly expressed during infection of pine needles, using a pipeline of genome analysis tools similar to those used for other fungi.

MATERIALS AND METHODS

A bioinformatics pipeline (Haddadi et al. 2016) was adapted to identify candidate effector genes from the *D. septosporum* genome that are highly expressed *in planta* (Figure 1): (1) The genome sequence, as well as gene models and predicted protein sequences (de Wit et al. 2012), were downloaded from the Joint Genome Institute

(JGI; <http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html>); (2) Effectors are secreted, so signal peptides that direct proteins through the classical fungal secretory pathway were predicted using the SignalP programme (Bendtsen et al. 2004); (3) Some integral membrane proteins, membrane-bound proteins, and various intracellular organelle proteins, are also known to possess a signal peptide, but are not effectors. Thus, proteins with a putative transmembrane domain, glycosylphosphatidylinositol (GPI) anchor site, or a predicted organelle localization profile, as predicted by the TMHMM (Krogh et al. 2001), BIG-PI (Eisenhaber et al. 2004), and TargetP (Emanuelsson et al. 2000) programmes respectively, were discarded; (4) Gene expression data from a *D. septosporum*–*P. radiata* transcriptome study (Bradshaw et al. 2016) were used to select effector candidates expressed above an arbitrary *in planta* threshold of 50 reads per million per kilobase (RPMK; a normalised expression value showing the abundance of RNA for each gene); (5) Effector candidates were then divided into four types commonly found to have effector activity in fungi (Sperschneider et al. 2015): Small (≤ 300 amino acids) secreted cysteine-rich (≥ 4) proteins (SSCPs), proteins with predicted carbohydrate-active enzyme activity (CAZys), and small and large secreted proteins (SSPs, LSPs); (6) The SSCP type, most prominent and well-studied among fungal effectors (Sperschneider et al. 2015), was further assessed for high or low expression at specific *in planta* stages of infection compared with *in vitro* culture at 3 weeks (early, biotrophic stage), 8 weeks (mid, necrotrophic), and 12 weeks (late, necrotrophic) post-inoculation. Those significantly up-regulated ($q < 0.05$; Storey (2003)) at least two-fold in the early stage were selected as biotrophic, and those in mid- or late-stage as necrotrophic effector candidates; (7) Other candidates were selected from SSPs, LSPs and CAZys if they had homology to known effectors or very high *in planta* expression (≥ 500 RPMK).

Table 1 Examples of selected *Dothistroma septosporum* effector candidates.

Protein ID ¹	<i>In vitro</i> ²	Early ²	Mid ²	Late ²	Type/annotation ³
Ds72737	41	491	245	8995	SSCP/C-type lectin
Ds70155	267	71	507	1303	SSCP/cerato-platanin
Ds68958	1595	182	753	3500	SSCP/unknown
Ds131290	399	637	89	174	SSCP/LysM
Ds62617	67	2608	137	16	CAZy/GH64
Ds75147	14	153	42	670	CAZy/GH13
Ds74062	1714	6386	390	1369	LSP/unknown
Ds73520	8	819	438	5421	SSP/unknown

¹*D. septosporum* protein identification number in the Joint Genome Institute database.

²Normalised expression of gene (reads per million per kilobase), for *in vitro* cultures or three different stages of *Pinus radiata* infection (Bradshaw et al, 2016).

³SSCP, small secreted cysteine-rich protein; CAZy, carbohydrate-active enzyme; LSP, SSP, large/small secreted protein; GH, glycosyl hydrolase.

RESULTS AND DISCUSSION

Of the predicted proteins in the *D. septosporum* genome, 1487 were predicted to have a signal peptide (Figure 1, steps 1 and 2). Of these, only 875 were predicted to be secreted (step 3) and 397 showed expression *in planta* above the arbitrary threshold (step 4) suggesting their potential to function in the *D. septosporum*–*P. radiata* interaction. Finally, 33 SSCPs showed significant up-regulation in at least one *in planta* stage compared with *in vitro* (step 6), and 22 other secreted proteins with sequence features of effectors such as LysM domains or glycoside hydrolase activity (O’Connell et al. 2012) or very high expression were also identified (step 7). A representative set of eight effector candidates, along with their annotations (predicted functions) and gene expression profiles, are shown in Table 1.

The effector candidates included some with predicted functions linked to virulence. Most highly expressed *in planta* was Ds72737, peaking at 222-fold up-regulation compared with *in vitro* and predicted to have a lectin domain. Lectins have high binding affinity for carbohydrates and can bind fungal cell wall chitin to mask the fungus from detection by the plant, as shown for the *C. fulvum* LysM effector Ecp6 (Bolton et al. 2008). Another candidate,

Ds70155, was a predicted cerato-platanin. Fungal cerato-platanin proteins have been suggested to be virulence factors of pathogens, including the conifer pathogen *Heterobasidion annosum*, from which the protein caused cell death in both *Pinus sylvestris* and *Nicotiana tabacum* (Chen et al. 2015). Specific up-regulation of Ds70155 in the late stage of pine infection may support a necrotrophic function for this candidate. Several CAZys, including glycoside hydrolase (GH) 13- and 64-type enzymes, were highly expressed and up-regulated *in planta* (Table 1). Such GHs could be involved in degradation of plant-derived glucan deposits that hinder fungal entry (O’Connell et al. 2012), or re-modelling of the fungal cell wall when switching from epi- to endophytic growth (Bradshaw et al. 2016).

Of the 55 candidate effectors, 40 did not show similarity to any characterised proteins, and thus are of unknown function. These potentially novel proteins, along with the 15 for which a function could be predicted, provide a rich resource for further study.

CONCLUSIONS

Potential effector genes were identified from the *D. septosporum* genome using a bioinformatics pipeline and gene expression data. To the best of our knowledge, this was the first time this type of

genome-wide approach had been used to identify effectors in a fungal pine pathogen. Effector candidates predicted to be highly expressed *in planta* included novel proteins but also some similar to known virulence factors. Identification of effector candidates in agricultural pathogens using similar methods underpinned efforts in host resistance screening (Haddadi et al. 2016) and thus may provide the first step towards effector-based resistance breeding in pines.

ACKNOWLEDGEMENTS

We thank Dr Carla Eaton and Dr Pierre-Yves Dupont for feedback on the pipeline design and Scion for funding this project.

REFERENCES

- Barnes I, Wingfield MJ, Carbone I, Kirisits T, Wingfield BD 2014. Population structure and diversity of an invasive pine needle pathogen reflects anthropogenic activity. *Ecology and Evolution* 4(18): 3642-61.
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* 340(4): 783-795.
- Bolton MD, van Esse HP, Vossen JH, de Jonge R, Stergiopoulos I, Stulemeijer IJE, van den Berg GCM, Borrás-Hidalgo O, Dekker HL, de Koster CG et al. 2008. The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. *Molecular Microbiology* 69(1): 119-136.
- Bradshaw RE, Guo Y, Sim AD, Kabir MS, Chettri P, Ozturk IK, Hunziker L, Ganley RJ, Cox MP 2016. Genome-wide gene expression dynamics of the fungal pathogen *Dothistroma septosporum* throughout its infection cycle of the gymnosperm host *Pinus radiata*. *Molecular Plant Pathology* 17(2): 210-224.
- Chen H, Quintana J, Kovalchuk A, Ubhayasekera W, Asiegbu FO 2015. A cerato-platanin-like protein HaCPL2 from *Heterobasidion annosum sensu stricto* induces cell death in *Nicotiana tabacum* and *Pinus sylvestris*. *Fungal Genetics and Biology* 84: 41-51.
- de Wit PJGM, van der Burgt A, Okmen B, Stergiopoulos I, Abd-Elsalam KA, Aerts AL, Bahkali AH, Beenen HG, Chettri P, Cox MP et al. 2012. The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PLoS Genetics* 8(11): e1003088.
- Eisenhaber B, Schneider G, Wildpaner M, Eisenhaber F 2004. 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*. *Journal of Molecular Biology* 337(2): 243-253.
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of Molecular Biology* 300(4): 1005-1016.
- Haddadi P, Ma L, Wang H, Borhan MH 2016. Genome-wide transcriptomic analyses provide insights into the lifestyle transition and effector repertoire of *Leptosphaeria maculans* during the colonization of *Brassica napus* seedlings. *Molecular Plant Pathology*; doi 10.1111/mpp.12356.
- Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* 305(3): 567-580.
- O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, Torres MF, Damm U, Buiate EA, Epstein I, Alkan N et al. 2012. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genetics* 44(9): 1060-1067.
- Sperschneider J, Dodds PN, Gardiner DM, Manners JM, Singh KB, Taylor JM 2015. Advances and challenges in computational

- prediction of effectors from plant pathogenic fungi. *PLoS Pathog* 11(5): e1004806-e1004806.
- Storey JD 2003. The Positive False Discovery Rate: A Bayesian interpretation and the q-value. *The Annals of Statistics* 31(6): 2013-2035.
- Vleeshouwers VGAA, Oliver RP 2014. Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. *Molecular Plant-Microbe Interactions* 27(3): 196-206.