Unraveling the mechanism of *ToxB* replication in *Pyrenophora tritici-repentis*

Ryan Gourlie[1,2], Megan McDonald[3], Dmytro Yevtushenko[2] and Reem Aboukhaddour[1]

[1]Agriculture and Agri-Food Canada, LRDC, Lethbridge, AB, Canada;
[2]University of Lethbridge, Lethbridge, AB, Canada;
[3]University of Birmingham, School of Biosciences, Edgbaston, Birmingham, United Kingdom

BioNet BIOINFORMATICS NETWORK ALBERTA

Agriculture and Agri-Food Canada

Introduction

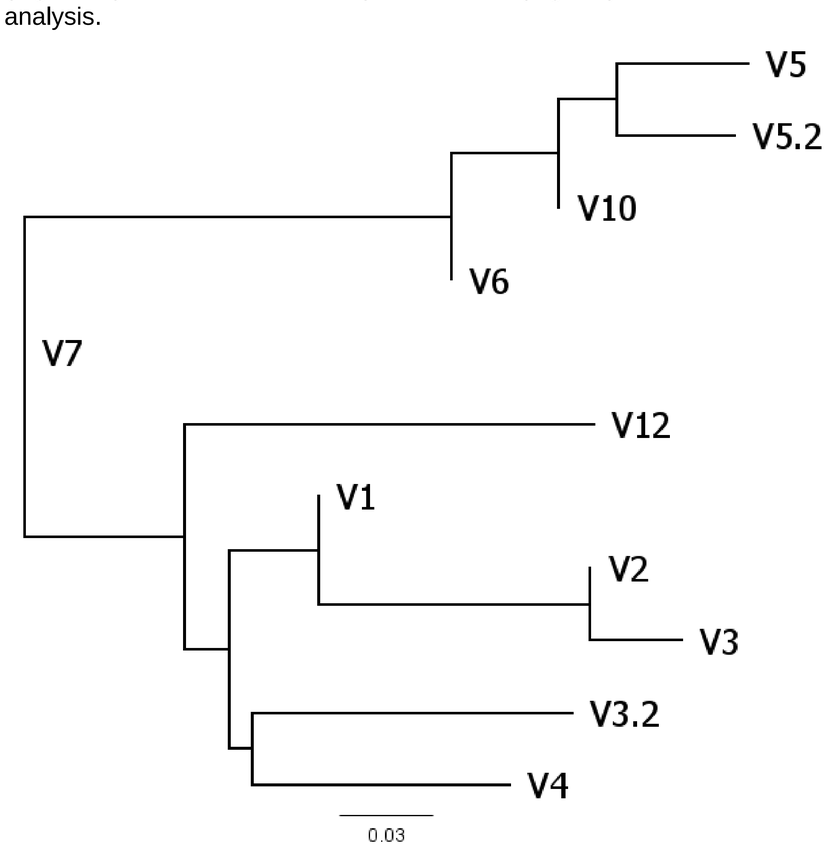
The necrotrophic fungal pathogen *Pyrenophora tritici-repentis* (Ptr) is a major foliar wheat pathogen with annual global losses estimated at ~4.3% (Savary et al., 2019). The Ptr-wheat interaction is used as a model system to help understand the evolution of virulence in necrotrophs. Ptr races are defined by their ability to produce different combinations of three effectors (host-selective toxins): ToxA, ToxB, and ToxC. Due to it's more damaging symptoms, the most extensively studied effector to-date has been *ToxA* which is a necrosis inducing toxin. Previously we have shown that the *ToxA* carrying transposon *ToxhAT* (McDonald et al., 2018; 2019) is itself nested within a larger *Starship* class transposon (*Horizon*) (Gourlie, et al., 2022). ToxB which causes chlorosis symptoms has been less well studied, perhaps due to its multi-copy nature. Previously, we used long-read assemblies to identify the genomic position of *ToxB*-copies in two isolates and found *ToxB* to be present within a different putative Starship transposon (*Icarus*) (Gourlie et al., 2022). However, the presence of *ToxB* within *Icarus* did not explain its multicopy nature, and further investigation was needed. In this work we sequenced eight (8) additional isolates of Ptr with PacBio RSII, including isolates with various copy numbers. With this new data we have finally begun to unravel the complex evolution and replication of the virulence gene *ToxB*.

Results and Discussion

Different avenues of investigation including self-alignments, full contig alignments, and alignments of regions surround *ToxB* caused possible edges of the replicative unit to emerge. The evidence for the edges was a large drop in synteny at the same locations in different isolates and different *ToxB* copies within the same isolate. Unexpectedly these edges (pink and blue boxes in Figure 1) did not match either directly or by reverse-compliment, as would be expected with many types of transposons (Wicker et al., 2007). After establishing putative edges, each copy of *ToxB* was compared edge-to-edge with the shortest version of the putative transposon to identify potential indels. These comparisons revealed many different insertions within the putative boundaries, with insertion sizes ranging from 170 bp up to 12.5 Kbp (Figure 1). Some of these insertions may themselves be transposons as nesting is not uncommon. Open-reading frames (ORFs) were identified within the putative edges and just beyond their borders to identify possible genes associated with transposon activity. After searching for domain functionality using UniProt, several ORFs of interest were noted. Primarily the presence of an ORF with reverse-transcriptase and RNase H type 1 domains present within the defined edges. Genes with these domains are associated with Class 1 retrotransposons (Wicker et al., 2007). RT domains are involved with converting RNA into cDNA in preparation for genome insertion, and RNase H domains are endonucleases that catalyze cleavage of RNA/DNA substrates. An ORF integrase was found just outside the putative edges as well, integrase is again associated with retrotransposons (and retroviruses) and is involved with the insertion (integration) of cDNA into the host. A BLAST search of RepBase (transposons database; Jurka et al., 2005) for the RT/RNase H, showed some similarity to *I-2_CH* (67% identity over 96% of the query) which is a non-LTR retrotransposon identified in the necrotrophic maize pathogen *Cochliobolus heterostrophus* (Santana et al., 2014). All these results point towards a possible Class 1 transposon is or was in the past associated with the replication of *ToxB*.

Figure 1 Putative replicative unit containing the *ToxB/toxb* gene in *Pyrenophora tritici-repentis*. Multiple variants are shown; insert sizes shown in red; conserved mismatching edges are pink or blue; retrotransposon related genes are green; conserved repeats in purple; *ToxB/toxb* in yellow; and loss of synteny with the reference (variant v1) and contig truncation shown in white.

Figure 2 Neighbour-joined phylogeny of *ToxB/toxb* replication unit variants. Sequences with a star in Figure 1 were included. Presumed incomplete (e.g. v1.2) and difficult to interpret variants (e.g. v9) were omitted from analysis.



A phylogeny of complete versions (i.e. edge-to-edge) of the putative transposon was created to identify evolutionary relationships between the various copies (Figure 2). Distance matrices used to create neighbour-joined phylogenies don't handle large insertions well in their analysis, and so the figure presented may not reflect an accurate evolutionary structure. The sequencing of more isolates, including closely related species in the *Pyrenophora* genus is underway to help confirm what we've found thus far. The evolutionary history of multi-copy *ToxB/toxb* is complicated and difficult to unravel and this research marks a significant step forward in our understanding of this virulence gene.

Materials and Methods

Sequencing and assembly: High-molecular weight DNA extracted with 'Genomic-tip 100/G Kit' (Qiagen) and gDNA sequenced with PacBio RS II at Genome Quebec. I-73-1 and D308 were previously assembled with Flye (Gourlie et al., 2022; Kolmogorov et al., 2015). DW5 assembly was retrieved from GenBank (Moolhuijzen et al., 2020). All others were assembled with CANU (v2.2; Koren et al., 2017).

Transposon identification and variant cataloguing: Various pair-wise linear alignments were made from contigs containing ToxB/toxb using Geneious Prime (Biomatters Ltd) extensions Mauve (v1.1.3; Darling et al., 2010) and LASTZ (v7.0.3; Haris 2007). Individual copies of ToxB along with their surrounding regions were extracted and smaller pair-wise alignments were made with MAFFT (v1.5.0; Katoh and Standley, 2013). Based on sequence similarity changes, putative edges were defined. Geneious ORF Finder was used to find open-reading frames within putative edges. UnitProt (UniProt Consortium) searches of identified ORFs using default parameters provided domain information. The variant with the shortest sequence length was selected to be the reference (variant v1) to which other variants would be compared. Inserts and syntenic blocks for each copy were individually categorized via alignment to the reference.

References

Savary et al., 2019. Nat Ecol Evol, 3(3), 430; McDonald et al., 2019. mBio. 10(5), e01515–19; McDonald et al., 2018. Mol Plant Pathol. 19(2), 432–9;Gourlie et al., 2022. BMC Biol. 20, 239; Wicker et al., 2007. Nat Rev Genet. 8, 973-982; Jurka et al., 2005. Cyto Gen Res. 110(1-4), 462-467; Santana et al., 2014. BMC Genomcs. 15, 536; Kolmogorov et al., 2019. Nat Biotech. 37(5), 540–546; Darling et al., 2010. PLoS ONE. 5(6), e11147; Moolhuijzen et al., 2020. BMC Genomics. 21(1), 1-12; Koren et al., 2017. Genome Res. 27(5), 722-736; Haris, R. 2007. PhD Thesis. Penn State; Katoh and Standley, 2013. Mol Biol and Evol. 30(4), 722-780; UniProt Consortium, 2015. Nucl Acid Res. 43(D1), D204-212

University of Lethbridge · Western Grains Research Foundation · Alberta Wheat Commission · Manitoba Wheat and Barley Growers Association · University of Birmingham · Sask Wheat Development Commission · Canada