COMPARATIVE GENOMICS OF TWO ISOLATES OF *S. SCITAMINEUM* WITH DIFFERENT LEVELS OF AGGRESSIVENESS IN SUGARCANE SMUT DISEASE

Pedro Fernando Vilanova Ferreira[1]

Hugo Rody Vianna Silva[1]

Claudia Barros Monteiro-Vitorello[1]

[1]College of Agriculture "Luiz de Queiroz"

pedrovilanova@usp.br

Objectives

Effectors are molecules secreted by pathogens to establish colonization by manipulating the plant immune system and metabolism. As a counterattack, the immune system of hosting plants evolve rapidly, mainly through the known R (resistance) genes, imposing selection pressure on pathogens and their effectors. Many pathogens attack important crops, and understanding the arsenal of effectors molecules and the mechanisms to surpass crop resistance may benefit worldwide breeding programs. Herein, we sequenced the whole-genome of an aggressive isolate (SSC04) of the biotrophic fungus *Sporisorium scitamineum*, the causing agent of sugarcane smut disease. We then performed whole-genome alignments between SSC04 and a reference genome (SSC39) (Taniguti et al., 2015) and investigated whether the predicted candidate effectors (CE) were polymorphic among isolates.

Materials and Methods

The whole-genome sequencing of the aggressive isolate SSC04 was performed by the Oxford/Nanopore MinION platform. Whole chromosome sequences were aligned to the reference genome SSC39 using the Mummer software v3.22 wrapped into the NucDiff v2.9.3 pipeline. Candidate effector (CE) genes obtained from Taniguti et al., 2015 and identified as polymorphic between isolates were selected. BLASTp searches followed by manual inspection of alignments were used to create orthogroup datasets composed of the CEs and its paralogs in SSC39, in addition to their corresponding best orthologs in SSC04. Conserved domains in sequences were investigated using InterproScan v5.33–72.0 (Jones et al., 2014). Maximum-Likelihood phylogenetic relationships among orthogroups were inferred using IQ-tree (Nguyen et al., 2015) software from amino acid alignments using MUSCLE (Edgar, 2004). The best-fit model of molecular evolution was selected by the ModelFinder software (Kalyaanamoorthy et al., 2017) with the parameter -m MFP. The consensus tree was visualized and edited using iTOL v4.3.2 (Letunic; Bork, 2019).

Results

The two isolates showed a high synteny level between ortholog genes (Figure 1). Whole chromosome sequences were 99.72% aligned, and 1.821 genes had polymorphisms in their coding sequence. The content of transposable elements representing 5.23% of the SSC04 genome differed between the isolates.
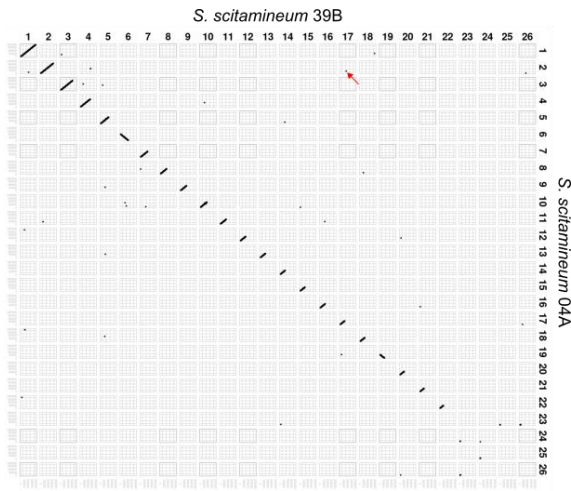
Figure 1. Ortholog best hits showing shared synteny between genes in the SSC39 and SSC04 *S. scitamineum* isolates. The red arrow indicates alignment in both chromosomes 2 and 17.

Six CEs showed non-synonymous substitutions between isolates. Further, BLAST searches showed four out of the six CEs having paralogs within genomes of isolates, and all the six CEs having orthologs between isolates. Phylogenies indicated the orthogroups (paralogs and orthologs) under distinct patterns of evolution, with paralogs harboring elevated variation compared to orthologs. The CE of g3085 (Figure 2) formed the largest orthogroup (N=16) and was highly conserved among isolates. In contrast, the intermediate size orthogroup of g1052 was found to have the largest number of polymorphisms between isolates.
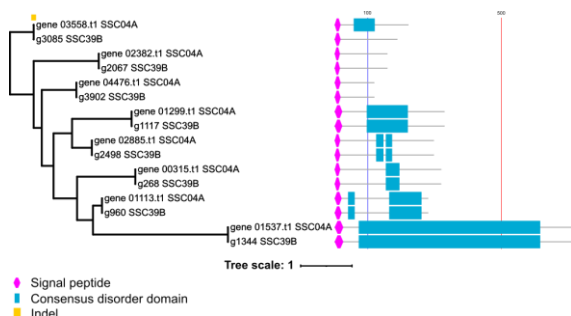

Figure 2. Unrooted Maximum Likelihood Phylogenetic consensus tree showing phylogenetic relationships among orthogroup of CE g3085.

Predicted domains in sequences are depicted alongside according to the legend. Horizontal lines in domains indicate sequence length according to reference vertical lines. Branch lengths are drawn to scale.

Conclusions

Our results suggest the CEs orthogroups evolving under different constraints, with significant sequence evolution occurring among paralogs before isolates split and may help further studies on effector biology and evolution.

References

EDGAR, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research, v. 32, n. 5, p. 1792–1797, 2004.

JONES, P. et al. InterProScan 5: genome-scale protein function classification. Bioinformatics, v. 30, n. 9, p. 1236, 2014.

KALYAANAMOORTHY, S. et al. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature methods, v. 14, n. 6, p. 587–589, 2017.

LETUNIC, I.; BORK, P. Interactive Tree of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Research, v. 47, n. W1, p. W256–W259, 2019.

NGUYEN, L.-T. et al. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution, v. 32, n. 1, p. 268–274, 2015.

TANIGUTI, L. M. et al. Complete genome sequence of *Sporisorium scitamineum* and biotrophic interaction transcriptome with sugarcane. PLoS ONE, v. 10, n. 6, 2015.