

# Sequence divergence in venom genes within and between montane pitviper (Viperidae: Crotalinae: *Cerrophidion*) species is driven by mutation-drift equilibrium

Ramses Alejandro Rosales-García<sup>a,\*</sup>, Rhett Rautsaw<sup>a</sup>, Erich P. Hofmann<sup>a,b</sup>, Christoph I. Grünwald<sup>c,d</sup>, Hector Franz-Chavez<sup>c,d</sup>, Ivan T. Ahumada-Carrillo<sup>c,d</sup>, Ricardo Ramirez-Chaparro<sup>c,d</sup>, Miguel Angel De la Torre-Loranca<sup>e</sup>, Jason L. Strickland<sup>a,f</sup>, Andrew J. Mason<sup>a,g</sup>, Matthew L. Holding<sup>a,h</sup>, Miguel Borja<sup>i</sup>, Gamaliel Castañeda-Gaytan<sup>i</sup>, Edward A. Myers<sup>a</sup>, Mahmood Sasa<sup>j</sup>, Darin R. Rokyta<sup>k</sup>, Christopher L. Parkinson<sup>a,\*</sup>

<sup>a</sup> Department of Biological Sciences, Clemson University, Clemson, SC 29634

<sup>b</sup> Science Department, Cape Fear Community College, Wilmington, NC 28401

<sup>c</sup> Herp.mx A.C., Colima, Mexico

<sup>d</sup> Biodiversa A. C., Chapala, Jalisco, Mexico 45900

<sup>e</sup> Instituto Lorancai, Ocotepéc, Veracruz, Mexico 24105

<sup>f</sup> Department of Biology, University of South Alabama, Mobile, AL 36688

<sup>g</sup> Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, Ohio 43210

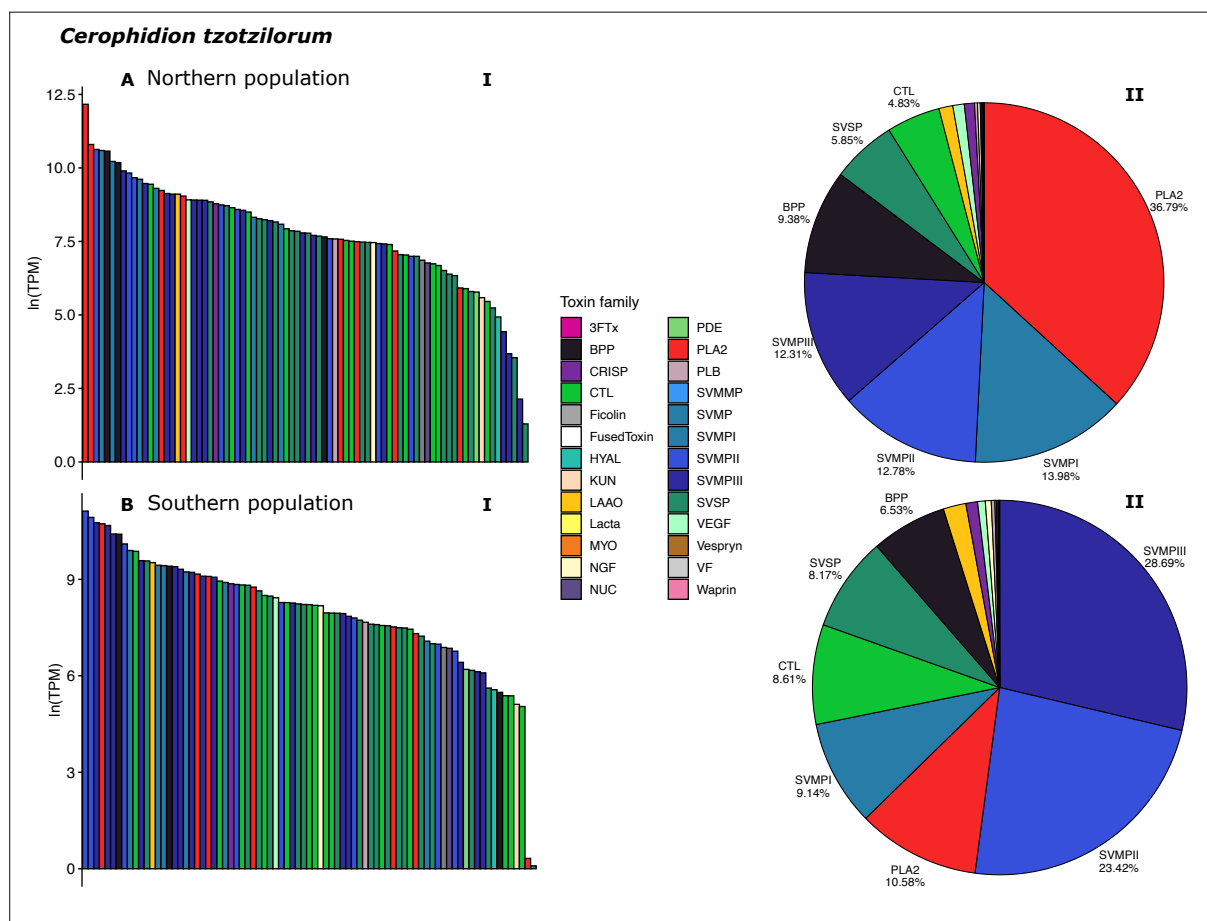
<sup>h</sup> Life Sciences Institute, University of Michigan, Ann Arbor, MI 48109

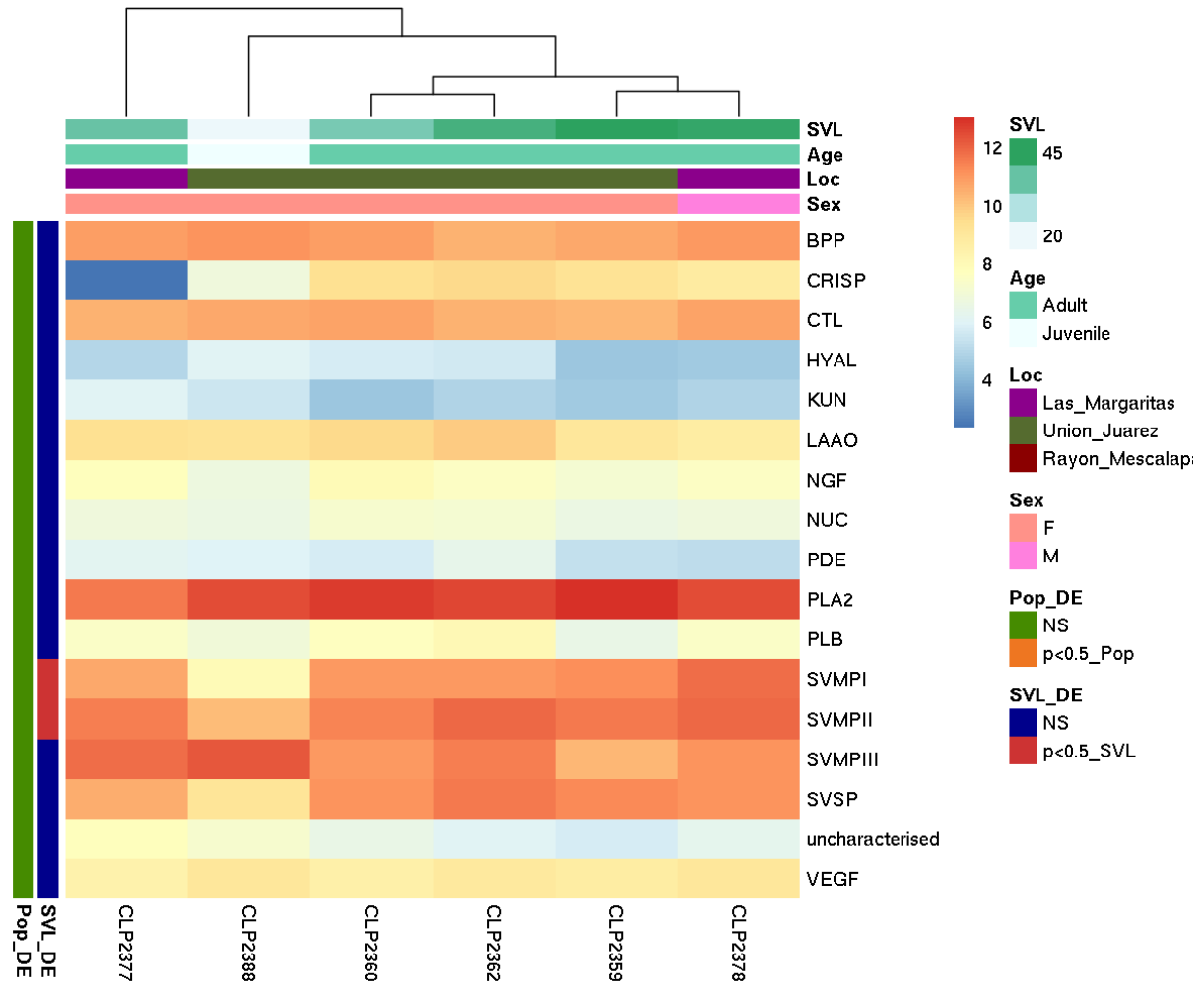
<sup>i</sup> Facultad de Ciencias Biológicas, Universidad Juárez del Estado de Durango, Gómez Palacio, Durango, Mexico, 35010

<sup>j</sup> Centro Investigaciones en Biodiversidad y Ecología Tropical and Instituto Clodomiro Picado, Universidad de Costa Rica, San José, Costa Rica

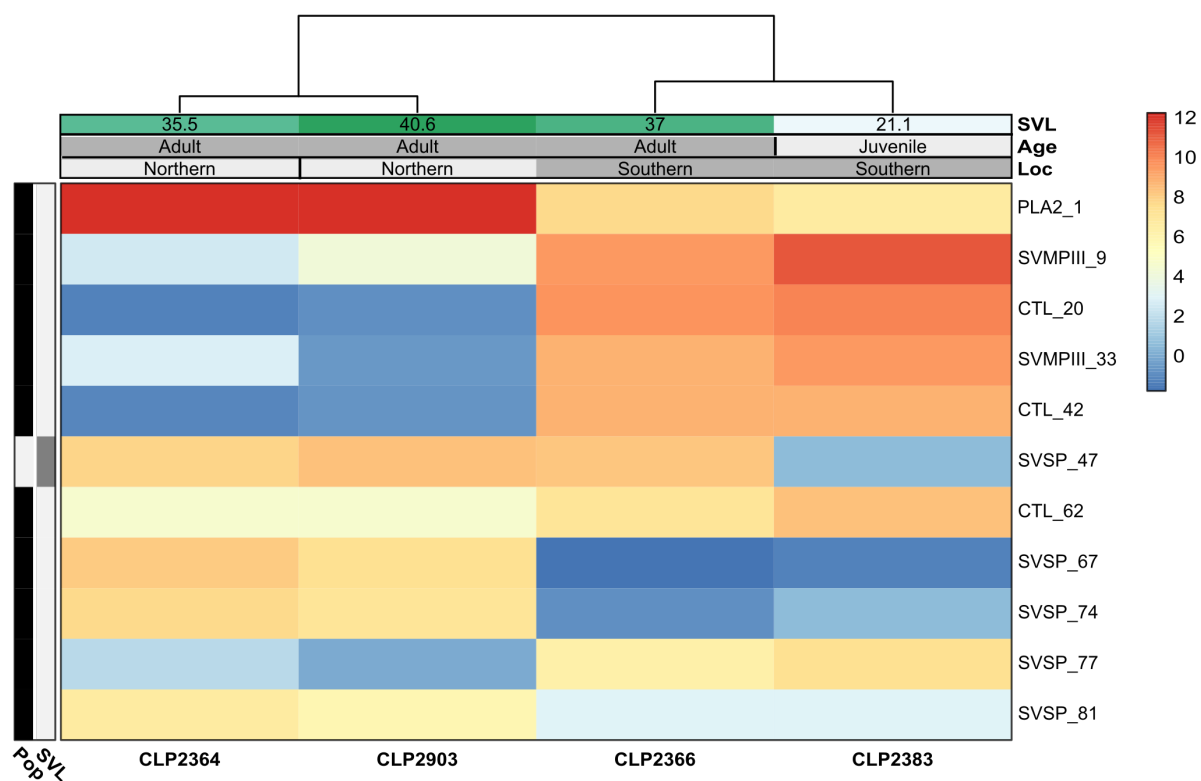
<sup>k</sup> Department of Biological Science, Florida State University, Tallahassee, Florida 32306

\* Corresponding author: viper@clemson.edu, ramsesr@g.clemson.edu

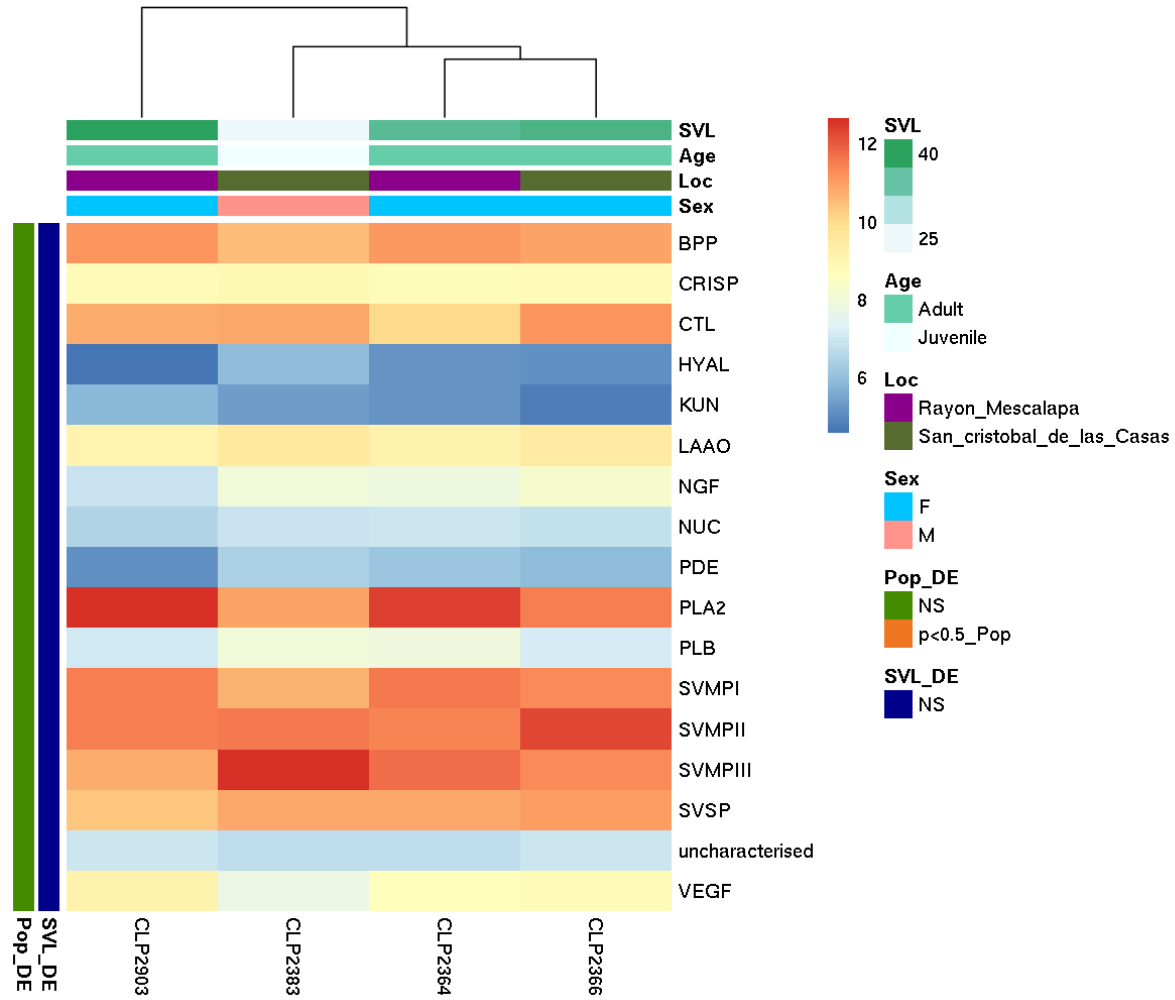




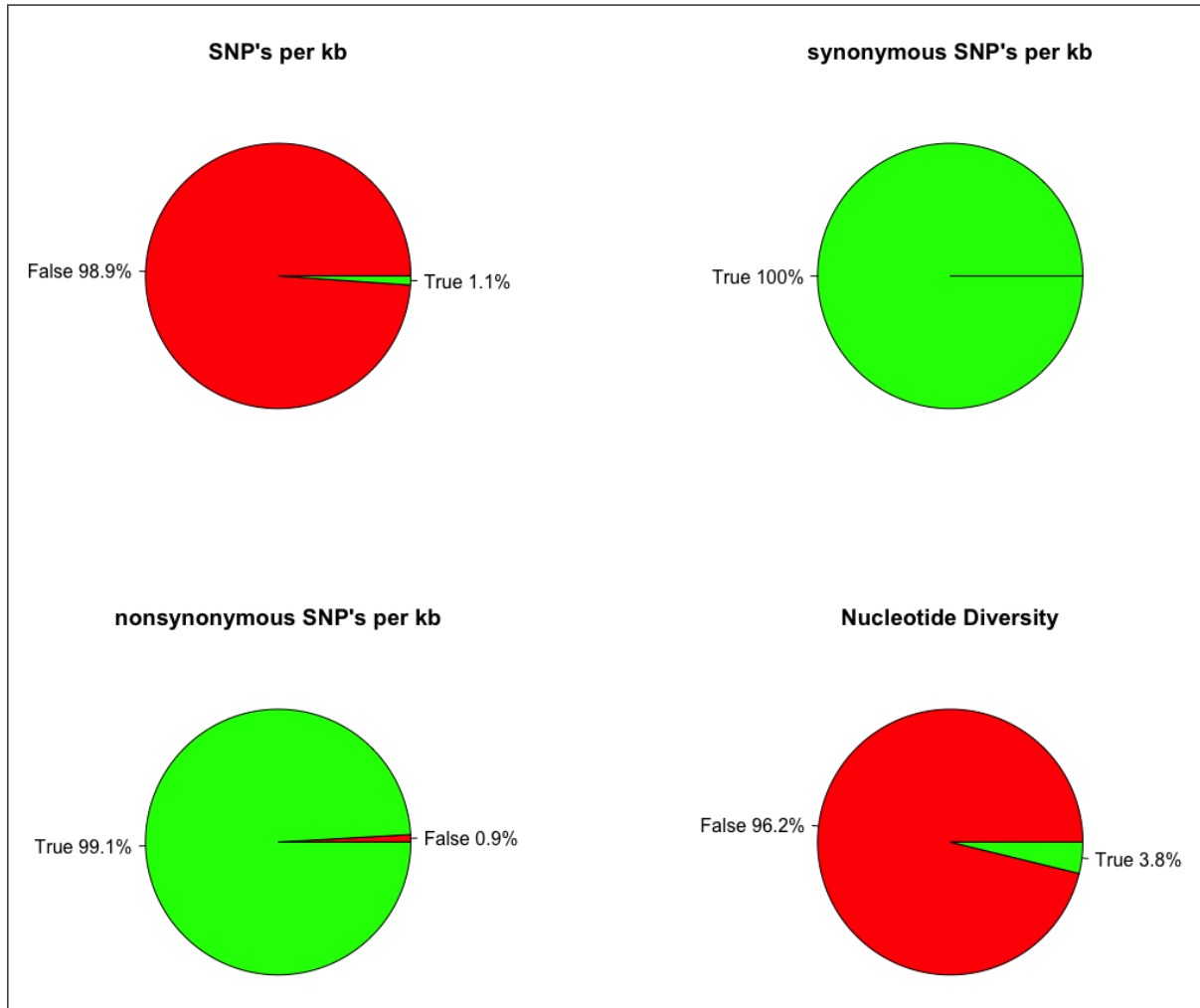
**Figure S2.** Heatmap showing the log TPM expression of toxin families in *C. godmani*. In the left columns (Pop & SVL) the orange and red colors respectively represent significant differential expression agreement by both DESeq2 and edgeR ( $FDR < 0.05$ ).



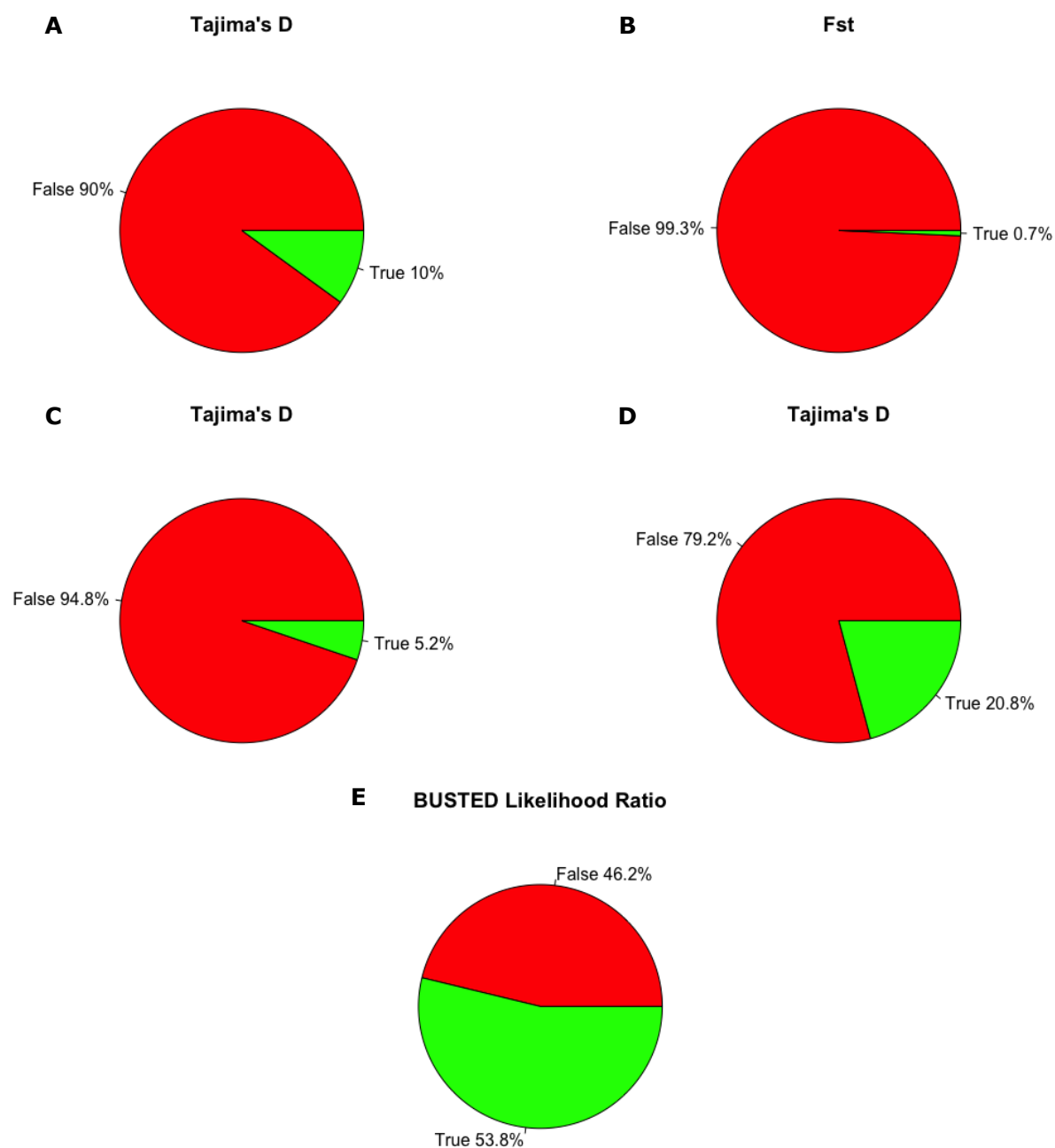
**Figure S3.** Heatmap showing the log TPM expression of toxins identified as differentially expressed in *C. tzotzilorum* ordered by the average expression. In the left columns (Pop & SVL) the darker colors represent significant differential expression agreement by both DESeq2 and edgeR ( $FDR < 0.05$ ).



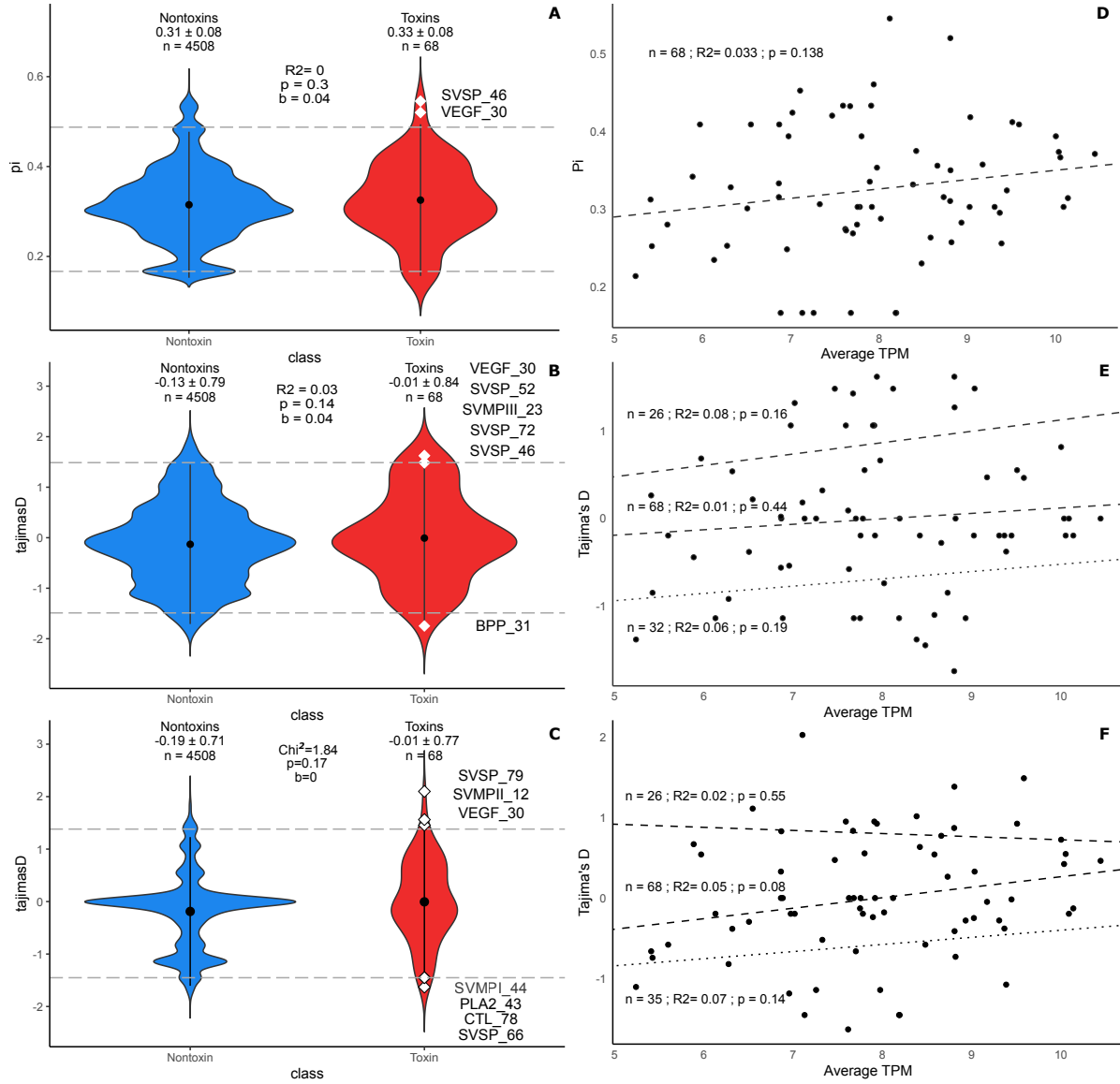
**Figure S4.** Heatmap showing the log TPM expression of toxin families in *C. tzotzilorum*. None of the toxin families were significantly differently expressed by both DESeq2 and edgeR ( $FDR < 0.05$ ), However PLA<sub>2</sub>s were differentially expressed in DESeq2 by population, and HYAL, SVMPIII, and VEGF were differentially expressed in DESeq2 by SVL.



**Figure S5.** Bootstrap pie charts, with 1000 repetitions of the linear model comparing toxins against nontoxins with equal sample size. True in green ( $p < 0.05$ ) and false in red ( $p > 0.05$ ).

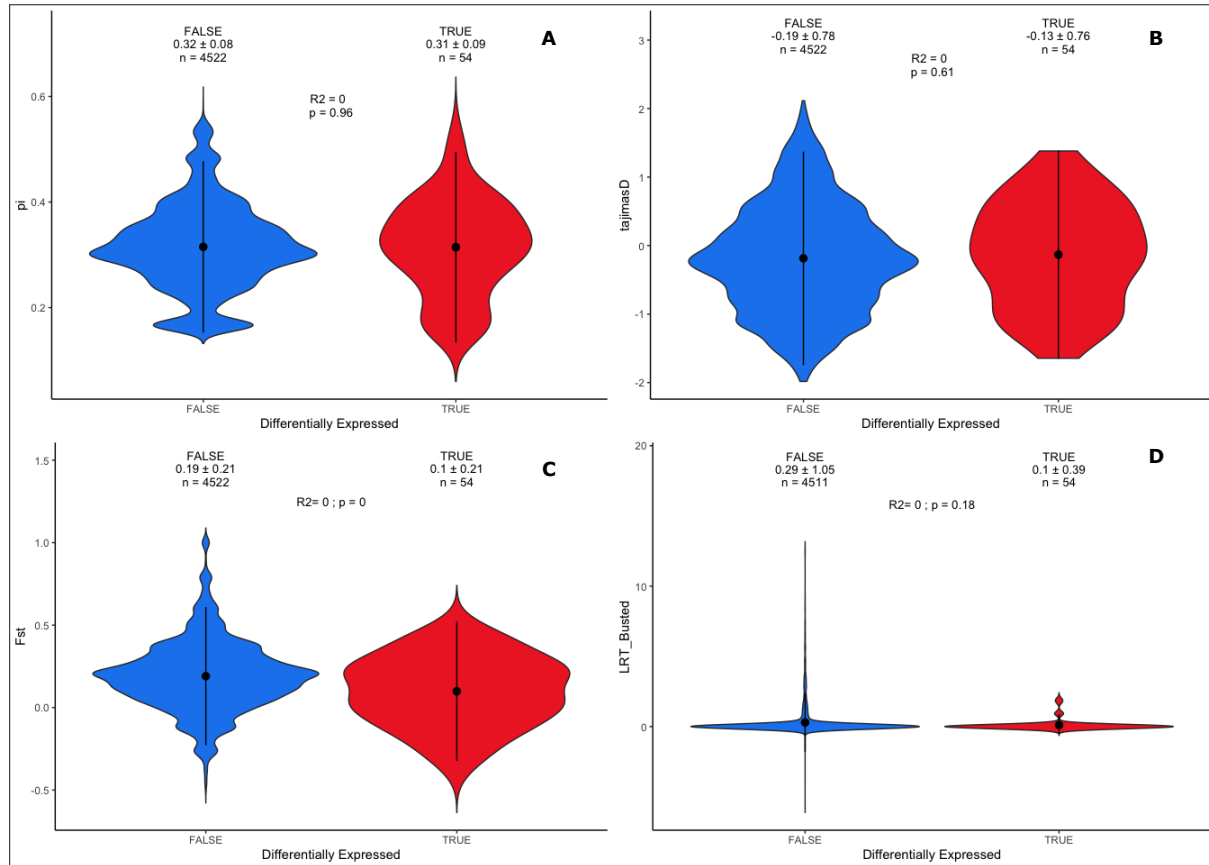


**Figure S6.** Bootstrap pie charts, with 1000 repetitions of the linear model comparing toxins against nontoxins with equal sample size. True in green ( $p < 0.05$ ) and false in red ( $p > 0.05$ ). **A** Tajima's D. **B**  $F_{ST}$ . **C** Synonymous SNPs Tajima's D. **D** Nonsynonymous SNPs Tajima's D. **E** BUSTED model LRT, tested with a non parametric Kruskal Wallis test.

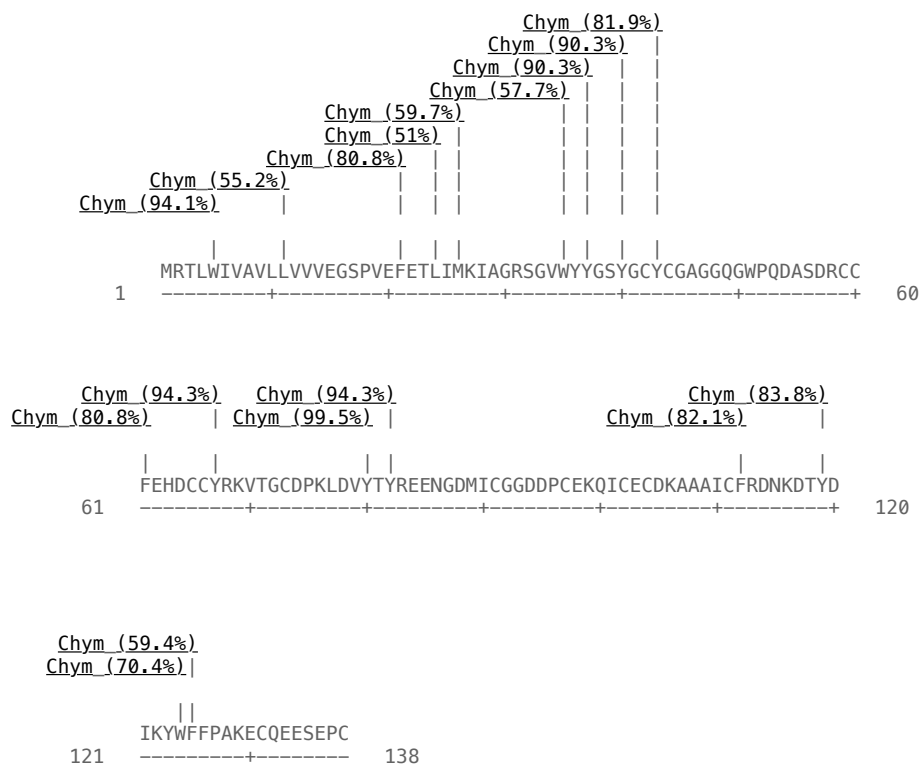


**Figure S7.** Selection plots. Left: estimates of selection using **A** Nucleotide Diversity ( $\pi$ ), **B** Synonymous SNPs Tajima's D, and **C** Nonsynonymous SNPs Tajima's D for toxins and nontoxins, each with the nontoxin 95th percentile (dotted lines) to identify outlier toxins. The toxin family and the rank based on highest-to-lowest average expression in the transcriptome are displayed for toxins which fall outside the 95th percentile. Right: Linear regressions of the Toxin's mean expression (Average TPM) and estimates of selection including **D** Nucleotide Diversity ( $\pi$ ), **E** Synonymous SNPs Tajima's D., and **F** Nonsynonymous SNPs Tajima's D. For Tajima's D, dotted lines are regressions of all the transcripts (center), just positive values (top) and just negative values (bottom).





**Figure S8.** Selection plots. Left: estimates of selection using **A** Nucleotide Diversity ( $\pi$ ), **B** Tajima's D, **C**  $F_{ST}$ , and **D** BUSTED model LRT for non differentially expressed genes (FALSE) and differentially expressed (TRUE).



**Figure S9.** Expasy Peptide Cutter Results ([https://web.expasy.org/peptide\\_cutter/](https://web.expasy.org/peptide_cutter/)) for Cgodm\_PLA2\_11. The peptide cutter tool was set to identify Chimotrypsin like cleavage residuals with cut of 50 % probability of cleavage.