Tips & Tools

Using Tobamoviruses for Phylogenetic Instruction in Undergraduate Biology Courses †

Kornelia Fillmer[1], Scott Adkins[2], Patchara Pongam[1], and Tom D'Elia[1],*
[1]Biology Department, Indian River State College, Fort Pierce, FL 34981,
[2]USDA, ARS, U.S. Horticultural Research Laboratory, Fort Pierce, FL 34945

INTRODUCTION

The rapid advancement and availability of sequencing technologies are necessitating the inclusion of a core set of bioinformatics competencies in biology curricula. Many curriculum guidelines have developed specific recommendations for key concepts and skill sets related to bioinformatics and molecular evolution analysis (1–4). To broaden student exposure to bioinformatics, we have developed an activity that uses amino acid sequences from Tropical soda apple mosaic virus (TSAMV) and other tobamoviruses to facilitate teaching multiple sequence alignment and phylogenetics in undergraduate biology courses. The activity introduces students to basic bioinformatics and provides a core skill set in fundamental sequence analysis to help students better understand the key relationships between phylogenetics and virus evolution. Tobamoviruses are a good teaching model because well-established phylogenetic groups form naturally, presenting results that can be readily compared with those published in scientific literature. Additionally, students can make important societal connections when analyzing the potential economic impact of this group of agriculturally important plant pathogens.

The activity described here was developed for an introductory bioinformatics course focusing on basic sequence analysis for biologists. In a computer lab, students use a freely available, user-friendly bioinformatics program to analyze the TSAMV replicase protein, and then use this information to better understand the evolution and ecology of this virus. To reach these goals, students complete multiple sequence alignments and phylogenetic analyses of the TSAMV replicase amino acid sequence in comparison with other tobamovirus sequences. The activity is broadly adaptable and easily introduced in courses that include microbiology, molecular biology, genetics, or evolution.

PROCEDURE

Preparing students for the activity

Tobamoviruses are rod-shaped RNA viruses that infect a broad range of plants including tomatoes, peppers, and cucumbers. Genomes of tobamoviruses are approximately 6.3–6.5 kb and consist of four open reading frames (ORFs): a 124- to 132-kDa replicase protein with methyltransferase and RNA helicase domains, a 181- to 189-kDa replicase protein with a polymerase domain, a movement protein (MP), and a coat protein (CP) (5–7). Tropical soda apple mosaic virus, a tobamovirus that was originally isolated from tropical soda apple (TSA), has also been shown to infect other solanaceous crops (8). Tropical soda apple is an invasive weed in the Solanaceae plant family and has spread throughout the southern United States, taking over large areas of cow pastures (9). Tropical soda apple is naturally infected by tobamoviruses and members of several other families of agriculturally important viruses and may facilitate virus dissemination. Tobamoviruses are excellent for introducing the principles of phylogenetics to students because three major phylogenetic groups naturally form based on the host plant family infected (Solanaceae, Cucurbitaceae, or Brassicaceae) (6, 10). Six single tobamovirus species branches also form, again based on host plant family. This activity utilizes the replicase amino acid sequence of TSAMV to determine the evolutionary relationship of this virus to other tobamoviruses. This analysis provides the required context needed for developing diagnostics and preventive control measures to limit the spread of TSAMV and related viruses.

Activity: Multiple sequence alignment and phylogenetic analysis

Students complete a multiple sequence alignment and phylogenetic analysis of 28 tobamovirus and 1 outgroup replicase amino acid sequences (see Appendix 1 for supplementary instructions). Sequences can be collected from NCBI using the accession numbers or amino acid text file provided (Appendices 2 and 3, respectively). Students then import sequences into MEGA7 for their analysis (11). After the sequences are imported into MEGA7, students should

note the number of sequences and their general lengths. The students will then use ClustalW to make the multiple sequence alignment. The alignment may take several minutes to complete. Once the sequence alignment is complete, students should be directed to evaluate their alignments. In general, observations can be made by looking for major regions of the sequences that are similar and the presence and distribution of gaps. After completing the multiple sequence alignment, students will use the same software to create a neighbor-joining phylogenetic tree. Once the tree is constructed, students can be directed to make their interpretations based on the groupings they observe within the phylogenetic tree. They should be able to determine which species on the tree belong to which tobamovirus phylogenetic group in Appendix 2. Based on this analysis, students will recognize that the species form three major groups and six single species branches based on the host family these viruses infect. The phylogenetic analysis will place TSAMV with Pepper mild mottle virus (PMMoV) within the Solanaceae group (10). Students can then be directed to predict other host plants that TSAMV may potentially infect. Through a literature search, or through the reference provided (8), students will discover that TSAMV can infect bell peppers and tomatoes. This is predicted in the tree, as TSAMV is most closely related to PMMoV and other viruses from peppers and tomatoes. The phylogenetic analysis reinforces the idea that the biology of TSAMV can be better understood through comparative molecular analysis. The entire activity can be completed in a single class, or spread over two class periods.

Extensions

Several extensions to this activity emerge once the basic processes of multiple sequence alignments and phylogenetics are understood. Students can use the genome accession numbers in Appendix 2 to collect complete genome sequences, or amino acid sequences for MP or CP, to generate a phylogenetic tree and compare the results with their replicase tree and literature results (6, 8, 10).

CONCLUSION

This activity has been implemented over the past four semesters, and more than 60 students have completed it. Students report a better understanding of molecular evolution by having the opportunity to see how actual similarities and differences in amino acid sequence alignments translate to phylogenetic relationships. The formation of groups based on host plant family provides a very simple framework for students to learn how to start evaluating phylogenetic trees. The ease of the activity demystifies sequence analyses for the students and opens their horizons to other applications including epidemiology, human genetics, and health and disease. Responses from students also make evident the value of this practical exercise in building their confidence for future bioinformatics analyses.

SUPPLEMENTAL MATERIALS

Appendix 1: Instructor notes and handout for classroom instruction
Appendix 2: Table of tobamoviruses, abbreviations, and accession numbers
Appendix 3: Text file of 28 tobamovirus and 1 outgroup replicase amino acid sequences

REFERENCES

1. American Association for the Advancement of Science. 2011. Vision and change in undergraduate biology education: a call to action: a summary of recommendations made at a national conference organized by the American Association for the Advancement of Science, July 15–17, 2009. Washington, DC.
2. Merkel S, Reynolds J, Hung K, Siegesmund A, Smith A, Smith H. 2012. ASM curriculum guidelines for undergraduate microbiology education. [Online.] http://www.asm.org/index.php/guidelines/curriculum-guidelines
3. White HB, Benore MA, Sumter TF, Caldwell BD, Bell E. 2013. Special section: foundational concepts and assessment tools for biochemistry and molecular biology educators, part 1: essential concepts and skills. Biochem Mol Biol Educ 41:297–301.
4. Welch L, Lewitter F, Schwartz R, Brooksbank C, Radivojac P, Gaeta B, Schneider MV. 2014. Bioinformatics curriculum guidelines: toward a definition of core competencies. PLOS Comput Biol 10(3):e1003496.
5. Adams MJ, Heinze C, Jackson AO, Kreuze JF, Macfarlane SA, Torrance L. 2012. Tobamovirus, p 1153–1156. In King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Virus taxonomy: ninth report of the international committee on taxonomy of viruses. Elsevier/Academic Press, London, United Kingdom.
6. Gibbs A. 1999. Evolution and origin of tobamoviruses. Philos Trans R Soc Lond B Biol Sci 354:593–602.
7. Lewandowski DJ, Dawson WO. 2000. Functions of the 126- and 183-kDa proteins of tobacco mosaic virus. Virology 271:90–98.
8. Adkins S, Kamenova I, Rosskopf EN, Lewandowski DJ. 2007. Identification and characterization of a novel tobamovirus from tropical soda apple in Florida. Plant Dis 91:287–293.
9. Coile NC. 1993. Tropical soda apple, Solanum viarum Dunal: the plant from hell. Florida Department of Agriculture and Consumer Services, Division of Plant Industry. Bot Circ 27.
10. Fillmer K, Adkins S, Pongam P, D'Elia T. 2016. The complete nucleotide sequence and genomic characterization of tropical soda apple mosaic virus. Arch Virol 161:2317–2320.
11. Kumar S, Stecher G, Tamura, K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874.