Using DNA Sequences and Phylogenetic Trees as Tools for Teaching Entomology to Undergraduate Students: A Simple Approach

Mohammad Ali Al-Deeb

Biology Department, United Arab Emirates University, Al Ain, UAE
Email: m_aldeeb@uaeu.ac.ae

Abstract

This technical note aims to show how any instructor teaching entomology can use the Basic Local Alignment Search Tool (BLAST) and the "one click" mode of Phylogeny.fr to teach undergraduate students about insect DNA similarity in a simple way. Teaching an entomology course requires the use of numerous tools to help students grasp different concepts. Knowing that there are more than one million described species of insects means that teaching students about insect identification and taxonomy can be challenging. However, here we present two easy exercises that could be used as classroom or take-home assignments to demonstrate various levels of DNA similarity among different insect taxa. Such exercises unlock students' creativity and break the barrier of fear of bioinformatics. Moreover, they open up new ways for them to understand insect taxonomy through molecular biology and allow them to develop new skills that contribute to strengthening their scientific performance in the future, especially when they do research as graduate students. Finally, this note is an example of how to integrate simple bioinformatics tools into the teaching of entomology.

Keywords

DNA Sequence, Phylogenetic Tree, GenBank, Teaching, Entomology

1. Introduction

Teaching entomology to students can be achieved by several methods. Undoubtedly, the conventional teaching method, which has limited innovation strategies had been the main method used so far by most teachers. However, some new and novel methods were proposed. For example, a group of scientists suggested

the use of games as a teaching aid and investigated the use of game-based activity as a tool for teaching entomology [1]. They found evidence supporting the usefulness of the game as a didactic activity in the classroom. In another study, the results showed that students were motivated to play a game and the learning test suggested an improvement in matters related to morphology, taxonomy, and the insect life cycle [2]. Furthermore, studies even suggested and investigated the teaching of entomology with postage stamps as a didactic resource [3]. No doubt, educational technologies have a role to play in teaching and learning entomology. Thus, an interactive album of veterinary acarology and entomology was presented and evaluated in a study, which revealed that students had better grades in a course with visualization technologies [4]. Overall, it should be emphasized that the instructional delivery method of any science, including entomology, can affect how students are exposed to new topics and their level of understanding.

Using technology in education helps students learn faster and make the learning process more enjoyable. The field of education was revolutionized by the use of technology, and with the onset of computers in education, teachers should encourage their students to acquire more knowledge [5]. Therefore, teaching insect identification and taxonomy with the help of *in silico* methods is always informative and an interesting experience for students. Moreover, such experience could be enhanced by the use of computers and online tools. On the other hand, mobile learning, which is the use of electronic devices to learn across contexts, connects individuals with media, educators, peers, and experts. Furthermore, mobile technologies can be employed to improve learning [6]. Thus, the teaching of entomology should benefit from mobile learning. Consequently, teachers can use DNA sequences and phylogenetic trees in the classroom. In addition, knowing how to interpret and reason from phylogenetic trees is a very important element of biology education in general [7]. A study tested the use of project-based learning (PBL) activity that was developed to teach phylogenetics for junior and senior-level biology students. The results indicated that the PBL approach helped the students and reinforced the theories and concepts which were taken previously along with incorporating bioinformatics [8]. In addition, another study emphasized the use of portable sequencing as a teaching tool in conservation and biodiversity research. It provided step-by-step protocols for a field training program that uses low-cost, portable devices to conduct direct genomics-based training [9]. Additionally, one study provided an educational guide for nanopore sequencing in the classroom. It proposed a model that bridges experimental and bioinformatics concepts using the Oxford Nanopore Technologies (ONT) sequencing platform [10]. However, it should be mentioned that many students are not familiar with DNA sequencing and feel that this topic is sophisticated. Therefore, practice is a key factor here to make things look easier over time. In addition, the role of a teacher is to make this topic more attractive to students. In this technical note, we demonstrate two exercises that help students learn how to use the Basic Local Alignment Search Tool (BLAST)

(https://www.ncbi.nlm.nih.gov/) and the "one click" mode of Phylogeny.fr (http://www.phylogeny.fr/simple_phylogeny.cgi) to learn insect DNA similarity levels.

2. Finding Insects' DNA Similarity Using BLAST

The first exercise uses the nucleotide BLAST. This tool helps students search in the GenBank database for DNA sequences that are similar to a DNA sequence that is provided to them by the teacher. Students should know that the GenBank has more tools besides BLAST and provides many services, however, the focus in this exercise is on BLAST only. The three most important things that a student needs to keep in mind during a BLAST search are: 1) to paste the DNA sequence or accession number in the "enter query sequence" window, 2) to make sure that the database is the standard database with the option nucleotide collection (nr/nt), and 3) to choose the "highly similar sequences (megablast)" option in the program selection. Usually, these are the default settings on the BLAST page. For example, a teacher may provide an accession number such as KJ022631 to students and ask them to write a report on the insect species of this accession number and show the DNA similarity between it and the most similar GenBank records. By pasting this accession number in the "enter query sequence" window, students will find out that this accession number is for a DNA sequence from the cytochrome c oxidase subunit 1 mitochondrial (cox1) gene extracted from a camel tick, *Hyalomma dromedarii*, that is found in the United Arab Emirates. In addition, students can get DNA sequences that are similar to this accession number in the GenBank and present a summary of the search results (Table 1). Furthermore,

Table 1. Summary of DNA similarity between a fragment of the cytochrome c oxidase subunit 1 mitochondrial gene (cox1) of the accession number KJ022631 and GenBank records using NCBI BLAST.

Scientific Name	Max Score	Total Score	Query Coverage	E value	Percent Identity	Accession Number
Hyalomma dromedarii	933	933	100%	0	100.00%	MT107484.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT107482.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT107481.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT093512.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT093511.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT093509.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT093505.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT066417.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT066416.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT062376.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MT040954.1
Hyalomma dromedarii	933	933	100%	0	100.00%	MK305816.1

students need to learn that a good match should have an "Expect" value (E-value), which is the number of hits one can find by chance, as zero or very close to zero (10^{-5} or lower). The lower the E-value is, the more significant the match. In addition, the "query coverage", which is the percentage of the input sequence (KJ022631 in this example) aligned to sequences of the GenBank, should be a high value, which means the target sequence in the database spans most or the whole query sequence. Also, the "percent identity", which is the extent (percent) to which the input sequence and the GenBank sequence have the same residues at the same positions in an alignment, should be 100% or close to 100% to provide species-level DNA similarity and the higher this percentage the better the match. In short, while working on the BLAST exercise students need to look for the highest values of the maximum score, total score, query coverage, and percent identity, except for the E-value where zero is the best score. Repeating this exercise several times during the semester builds a good level of confidence in students and encourages them to use the GenBank data without difficulty in an entomology course as well as in other courses.

3. Constructing a Simple Phylogenetic Tree

In the second exercise, we give the students several accession numbers (Table 2) and ask them to get their nucleotide sequences in FASTA format from the GenBank.

These accession numbers are for DNA sequences of the cox1 gene of eight insect species. Each DNA sequence can be retrieved easily from the GenBank by writing the accession number in the search window while keeping the default settings without change. Once a target record is found, its nucleotide sequence can be downloaded in FASTA format as a text file. In this example, students have eight files to download. First, students should download and save each file separately then merge the files into one file to be used later in making a phylogenetic tree. To do this, a text file (Microsoft NotePad file) can be created (Figure 1). Within the file, the accession number and the name of each insect

Table 2. GenBank accession numbers of DNA sequences of the cytochrome c oxidase subunit 1 mitochondrial gene (cox1) of eight insect species from four orders.

Species Name	Order	Accession Number
Apis mellifera	Hymenoptera	MN563103.1
Apis nigrocincta	Hymenoptera	MK880239.1
Musca domestica	Diptera	MZ513096.1
Lucilia porphyrina	Diptera	MW566109.1
Leptinotarsa decemlineata	Coleoptera	MW348766.1
Sitophilus oryzae	Coleoptera	MT872411.1
Aphis nerii	Hemiptera	MZ091378.1
Schizaphis graminum	Hemiptera	MT011383.1

>MN563103.1 Apis mellifera

GCTTTATGATCTGGAATACTAGGATCATCAATGAGACTTATTATTCGAATAGAATTAAGATCCCCAGGAT
CATGAATTAGCAATGATCAAATTTATAATACAATTGTTACTAGTCATGCATTCCTAATAATTTTTTTTAT
AGTTATACCATTTTTAATTGGAGGATTTGGAAATTGGCTTATTCCTTTAATACTAGGATCACCTGATATA

>MK880239.1 Apis nigrocincta

ATTGGAATTTTATATATTTTTTTAGCTTTATGATCAGGAATGTTAGGCTCATCAATAAGATTAATTATTC
GTATAGAATTAAGATCTCCTGGTTCTTGAATTAATAATGATCAAATTTATAATACAATTGTTACAAGACA
TGCATTTCTCATAATTTTTTTTATAGTAATACCATTTTTAATTGGAGGTTTTGGAAATTGATTAATTCCA

>MZ513096.1 Musca domestica

TGATTAGTTCCTTTAATATTAGGAGCTCCAGATATAGCATTCCCTCGAATGAATAATATAAGTTTTTGAT
TATTACCTCCTGCATTAACTCTATTATTAGTAAGAAGTATAGTAGAAAAGGGAGCTGGAACAGGTTGAAC
TGTTTATCCACCTTTATCATCAATTATTGCTCATGGTGGAGCTTCAGTTGATTTAGCTATTTTCTCTCTT

>MW566109.1 Lucilia porphyrina

AAAGATATTGGTACTTTATACTTTATTTTTGGAGCTTGATCCGGTATAATTGGAACTTCATTAAGAATTC
TAATTCGAGCTGAATTAGGACACCCTGGTGCACTAATTGGAGATGACCAAATTTATAATGTAATTGTTAC
AGCTCATGCTTTTATTATAATTTTCTTTATAGTAATACCAATTATAATTGGAGGATTTGGTAATTGATTA

>MW348766.1 Leptinotarsa decemlineata

CATAAAGATATTGGAACTTTATATTTTATTTTTGGTGTTTGAGCTGGAATATTAGGTACTTCTTTAAGAA
TTCTAATTCGGGCAGAATTAGGGAATCCTGGTACTTTAATTGGCAATGATCAAATTTATAATGTTATTGT
TACCGCTCATGCTTTTGTAATAATTTTTTTTATAGTTATGCCAATTATAATTGGGGGATTTGGTAATTGG

>MT872411.1 Sitophilus oryzae

TCAGGAATAGTAGGTACATCTTTAAGTTTGCTAATTCGGGCAGAACTAGGAAATCCTGGATCACTAATTG
GAAATGACCAAATTTATAATACTATTGTCACAGCACATGCATTCATTATAATTTTCTTTATAGTAATACC
AATTATAATTGGAGGATTTGGAAACTGATTAATCCCATTAATATTAGGAGCCCCAGATATAGCATTCCCC

>MZ091378.1 Aphis nerii

GGATCTTCTCTTAGAATTTTAATTCGACTAGAATTAAGTCAAATTAATTCAATTATTAATAATAATCAAT
TATATAATGTAATTGTTACAATTCATGCTTTCATTATAATTTTTTTTATAACAATACCAATTGTAATTGG
AGGATTTGGAAACTGGTTAATTCCTATAATAATAGGATGTCCAGATATATCTTTTCCACGATTAAATAAT

>MT011383.1 Schizaphis graminum

AAGATATTGGAACTTTATATTTTTTTATTTGGTATTTGATCAGGTATAATTGGTTCATCTTTAAGAATTTT
AATTCGACTTGAATTAAGACAAATTAATTCAATTATTAATAATAATCAATTATATAATGTAATCGTTACA
ATTCATGCTTTTATTATAATTTTTTTTATAACTATACCAATTGTTATTGGTGGTTTTGGAAATTGATTAA

Figure 1. Content of the text file with DNA sequences in FASTA format, which were downloaded from the GenBank to be used in making a phylogenetic tree. Sequences were shortened here (3 lines each) to fit in one page, however, in actual work each DNA sequence should be copied and pasted according to its length in the GenBank.

species should be preceded by the greater than (>) symbol.

Students do not need to be experts in phylogenetics to construct an informative phylogenetic tree. This is most simply done using an online tool such as Phylogeny.fr (http://www.phylogeny.fr/simple_phylogeny.cgi) using the "one click" mode [11]. The text file, which was already created and saved by each student, should now be opened using the "browse" option on Phylogeny.fr website

or by copying and pasting the file content (**Figure 1**) in the submission window. As soon as the data are submitted, the one-click mode will automatically perform sequence alignment, curation, phylogenetic analysis, and tree rendering. As a result, a phylogenetic tree is produced. Under the "display" options, students can select "Display branch support values in %" because using percentages is always much easier to be understood. The tree can be downloaded (**Figure 2**) in different formats by first clicking either PNG, PDF, SVG, or Newick option and then by right-clicking on the tree image to save it using the "save image as" option. This tree gives the students a visual demonstration of the degree of DNA similarity among the eight sequences. For instance, the tree shows that the two species of bees (*Apis mellifera* and *Apis nigrocincta*) were in one cluster and have the highest branch support value (100%) compared to the other species. The same can be said almost about all the other insect species that belong to one order on the tree. In short, this exercise shows students how insects in one order have a higher degree of DNA similarity compared to insects from other orders. By following this simple workflow (**Figure 3**), different exercises could be created as classroom or take-home assignments to demonstrate various levels of DNA similarity in different insect taxa at any level. One point that should be mentioned here is that students need to know that DNA-based insect identification in some cases does not lead to the right species name and consequently, it should be supplemented by morphology and this is known as integrative taxonomy [12]. Some people may consider that searching for an accession number in the GenBank using BLAST is a very simple task and does not offer much to students. However, over the years we found that it is exactly the opposite because this search breaks the barrier of fear of bioinformatics and opens up new ways for students to understand some concepts in entomology through molecular biology, and allows them to develop new skills that contribute to strengthening their scientific performance in the future, especially when they do research. Furthermore, constructing a simple phylogenetic tree can give students a basic understanding of the fact that there are common genes among living things, including insects, which can be used to show the degree of relatedness among living things. In addition, this understanding will be the basis on which students
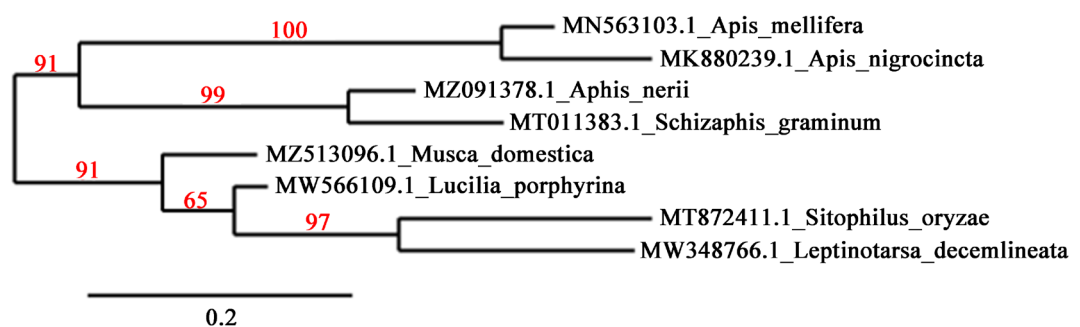

Figure 2. Phylogenetic tree constructed using DNA sequences of the cox1 gene of eight insect species from four orders. Red numbers are the branch support values presented as percentages. The analysis was performed using Phylogeny.fr (http://www.phylogeny.fr).
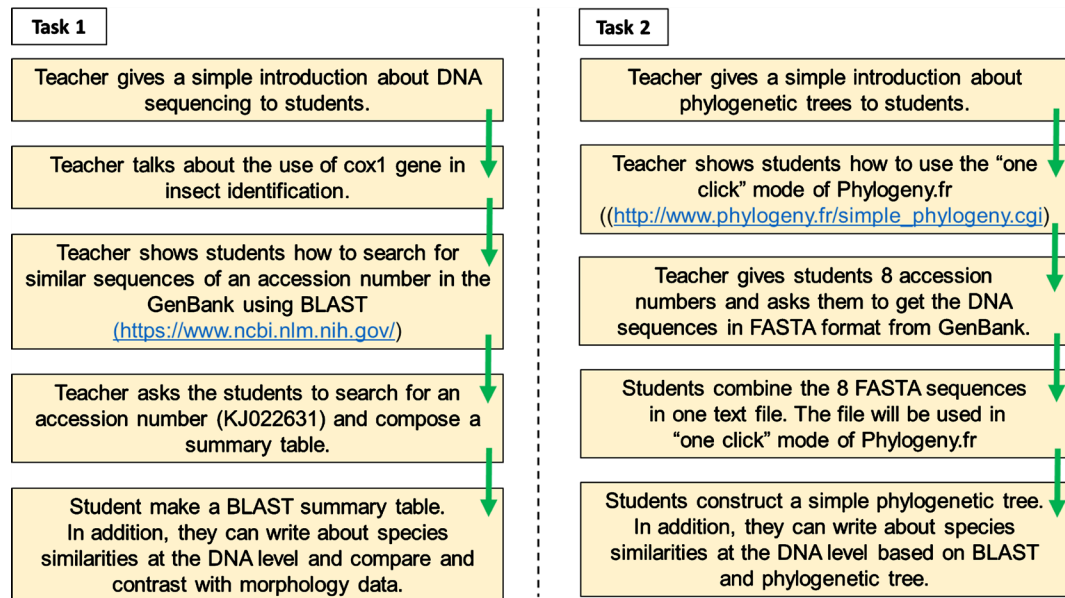
Figure 3. Summary of the workflow of the two tasks. Task 1 shows how students are given an accession number to find similar sequences in the GenBank using BLAST. Task 2 shows how students can make a simple phylogenetic tree with eight insect DNA sequences using Phylogeny.fr (http://www.phylogeny.fr).

build deeper awareness of insect taxonomy.

4. Conclusion

To conclude, using DNA sequences of the cox1 gene from the GenBank and the "one click" mode of Phylogeny.fr to construct simple phylogenetic trees are good tools for teaching entomology to undergraduate students. Moreover, both the teacher and the students do not need to be experts in phylogenetics to benefit from these tools and enjoy learning about DNA similarities in insects. In addition, besides aiding the learning of entomology these tools expand students' scientific horizons and unlock creativity.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Cosme Jr., L., Turchen, L.M. and Guedes, R.N.C. (2020) Insect World: Game-Based Learning as a Strategy for Teaching Entomology. *The American Biology Teacher*, **82**, 210-215. https://doi.org/10.1525/abt.2020.82.4.210

[2] Wommer, F.G.B., Sepel, L.M.N. and Loreto, E.L.S. (2021) Insects GO: A Gaming Activity for Entomology Teaching in Middle School. *Research in Science & Technological Education*, **39**, 1-15. https://doi.org/10.1080/02635143.2021.1921724

[3] Turienzo, P. (2018) Teaching Entomology with Postage Stamps as Didactic Resource. *IDESIA*, **36**, 119-129.

[4] Vega-Garzón, J.C., Robayo-Sánchez, L.N., Cruz-Maldonado, O.A. and Cortés-Vecino, J.A. (2021) Visualization Technologies for Learning and Teaching Veterinary Aca-

rology and Entomology. *Journal of Veterinary Medical Education*, **48**, e20200034. https://doi.org/10.3138/jvme-2020-0034

[5] Raja, R. and Nagasubramani, P.C. (2018) Impact of Modern Technology in Education. *Journal of Applied and Advanced Research*, **3**, 33-35. https://doi.org/10.21839/jaar.2018.v3iS1.165

[6] Bernacki, M.L., Greene, J.A. and Crompton, H. (2020) Mobile Technology, Learning, and Achievement: Advances in Understanding and Measuring the Role of Mobile Technology in Education. *Contemporary Educational Psychology*, **60**, 101827. https://doi.org/10.1016/j.cedpsych.2019.101827

[7] Dees, J., Bussard, C. and Momsen, J.L. (2018) Further Effects of Phylogenetic Tree Style on Student Comprehension in an Introductory Biology Course. *CBE-Life Sciences Education*, **17**, ar17. https://doi.org/10.1187/cbe.17-03-0058

[8] David, A.A. (2018) Using Project-Based Learning to Teach Phylogenetic Reconstruction for Advanced Undergraduate Biology Students: Molluscan Evolution as a Case Study. *The American Biology Teacher*, **80**, 278-284. https://doi.org/10.1525/abt.2018.80.4.278

[9] Watsa, M., Erkenswick, G.A., Pomerantz, A. and Prost, S. (2020) Portable Sequencing as a Teaching Tool in Conservation and Biodiversity Research. *PLoS Biology*, **18**, e3000667. https://doi.org/10.1371/journal.pbio.3000667

[10] Salazar, A.N., Nobrega, F.L., Anyansi, C., Aparicio-Maldonado, C., Costa, A.R., Haagsma, A.C., Hiralal, A., Mahfouz, A., McKenzie, R.E., van Rossum, T. and Brouns, S.J. (2020) An Educational Guide for Nanopore Sequencing in the Classroom. *PLoS Computational Biology*, **16**, e1007314. https://doi.org/10.1371/journal.pcbi.1007314

[11] Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M. and Claverie, J.M. (2008) Phylogeny. fr: Robust Phylogenetic Analysis for the Non-Specialist. *Nucleic Acids Research*, **36**, W465-W469. https://doi.org/10.1093/nar/gkn180

[12] Almansoori, N., Enan, M.R. and Al-Deeb, M.A. (2019) Barcode of Life Data Systems (BOLD) versus GenBank Molecular Identification of a Dragonfly from the UAE in Comparison to the Morphological Identification. *OnLine Journal of Biological Sciences*, **19**, 237-244. https://doi.org/10.3844/ojbsci.2019.237.244