

Using Phylogenetic Analysis to Detect Market Substitution of Atlantic Salmon for Pacific Salmon: An Introductory Biology Laboratory Experiment

ERICA CLINE, JENNIFER GOGARTEN

ABSTRACT

We describe a laboratory exercise developed for the cell and molecular biology quarter of a year-long majors' undergraduate introductory biology sequence. In an analysis of salmon samples collected by students in their local stores and restaurants, DNA sequencing and phylogenetic analysis were used to detect market substitution of Atlantic salmon for Pacific salmon. This allowed students to apply molecular methods such as polymerase chain reaction (PCR) and DNA sequencing to a socially relevant issue.

Key Words: Undergraduate introductory biology education, salmon market substitution, phylogenetic analysis, DNA sequencing.

In the introductory biology curriculum, molecular and cellular biology present a unique challenge; the technology in these fields is advancing so quickly that the laboratory experience often lags behind the lecture material. As a result, many instructors struggle to maintain a compelling connection between lecture information and laboratory exercises. There is a tendency to emphasize basic techniques over concepts (the gel electrophoresis lab, the PCR lab) or to (re-)confirm widely known information (such as visualizing DNA in cells) rather than place the emphasis on true inquiry, modeling the process of the scientific method. A recent "call to action" by the American Association for the Advancement of Science (AAAS, 2011) emphasizes the importance of inquiry-based exercises at the introductory level as one of the top priorities for engaging students in science. As Alberts (2009) argued, we must engage students in authentic scientific inquiry from the very beginning of their college experience (if not earlier).

Student engagement can also be enhanced by taking inquiry outside the classroom and connecting students to their local neighborhood, community, or broader regional issues. Projects that stimulate students to consider the broader societal context and address

questions that have real-world implications can transform the student experience because they are not just learning about science, they are actually doing science, and thereby becoming part of the scientific community. In our experience, projects that connect students to their local community engage them at a much more visceral level.

We developed this laboratory exercise to allow students to apply DNA sequencing and phylogenetic analysis to the socially relevant goal of detecting market substitution of Atlantic salmon, which is mostly farmed, for Pacific salmon, which is mostly wild-caught. Pacific salmon include six closely related species within the genus *Oncorhynchus*: chum (*Oncorhynchus keta*), coho (*O. kisutch*), chinook (*O. tshawytscha*), pink (*O. gorbuscha*), sockeye (*O. nerka*), and Japanese cherry (*O. masou*). Of the Pacific salmon, only coho and chinook are farmed, and only on a limited scale, whereas >90% of farmed salmon are Atlantic salmon (*Salmo salar*; see http://www.fao.org/fishery/culturedspecies/Salmo_salar).

Market substitution of seafood is surprisingly common; for example, using DNA sequence analysis, Wong and Hanner (2008) found that 25% of seafood samples marketed in the United States were potentially mislabeled. A high school independent project used similar methods to detect market substitution in 23% of seafood samples from Manhattan restaurants and stores (Stoeckle & Strauss, 2008). A Bellingham, Washington, fish processor was recently sentenced to prison and a substantial fine for substituting >160,000 lbs of coho for king salmon over a 2-year period (Durkan, 2011), but this is the exception; in most cases, market substitution is undetected and unpunished. Market substitution can be harmful to society on multiple levels: (1) financial losses of consumers paying too much for the product (consumer fraud), (2) reduced purchases of environmentally sustainable wild salmon in favor of environmentally damaging farmed salmon, and (3) undermining of consumer confidence in the accuracy of eco-labeling (Jacquet & Pauly, 2008).

As a result, many instructors struggle to maintain a compelling connection between lecture information and laboratory exercises.

Table 1. Scientific and common names of Pacific salmon and close relatives.

Latin Name	Common Names
<i>Oncorhynchus nerka</i>	Sockeye, red, redfish, or blueback salmon; kokanee
<i>O. tshawytscha</i>	Chinook, king, chub, black, blackmouth, tyee, spring, springer, winter, ivory, Columbia River, hook bill, or quinnat salmon
<i>O. keta</i>	Chum, calico, dog, silverbrite, or keta salmon
<i>O. kisutch</i>	Coho, silver, or silverside salmon
<i>O. gorbuscha</i>	Pink or humpback salmon; humpies
<i>O. masou</i>	Cherry, Japanese, Masu, Taiwanese, or Formosan salmon; seema.
<i>O. mykiss</i>	Steelhead trout or rainbow salmon
<i>O. clarkii</i>	(Coastal/sea run) cutthroat or harvest trout

Farmed salmon negatively affect the environment through nitrogen and phosphorus contamination from feed and fish waste (Wu, 1995). Farmed salmon have up to 10× higher concentrations of fat-soluble organic pollutants such as dioxins, polychlorinated biphenyls (PCBs), and polybrominated diphenyl ethers (PBDEs) (Hites et al., 2004a, b; Hamilton et al., 2005; Hayward et al., 2007). Salmon farms can have detrimental effects on native salmon populations by spreading parasites and diseases (McVicar, 1997), and escaped Atlantic salmon can interbreed with Pacific salmon, and compete directly with native fish for space and food (Gross, 1998). Some of these impacts could be minimized through more sustainable farming practices (e.g., Leschin-Hoar, 2010).

Our motivation in developing this laboratory exercise was to allow students to apply phylogenetic analysis techniques to a socially relevant question, market substitution, and thereby increase student engagement and promote learning of key concepts in cell and molecular biology.

○ Learning Objectives

1. To help the student master the fundamental concepts of cell and molecular biology with application to a unifying theme and a socially relevant project.
2. To provide a context for understanding the utility of molecular biology applications to an environmental issue by using these techniques to analyze samples from their local community.

Table 2. Lab modules and required time, equipment, and expenses.

Lab Exercise	Minimum Time Required	Equipment & Supplies	Consumable Expenses
DNA extraction and PCR	2.5 hours, with PCR run started at the end of the lab period	Microcentrifuge, vortexer, DNA extraction kit, thermal cyclor	Extraction kit, \$2/sample; PCR reagents, \$1/sample
Gel electrophoresis and setup sequencing	2.5 hours	Gel boxes and power supplies, agarose, DNA standards, 96 well plates	\$3/sample
Phylogenetic analysis	2.5 hours	One computer per student; free online software	None

○ Materials & Methods

Small plastic Petri dishes were provided to each student, along with detailed instructions for recording information about each sample on a pre-lab assignment sheet due at the beginning of the lab period. Students were asked to obtain one sample of either *O. nerka* or *O. tshawytscha*, and one sample of *O. keta*, *O. kisutch*, or *O. gorbuscha*, to ensure a broad assortment of species in the resulting phylogenetic tree (Table 1). These requirements might need to be relaxed in regions in which Pacific salmon is less available than in the Pacific Northwest. Nevertheless, even if some of the Pacific salmon species are not sampled, those branches will still be visible because of the standard reference sequences, so this will not detract from the educational value of the phylogenetic tree. Most students were able to obtain free samples from stores and restaurants, since this procedure required <1 g of salmon per sample. Students were encouraged to examine the salmon muscle tissue under a microscope, record detailed observations, and compare their samples with other students' samples of the same species to look for obvious differences in color and other features.

○ Laboratory 1: DNA Extraction & Polymerase Chain Reaction

Equipment and reagents required for DNA extraction and polymerase chain reaction (PCR) are listed with estimated costs per sample (Table 2). To avoid cross-contamination, students used sterile scalpels – cleaned carefully between samples – to excise an interior section of salmon muscle tissue from each sample. The MoBio UltraClean Tissue DNA extraction kit (<http://www.mobio.com>) was used, following the kit directions. Standard fish DNA barcoding primers were ordered from Invitrogen and used to amplify a section of the mitochondrial cytochrome oxidase I (COI) gene; the forward (sense) primer was FishF1 (5' TCAACCAACCACAAAGA-CATTGGCAC 3'), and the reverse (antisense) primer was Fish R1 (5' TAGACTTCTGGGTGGCCGAAGAATCA 3') (Ward et al., 2005). We selected this gene because the DNA Barcode of Life (BOL) project has an extensive database of COI sequences from fish (<http://www.fishbol.org>; Ratnasingham and Hebert, 2007).

The PCR master mix was made up as a lab demo. Reactions were performed in 50-µL volumes containing final reaction concentrations of Promega GoTaq Flexi DNA Polymerase (2.5 U), MgCl (2.5 mM), 1X buffer, Promega dNTPs (0.2 mM), Fish F1 primer (25 nM), and FishR1 primer (25 nM). Each student diluted his or her DNA extracts 1:200 with DNase-free water and added 25 µL of diluted DNA extract to 25 µL of the master mix. The PCR

reaction conditions were as follows: 96°C for 5 minutes, 42 cycles of 94°C for 60 seconds, 51°C for 45 seconds, and 72°C for 60 seconds, followed by a final extension of 72°C for 5 minutes. The PCR reactions were performed using a ThermoElectron Px2 thermal cycler.

○ Laboratory 2: Gel Electrophoresis & DNA Sequencing

Prepoured mini-gels were provided to groups of three or four students, with 1% agarose (Fisher Scientific) in Tris-Boric Acid-EDTA buffer (Fisher Scientific), and 1:10,000 diluted SybrSafe dye (Invitrogen). The SybrSafe dye is as sensitive as the traditional ethidium bromide stain visualized using UV light, but preferable for student use because of the low toxicity of the stain and the nonharmful blue light. Each student loaded 5 μ L of PCR product from each of his or her samples; because loading dye was included in the PCR mix, it was not necessary to add it as a separate step. Gels were electrophoresed at 80 V for 30 minutes, then examined on a blue-light DarkReader transilluminator (Clare Chemical Research) using a Kodak GelLogic 100 Imaging System and/or a blue light transilluminator and gel visualizing system (Vernier). The Vernier system has the advantage of being portable and is adequate for student use without the greater expense of a research-quality system. An example of a typical student gel is provided in Figure 1.

Successful PCR products (i.e., a single clear band between 500 and 800 base pairs in size) were added to 96-well plates (Fisher) and submitted, along with the FishF1 sequencing primer, to the University of Washington High-Throughput Sequencing facility (<http://www.htseq.org>). We hand delivered the plates, but the facility also accepts mailed plates. Although the sequencing facility recommends optimizing the amount of PCR product, we have had good success with omitting this adjustment, instead submitting 10 μ L of undiluted PCR product regardless of the intensity of the gel band, to save class time. A strong band on the gel corresponds to \sim 1 μ g of DNA, whereas a weak band is \sim 100 ng of DNA. Pricing for post-PCR cleanup and sequencing was \$282/plate,

a flat rate for any number of samples up to 96. This was sufficient for several samples per student, in a single or double laboratory section. The sequencing facility can also perform PCR reactions for an additional \$155/plate, which would greatly reduce the required laboratory time and necessary equipment, making the experiment feasible for institutions that lack a thermal cycler. This might make the experiment suitable for adoption at a wider range of institutions, including high schools.

○ Laboratory 3: Phylogenetic Analysis

Sequencing results from the sequencing facility were viewed using the free software package FinchTV (<http://www.geospiza.com>); a representative sequence trace is shown in Figure 2. Students were instructed to curate sequences appropriately, deleting regions of low-quality and rectifying ambiguities. High-quality curated sequences were then shared in a communal document, added to reference sequences identified by searches of NCBI using the BLAST algorithm (Altschul et al., 1997). Alternatively, in some quarters when time was limited, the instructor compiled all sequences from the entire class into a single FASTA file provided directly to the students. Students generated a sequence alignment using the free software package ClustalW (Larkin et al., 2007) and a bootstrapped neighbor-joining phylogenetic tree using the free software package MEGA5 (Tamura et al., 2007). The results were summarized by each group in an informal laboratory report (due at the end of the lab period) and a formal reflection paper (due 1 week later).

○ Assessment

Student learning was assessed through informal and formal laboratory reports, quizzes, and essay exams. We also asked students to assess their own learning in two ways: first through surveys, and second through graded reflection papers. Anonymous surveys were administered in the final week of the quarter; students were free to decline to participate. (This research has been approved as exempt by the University of Washington Institutional Review Board, IRB no. 33977.)

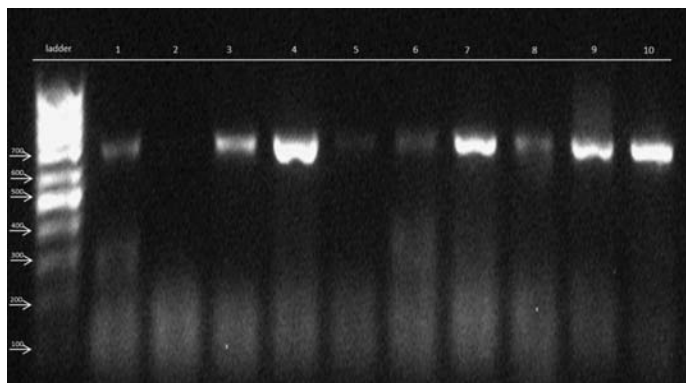


Figure 1. An example of an agarose gel showing typical student results from PCR reactions (samples 1–10); the first lane contains a low-molecular-weight DNA marker mix (“ladder”).

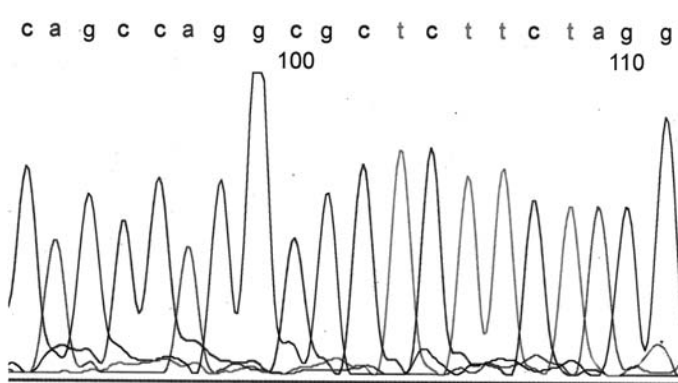


Figure 2. An example of a portion of a typical sequence trace, viewed with FinchTV. Individual peaks correspond to the DNA bases (A, G, C, T).

○ Results

Student Project Results

Over a 2 year period (2009 to winter 2011), we have used this exercise with four single-section lab courses. Overall, 53% of student samples yielded clear DNA sequence results. Of the 81 samples for which we had interpretable sequence data, 11 (14%) were labeled as wild Pacific salmon but proved to be Atlantic salmon (Figure 3). The substitution rate was much higher in restaurant samples: 29% compared to only 2% of grocery store samples (Figure 3). Eight Pacific salmon samples were substituted for another Pacific salmon species; again, the substitution rate was higher in restaurants than in stores (18% vs. 5%, respectively). Of these, four were a cheaper fish substituted for a more expensive species, and the other four were substitutions of fish of approximately the same value. An example phylogenetic tree from a recent class is shown in Figure 4.

○ Assessment

In assessing student learning through informal and formal laboratory reports, quizzes, essay exams, and a reflection paper, we focused on three key areas of students' understanding: (1) the mechanics of the procedures, linked to lecture concepts such as the structure of DNA and the mechanism of DNA replication; (2) how molecular methods can be applied; and (3) basic principles of phylogenetic analysis. We consistently witnessed an increase in the depth and sophistication of the students' answers and explanations within these three key areas as they carried out the project.

In anonymous surveys, students consistently ranked the salmon project as highly effective in promoting their learning of key concepts: 71% of students listed the salmon project as the lab experience that contributed most to their learning in the class, and 80% considered it useful or highly useful for "learning fundamental concepts in cell and molecular biology and synthesizing and applying these concepts in novel situations." The full surveys can be copied and adopted from the Student Assessment of their Learning Gains website (<http://www.salgsite.org>; search for "TESC130 Intro Bio 2" for recent examples of pre- and posttests).

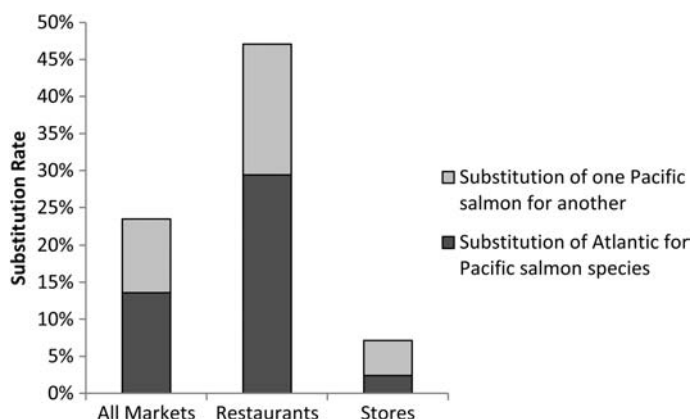


Figure 3. Rates of substitution of Atlantic salmon for Pacific salmon, and one species of Pacific salmon for another, in restaurants or stores, based on phylogenetic analysis, from winter 2009 to winter 2011.

The students' reflections captured the excitement they felt at engaging in true inquiry rather than a "cookbook" project, collecting their own samples, and participating in an integrated sequence of labs building toward an end-of-quarter climax (Box 1).

○ Discussion

Over the past 2 years that we have used this laboratory exercise, we have found it to be highly engaging and effective in promoting student learning in introductory biology courses. To help other faculty adopt this laboratory exercise, we have provided a rich array of supporting materials, including student lab handouts, an instructor manual, and sample data sets, freely downloadable and ready for modification to fit the instructor's needs, as online supplemental materials.

These laboratory exercises have been successfully used as the cornerstone of a college introductory biology course for science majors but could easily be adapted to an advanced high school biology course. We have found salmon to provide a particularly powerful unifying theme and a connection point for a series of learning modules related to multiple content areas throughout the biology sequence. During the evolution section, connections include phylogenetic tree-building algorithms and statistical analysis, and population genetics of salmon. During the animal physiology section, muscle physiology and osmoregulation changes from saltwater to freshwater can be incorporated. During the cell and molecular biology sections of the biology sequence, the labs provide a hands-on introduction to basic methods in molecular biology. During ecology, the bioaccumulation of toxins and environmental consequences of farming versus fishing wild animals can be discussed. The salmon theme resonates with students because it involves a universal need (food) and brings the techniques of molecular biology out of the classroom and into the local community by allowing students to test their local stores and restaurants. This community monitoring can provide a valuable public service. For example, results of our study were published by several newspapers (Blankinship, 2011; Gillie, 2011), and in the scientific literature (Cline, 2012), enhancing public awareness of this issue.

The potential for using this information to accuse local stores or suppliers of consumer fraud reinforces for the students the importance of keeping track of the identity of their samples and thoroughly documenting where the samples came from. It is important to note that the PCR procedure is highly vulnerable to cross-contamination. Therefore, it is essential to use careful lab technique at all times by

- stressing accountability and documentation of the sample collection,
- cleaning utensils between samples and sampling only from interior tissue, and
- storing all samples in a freezer for later confirmation of results.

○ Accessing Materials

An instructor handbook containing detailed instructions for preparation and materials required for each modular lab exercise, a student lab manual, including prelab and postlab questions, introduction and background information, and detailed instructions for each modular lab exercise, and a database of COI sequences and reference sequences are freely available at <https://sites.google.com/a/uw.edu/erica-cline-uw-faculty-webpage/home/syllabi-and-teaching-materials>.

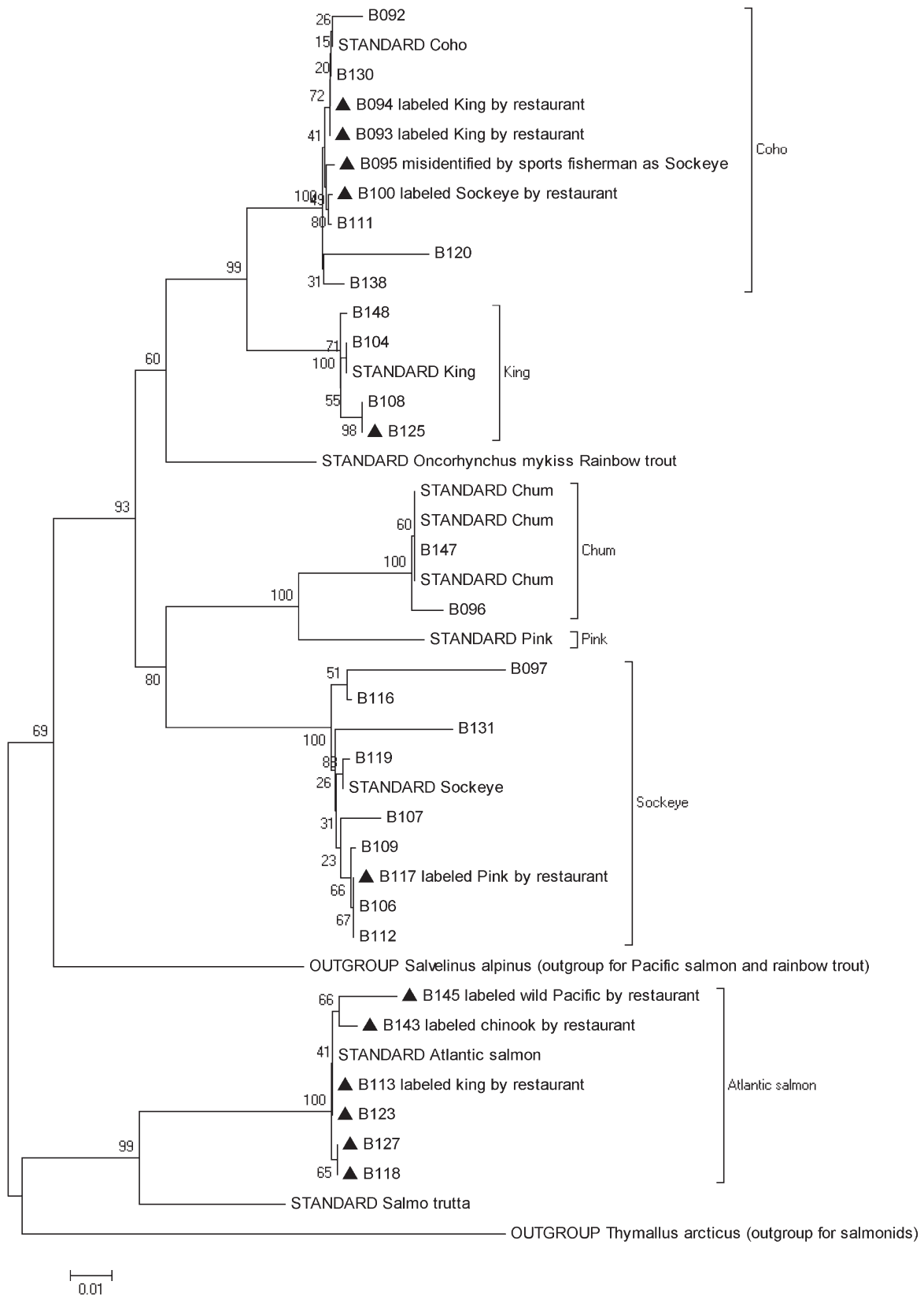


Figure 4. Example of a neighbor-joining tree constructed using MEGA5 (Tamura et al., 2007) based on sequencing of a 645-base-pair region of the COI gene by a single-section class of 20 students from winter 2011. Reference sequences from GenBank (taxa labeled as STANDARD) are compared with COI sequences generated by the University of Washington, Tacoma, TESC130 students (taxa labeled with B#). Bootstrap values (with 500 iterations) are included as node labels for genus and species clades. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Erroneously labeled species are labeled with the name of the species under which they were sold.

Box 1. Student reflections.

I really liked the continuity of working on the project over a long series of labs that pulled in many different skills and types of lab work. This lab was a great tie-in to our study of DNA this quarter and gave us a real-world example of the value and use of what we were learning. . . . I also liked testing the local stores where I sometimes buy fish and was happy to see that they are not cheating me.

The best part of the “cheaters” exercise was seeing all of the data laid out in the phylogenetic tree. Seeing the final product of weeks’ worth of labor made me feel like I had actually done real science.

Putting together a phylogenetic tree using the sequencing definitely gave me a better understanding of how they work. [T]o see the tree fit together as a result of hands-on experience just kind of helps solidify the concept more. It also gave me a better understanding of how molecular biology can be applied to real-world issues.

I even took my lab report and showed it to all my friends at work.

○ Acknowledgments

We thank Jean MacGregor (Curriculum for the Bioregion) for supporting the development of the teaching materials, Amanda Watts (lab assistant), and the students of TESC130 Introductory Biology II for their diligence and enthusiasm.

References

- AAAS. (2011). Vision and change in undergraduate biology education: a call to action. Available online at <http://visionandchange.org/viewfinalreport/>.
- Alberts, B. (2009). Redefining science education. *Science*, 323, 437.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25, 3389–3402.
- Blankinship, D. (2011). It’s not the salmon you thought it was. *Oregonian*, 9 July, B5.
- Cline, E.T. (2012). Marketplace substitution of Atlantic salmon for Pacific salmon in Washington State detected by DNA barcoding. *Food Research International*, 45, 388–393.
- Durkan, J.A. (2011). Fish processor sentenced to prison for selling falsely labeled salmon. U.S. Department of Justice briefing, Western district of Washington, 1 April.
- Gillie, J. (2011). Salmon often mislabeled. *News Tribune* (Tacoma, WA), 7 July, A1.
- Gross, M.R. (1998). One species with two biologies: Atlantic salmon (*Salmo salar*) in the wild and in aquaculture. *Canadian Journal of Fisheries and Aquatic Sciences*, 55 (Supplement 1), 1–14.

- Hamilton, M.C., Hites, R.A., Schwager, S.J., Foran, J.A., Knuth, B.A. & Carpenter, D.O. (2005). Lipid composition and contaminants in farmed and wild salmon. *Environmental Science & Technology*, 39, 8622–8629.
- Hayward, D., Wong, J. & Krynitsky, A.J. (2007). Polybrominated diphenyl ethers and polychlorinated biphenyls in commercially wild caught and farm-raised fish fillets in the United States. *Environmental Research*, 103, 46–54.
- Hites, R.A., Foran, J.A., Carpenter, D.O., Hamilton, M.C., Knuth, B.A. & Schwager, S.J. (2004a). Global assessment of organic contaminants in farmed salmon. *Science*, 303, 226–229.
- Hites, R.A., Foran, J.A., Schwager, S.J., Knuth, B.A., Hamilton, M.C. & Carpenter, D.O. (2004b). Global assessment of polybrominated diphenyl ethers in farmed and wild salmon. *Environmental Science & Technology*, 38, 4945–4949.
- Jacquet, J.L. & Pauly, D. (2008). Trade secrets: renaming and mislabeling of seafood. *Marine Policy*, 32, 309–318.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H. & others. (2007). ClustalW and ClustalX version 2. *Bioinformatics*, 23, 2947–2948.
- Leschin-Hoar, C. (2010). Sea change: environmental group gives first-time nod to sustainable salmon-farming method. *Scientific American*, January, p. 23.
- McVicar, A.H. (1997). Disease and parasite implications of the coexistence of wild and cultured salmon populations. *ICES Journal of Marine Science*, 54, 1093–1103.
- Ratnasingham, S. & Hebert, P.D.N. (2007). Bold: the barcode of life data system. *Molecular Ecology Notes*, 7, 355–364.
- Stoeckle, K. & Strauss, L. (2008). High school students track down fish fraud. *Pacific Fishing*, September, p. 34.
- Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24, 1596–1599. [Publication PDF at <http://www.kumarlab.net/publications>; software at <http://www.mega.org>.]
- Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R. & Hebert, P.D.N. (2005). DNA barcoding Australia’s fish species. *Philosophical Transactions of the Royal Society of London, Series B*, 360, 1847–1857.
- Wong, E.H.-K. & Hanner, R.H. (2008). DNA barcoding detects market substitution in North American seafood. *Food Research International*, 41, 828–837.
- Wu, R.S.S. (1995). The environmental impact of marine fish culture: towards a sustainable future. *Marine Pollution Bulletin*, 31, 159–166.

ERICA CLINE is Assistant Professor of Interdisciplinary Arts and Sciences in the Environmental Sciences Program at the University of Washington Tacoma, Box 358436, Tacoma, WA 98402; e-mail: ecline@uw.edu. JENNIFER GOGARTEN is a Lecturer of Interdisciplinary Arts and Sciences in the Environmental Sciences Program at the University of Washington Tacoma, Box 358436, Tacoma, WA 98402; e-mail: gogarten@uw.edu.