Letter to the Editor

Integrating Genomics Research throughout the Undergraduate Curriculum: A Collection of Inquiry-Based Genomics Lab Modules

Lois M. Banta,*[†] Erica J. Crespi,[†‡§] Ross H. Nehm,[†‖] Jodi A. Schwarz,[†‡] Susan Singer,[†¶] Cathryn A. Manduca,[†#] Eliot C. Bush,[@] Elizabeth Collins,[‡] Cara M. Constance,** Derek Dean,* David Esteban,[‡] Sean Fox,[#] John McDaris,[#] Carol Ann Paul,[††] Ginny Quinan,[††] Kathleen M. Raley-Susman,[‡] Marc L. Smith,[‡‡] Christopher S. Wallace,[§§] Ginger S. Withers,[§§] and Lynn Caporale[‖‖]

*Department of Biology, Williams College, Williamstown, MA 01267; [‡]Department of Biology, Vassar College, Poughkeepsie, NY 12604; [‖]College of Education and Human Ecology, Ohio State University, Columbus, OH 43210; [¶]Department of Biology and [#]Science Education Resource Center, Carleton College, Northfield, MN 55057; [@]Department of Biology, Harvey Mudd College, Claremont, CA 91711; **Department of Biology, Hiram College, Hiram, OH 44234; [††]Department of Biological Sciences, Wellesley College, Wellesley, MA 02481; [‡‡]Computer Science Department, Vassar College, Poughkeepsie, NY 12604; [§§]Department of Biology, Whitman College, Walla Walla, WA 99362; [‖‖]Consultant; caporale@usa.net

We wish to let *CBE—Life Sciences Education* readers know about a portal to a set of curricular lab modules designed to integrate genomics and bioinformatics into commonly taught courses at all levels of the undergraduate curriculum. Through a multi-year, collaborative process, we developed, implemented, and peer-reviewed inquiry-based, integrated instructional units (I^3Us) adaptable to a range of teaching settings, with a focus on both model and nonmodel systems. Each of the products is built on vetted design principles: 1) they have clear pedagogical objectives; 2) they are integrated with lessons taught in the lecture; 3) they are designed to integrate the learning of science content with learning about the process of science; and 4) they require student reflection and discussion (Figure 1; National Research Council [NRC], 2005). Eleven I^3Us were designed and implemented as multi-week modules within the context of an existing biology course (e.g.,

microbiology, comparative anatomy, introduction to neurobiology), and three I^3Us were incorporated into interdisciplinary biology/computer science classes. Our collection of
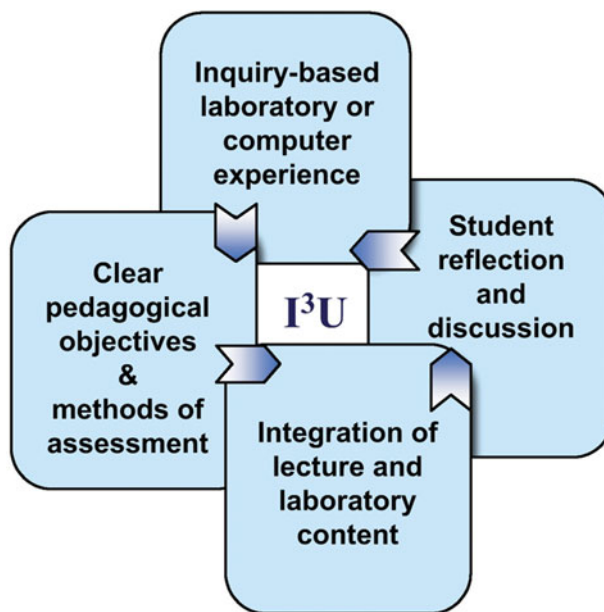
Figure 1. Pedagogical elements of the I^3U, which was based on the findings of America's Lab Report (NRC, 2005) and was used as the primary curricular design framework for this project.

Comparison of a Highly Polymorphic Olfactory Receptor Gene Subfamily in Genetically Diverse Dog Breeds

Lois Banta, Norm Bell, and Duane Bailey
Williams College
Williamstown, MA 01267

Summary

In this three or four week project, students learn about single nucleotide polymorphisms (SNPs) by amplifying and generating sequence data on a highly polymorphic gene subfamily in a diverse population of subjects (dogs) with which many students have considerable familiarity and affinity. In the first week, students make use of previously acquired knowledge of phylogenetic relationships and experience in sequence alignment to design primers specific for one subfamily of canine olfactory receptor genes. In the second week, each student uses his/her primers in a polymerase chain reaction (PCR) to amplify the corresponding DNA from one dog's cheek cells. During the following lab, PCR products are purified and the yield is confirmed on a gel. In the final week, commercial or in-house sequencing is used to determine the sequence of the PCR product. The data analysis draws on a published microsatellite genotype-based population structure of 85 domestic dog breeds, allowing the students to compare a phylogenetic tree estimated from a single gene with data obtained through a genome-wide analysis. An optional bioinformatics module introduces existing web resources to predict transmembrane domains and/or provides students with a short programming assignment in which they write a Perl script to perform this analysis on an olfactory receptor sequence.

Learning Goals

- Students will put into practice skills they have learned in previous bioinformatics labs, by analyzing and generating phylogenetic trees and by performing sequence alignments.

- Students will appreciate the difficulties in obtaining sequence data for an individual gene in a highly conserved family and the value of single-molecule sequencing strategies.

- Students will gain an understanding of the role of single nucleotide polymorphisms in gene family structure, the evolutionary processes that lead to genetic duplication and diversification, and the limitations in correlating function with specific SNPs.

- Students without prior wet-lab experience will gain familiarity with the rudiments of molecular biology tools and techniques in a relatively straightforward protocol.

Context for Use

This activity is part of an upper-level elective (12-14 students) in genomics and bioinformatics. It could also be used in a molecular evolution or a genetics course to integrate with classroom-based lessons on SNPs, PCR and/or DNA sequencing. The adaptability to larger classes is limited only by the availability of gel electrophoresis equipment and the cost of the PCR reagents and sequencing; students could also work in pairs to reduce expenses. Ideally, faculty would have access to as many different dogs as there are students or pairs of students; the larger the data set the more interesting the data analysis. However, multiple students or groups could also use the same dog's cheek cell sample. The lab presupposes no previous hands-on experience in molecular biology, and the only wet-lab manipulations (Weeks 2-3) involve setting up a PCR reaction, purifying the product, and running an agarose gel to confirm the yield. This lab was designed specifically to be a good first introduction to molecular biology, simple gel electrophoresis, and the use of pipettors for computer science, chemistry, physics and math students. The in silico investigation (week 1) builds on prior experience with Clustal and phylogenetic tree analysis, and the final data analysis reinforces the students' experience with Phylip or another tree estimation program (although Clustal can also be used to generate the neighbor-joining tree). The DNA sequencing can be performed in house, if the institution has its own sequencer, but the PCR products can also be sequenced commercially. The data analysis can be completed outside of lab.

Description and Teaching Materials

Week 1: Instructor introduces project, students design primers (computer lab needed)

- Behind the scenes between weeks 1 and 2: Instructor orders primers

Week 2: Instructor or students collect dog cheek cells from multiple dog donors and prepare genomic DNA (extract can be stored refrigerated for at least 1-2 days)
Students set up PCR reactions

Figure 2. Excerpt from an activity sheet from the Genomics Instructional Units Minicollection describing one of the curricular modules developed within the Bringing Big Science to Small Colleges program (for the complete activity sheet, see http://serc.carleton.edu/genomics/units/19163.html).

Table 1. List of I^{3}Us generated in the Bringing Big Science to Small Colleges collaborative project, grouped by the general level in the curriculum in which they were originally taught

ID	I^{3}U title	Conceptual content
	Introductory level	
A	Reconstructing the Evolution of Cauliflower and Broccoli	Plant development Evolution Bioinformatics
B	Human Single Nucleotide Polymorphism Determination	Genetics Human evolution Bioinformatics
C	Local Population Structure and Behavior of the Wood Frog *Rana sylvatica*	Population genetics Behavioral ecology
	Intermediate level	
D	Comparison of Protein Sequences: BLAST Searching and Phylogenetic Tree Construction	Molecular biology Molecular evolution Bioinformatics
E	Phylogenetic Analysis of Bony Fishes: Morphological and mtDNA Sequence Comparisons	Phylogenetics Vertebrate biology Bioinformatics
F	Molecular Evolution of Gene Families	Genetics Molecular evolution Bioinformatics
G	Exploring the *Chamaecrista fasciculata* Gene Space	Plant genetics Molecular evolution Bioinformatics
H	Metagenomic Analysis of Winogradsky Columns	Microbial metabolism Community ecology Ecosystems studies Bioinformatics/programming
I	Behavior, Neuroanatomy, Genomics: What Can We Learn from Mouse Mutants?	Neurobiology Behavioral genetics Bioinformatics
J	Expression of Gerontogenes in Neurons: A Comparative Genomic Approach to Studying the Role of the Nervous System in Lifespan/Aging (Raley-Susman and Gray, 2010)	Molecular evolution Behavioral neuroscience Bioinformatics
K	Comparison of a Highly Polymorphic Olfactory Receptor Gene Subfamily in Genetically Diverse Dog Breeds[a]	Molecular evolution Phylogenetics Sensory biology Bioinformatics/programming
	Advanced level	
L	Integrative Activities to Study the Evolution of Nervous System Function	Neurobiology Evolution Bioinformatics
M	Modeling Molecular Evolution[a]	Molecular evolution Bioinformatics/programming Computer science
N	Constructing and using a PAM-Style Scoring Matrix[a]	Molecular evolution Bioinformatics/programming Computer science

[a]I^{3}U was implemented in an interdisciplinary biology/computer science course.

Table 2. Pedagogical attributes (scale of biological organization, genomic level of analysis, and bioinformatic skills taught) of I³Us developed in this project and disseminated on the project's website

I³Uª	Focal taxa	Evolution	Behavior	Physiology	Morphology	Cell	Molecular	Biochemistry	Functional	Comparative	Metagenomic	Species-specific DNA database	Multi-species genome browser	Protein modeling	BLAST/alignment	Phylogenetics	Programming
					Questions asked at the level of					Analysis			Bioinformatics skills/tools				
A	*Brassica*				✓		✓			✓					✓		
B	Human	✓					✓			✓							
C	Wood frog	✓	✓				✓			✓							
D	Fish/vertebrates	✓					✓	✓	✓	✓					✓		
E	Fish/vertebrates	✓			✓		✓			✓					✓	✓	
F	*Xenopus*	✓					✓		✓				✓		✓	✓	
G	Pea/various	✓		✓			✓		✓			✓					
H	Eubacteria			✓			✓	✓			✓				✓	✓	
I	Mouse		✓		✓	✓			✓			✓	✓		✓		
J	Worm		✓							✓		✓			✓	✓	
K	Dog	✓					✓			✓				✓	✓	✓	✓
L	Various	✓	✓			✓	✓			✓					✓	✓	
M	Various	✓					✓			✓		✓			✓		✓
N	N/A						✓			✓					✓		✓

ªLetters denote I³U units as follows: A: Reconstructing the Evolution of Cauliflower and Broccoli; B: Human Single Nucleotide Polymorphism Determination; C: Local Population Structure and Behavior of the Wood Frog *Rana sylvatica*; D: Comparison of Protein Sequences: BLAST Searching and Phylogenetic Tree Construction; E: Phylogenetic Analysis of Bony Fishes: Morphological and mtDNA Sequence Comparisons; F: Molecular Evolution of Gene Families; G: Exploring the *Chamaecrista fasciculata* Gene Space; H: Metagenomic Analysis of Winogradsky Columns; I: Behavior, Neuroanatomy, Genomics: What Can We Learn from Mouse Mutants?; J: Expression of Gerontogenes in Neurons: A Comparative Genomic Approach to Studying the Role of the Nervous System in Lifespan/Aging; K: Comparison of a Highly Polymorphic Olfactory Receptor Gene Subfamily in Genetically Diverse Dog Breeds; L: Integrative Activities to Study the Evolution of Nervous System Function; M: Modeling Molecular Evolution; N: Constructing and Using a PAM-Style Scoring Matrix.

genomics instructional units, together with extensive supporting material for each module, is accessible on a dedicated website (http://serc.carleton.edu/genomics/activities.html) that also provides links to bioinformatics tools and online assessment and pedagogical resources for teaching genomics.

Rapid advances in genome sequencing and analysis offer unparalleled opportunity and challenge for biology educators. More data are being generated than can be analyzed and contextualized in traditional teaching or research models. Indeed, this explosion of data has spawned rapid growth in the discipline of bioinformatics, which is focused on development of the computational tools and approaches for extracting biologically meaningful insights from genomic data. At the same time, access to vast quantities of genomic data stored in publicly available databases can offer educators ways to engage undergraduates in authentic research and to democratize research that was previously possible only

at research-intensive universities with massive instrumentation infrastructures. The integration of genomic and bioinformatic approaches into undergraduate curricula represents one response to the national calls for biology teaching that is more quantitative and promotes deeper understanding of biological systems through interdisciplinary analyses (National Academy of Sciences, 2003; Association of American Medical Colleges and Howard Hughes Medical Institute [HHMI], 2009; NRC, 2009; American Association for the Advancement of Science, 2011). Yet relatively few faculty members who teach undergraduate biology have expertise in the fields of genomics or bioinformatics, and they may therefore feel inadequately prepared to develop their own new curricular modules capitalizing on this dispersed abundance of available resources.

Our Teagle Foundation–funded genomics education initiative, Bringing Big Science to Small Colleges: A Genomics Collaboration, was designed to address the challenges of helping

faculty members integrate genome-scale science into the undergraduate classroom. The goal of the project was to create and disseminate self-contained curricular units that stimulate students and faculty alike to think in new ways and at different scales of biological inquiry. To this end, a series of three workshops over 3 yr brought together a total of 34 faculty participants from 19 institutions and a diverse array of disciplines—including biology, computer science, and science education—to develop a set of lab modules containing a substantial genomics component. We believe that these modules are suitable for integration into existing courses in the biology curriculum and are adaptable to a variety of teaching settings.

The project website serves as a portal to activity sheets describing each I[3]U, complete with learning goals, teaching tips, and links to teaching materials, as well as downloadable resources and assessment tools (Figure 2), that can be customized by any interested educator. Each I[3]U was peer-reviewed by fellow participants, as well as by a professional project consultant who has extensive experience with Web-based description of teaching materials using this format (Manduca *et al.*, 2006). The goals of this review process were to ensure that each I[3]U met the design criteria articulated above, and to evaluate whether the activity sheet provided both an easily accessible overview of the content and enough detailed information for other instructors to adapt and implement the material and its associated assessment strategies. This peer review was complemented by each participant's own explicitly framed evaluation of his/her I[3]U through a formal reflection form (accessible at http://serc.carleton.edu/genomics/workshop09/index.html). Although these I[3]Us were designed for courses currently taught by the project participants within the specific institutions' curricula, we propose that they can be inserted into other courses encompassing similar content (Tables 1 and 2) and/or learning goals (e.g., Figure 2). We have received many communications from colleagues at other institutions who have adapted our I[3]Us for their courses.

One fundamental characteristic of each I[3]U in our collection is the focus on guided inquiry. The benefits to an undergraduate of hands-on participation in research are well documented (Nagda *et al.*, 1998; Gafney, 2001; Hunter *et al.*, 2007; Kardash *et al.*, 2008; Lopatto, 2009). Integrating authentic research experiences into the undergraduate curriculum allows this powerful learning model to be scaled from use with only a few students to use with entire laboratory sections (Lopatto 2009; Lopatto *et al.* 2008). Like other national participatory genomic teaching initiatives (Campbell *et al.*, 2006, 2007; Ditty *et al.*, 2010; Shaffer *et al.*, 2010; HHMI, 2011), our model for curriculum development in genomics emphasizes synergies between student-centered research and education. However, in contrast with some of these other projects, our grassroots approach leveraged a wealth of existing expertise by providing opportunities for individual faculty members to develop, implement, modify, evaluate, and share undergraduate teaching modules that stem from their own research and/or teaching interests. In this regard, our project most closely resembles the Genome Consortium for Active Teaching, which provides faculty members and their undergraduates access to microarrays from a variety of organisms, allowing participants to define their own research questions in a model system with which they are already familiar (Campbell *et al.*, 2006, 2007).

Our collaborative effort among biologists, computer scientists, and science educators has yielded a collection of pedagogical resources that can be adapted for use in a wide variety of educational settings. We invite other biologists to begin building on our work by using and providing feedback on our I[3]Us. Faculty who have tested materials that exemplify our design principles are encouraged to add them to our collection. For further information, please contact the corresponding author.

ACKNOWLEDGMENTS

REFERENCES

American Association for the Advancement of Science (AAAS) (2011). Vision and Change in Undergraduate Biology Education, Washington, DC.

Association of American Medical Colleges (AAMC) and Howard Hughes Medical Institute (HHMI) (2009). Scientific Foundations for Future Physicians, Washington, DC.

Campbell AM, Eckdahl TT, Fowlks E, Heyer LJ, Mays Hoopes LL, Ledbetter ML, Rosenwald AG (2006). Genome Consortium for Active Teaching (GCAT). Science *311*, 1103–1104.

Campbell AM, Ledbetter MLS, Hoopes LLM, Eckdahl TT, Heyer LJ, Rosenwald AG, Fowlks E, Tonidandel S, Bucholtz B, Gottfried G (2007). Genome Consortium for Active Teaching: Meeting the goals of BIO2010. CBE Life Sci Educ *6*, 109–118.

Ditty JL *et al.* (2010). Incorporating genomics and bioinformatics across the life sciences curriculum. PLoS Biol *8*, e1000448.

Gafney L (2001). The impact of research on undergraduates' understanding of science. Counc Undergrad Res Q *21*, 172–176.

HHMI (2011). Science Education Alliance. www.hhmi.org/grants/sea (accessed 3 May 2012).

Hunter A-B, Laursen SL, Seymour E (2007). Becoming a scientist: the role of undergraduate research in students' cognitive, personal, and professional development. Sci Educ *91*, 36–74.

Kardash C-A, Wallace M, Blockus L (2008). Science undergraduates' perceptions of learning from undergraduate research experiences. In: Developing, Promoting, and Sustaining the Undergraduate Research Experience in Psychology, ed. RL Miller, RF Rycek, E Balcetis, ST Barney, BC Beins, SR Burns, R Smith, ME Ware, Washington, DC: Society for the Teaching of Psychology. http://teachpsych.org/resources/e-books/ur2008/ur2008.php (accessed 3 May 2012).

Lopatto D (2009). Science in Solution: The Impact of Undergraduate Research on Student Learning, Tuscon, AZ: Research Corporation for Science Advancement.

Lopatto D *et al.* (2008). Genomics Education Partnership. Science *322*, 684–685.

Manduca CA, Fox S, Iverson ER (2006). Digital library as network and community center. D-Lib Magazine, *12*. doi:10.1045/december2006-manduca.

Nagda BA, Gregerman SR, Jonides J, von Hippel W, Lerner JS (1998). Undergraduate student-faculty research partnerships affect student retention. Rev Higher Educ *22*, 55–72.

National Academy of Sciences (2003). BIO 2010: Transforming Undergraduate Education for Future Research Biologists, Washington, DC: National Academies Press.

National Research Council (NRC) (2005). America's Lab Report, Washington, DC: National Academies Press.

NRC (2009). A New Biology for the 21st Century: Ensuring the United States Leads the Coming Biology Revolution, Washington, DC: National Academies Press.

Raley-Susman KM, Gray JM (2010). Exploration of gerontogenes in the nervous system: a multi-level neurogenomics laboratory module for an intermediate neuroscience and behavior course. J Undergrad Neurosci Educ *8*, A108–A115.

Shaffer CD *et al.* (2010). The Genomics Education Partnership: successful integration of research into laboratory classes at a diverse group of undergraduate institutions. CBE Life Sci Educ *9*, 55–69.