

Transposing from the Laboratory to the Classroom to Generate Authentic Research Experiences for Undergraduates

James M. Burnette, III, and Susan R. Wessler¹

Department of Botany and Plant Sciences, University of California, Riverside, California 92521

ABSTRACT Large lecture classes and standardized laboratory exercises are characteristic of introductory biology courses. Previous research has found that these courses do not adequately convey the process of scientific research and the excitement of discovery. Here we propose a model that provides beginning biology students with an inquiry-based, active learning laboratory experience. The Dynamic Genome course replicates a modern research laboratory focused on eukaryotic transposable elements where beginning undergraduates learn key genetics concepts, experimental design, and molecular biological skills. Here we report on two key features of the course, a didactic module and the capstone original research project. The module is a modified version of a published experiment where students experience how virtual transposable elements from rice (*Oryza sativa*) are assayed for function in transgenic *Arabidopsis thaliana*. As part of the module, students analyze the phenotypes and genotypes of transgenic plants to determine the requirements for transposition. After mastering the skills and concepts, students participate in an authentic research project where they use computational analysis and PCR to detect transposable element insertion site polymorphism in a panel of diverse maize strains. As a consequence of their engagement in this course, students report large gains in their ability to understand the nature of research and demonstrate that they can apply that knowledge to independent research projects.

BIOLGY instruction in introductory courses for life science majors has traditionally used a passive lecture format that is reinforced through standardized laboratory exercises. A consequence of this mode of instruction is that students often memorize a series of facts and vocabulary unrelated to any particular living system. Further, because the scientific method is not emphasized, students emerge poorly equipped to design properly controlled experiments or to interpret data. Students who are inadequately engaged by these teaching approaches may fail to achieve the best possible learning outcomes. Consequently, there is a high attrition rate from life science degree programs, and students

who do graduate are often underprepared for their chosen careers (Harrison *et al.* 2011). A number of recent studies have indicated that this conventional mode of instruction is ineffective and that meaningful learning instead requires an intense inquiry-based experience that uses scientific teaching methodologies and research-focused instruction (American Association for the Advancement of Science 2011).

A problem-solving approach to learn science is best implemented when instruction mirrors the research process and students are engaged in addressing biological questions. This scientific approach to teaching applies active learning, immediate feedback, and diversity of instruction techniques to foster critical thinking skills and a richer understanding of the content (Handelsman *et al.* 2004). While effective at all levels, inquiry-based learning is especially beneficial to undergraduate freshmen and sophomores (Seymour *et al.* 2004; Derting and Ebert-May 2010). Scientific teaching strategies provide lower division students with a firm foundation for advanced course work (Derting and Ebert-May 2010) and informs these students about career choices in science, technology, engineering, and math (STEM) (Harrison *et al.* 2011). Several inquiry-based courses have been reported previously where students participate in semester-long

Copyright © 2013 by the Genetics Society of America

doi: 10.1534/genetics.112.147355

Manuscript received June 25, 2012; accepted for publication November 13, 2012

Available freely online through the author-supported open access option.

Supporting information is available online at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.112.147355/-/DC1>.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. JQ808121–JQ808128.

Arabidopsis Biological Resource Center accession nos. CS67774–CS67779, Education Kit CS19995.

IRB Project Number 2009-10835-0

¹Corresponding author: Dept. of Botany and Plant Sciences, University of California Riverside, 4107A Genomics, Riverside, CA 92521. E-mail: susan.wessler@ucr.edu

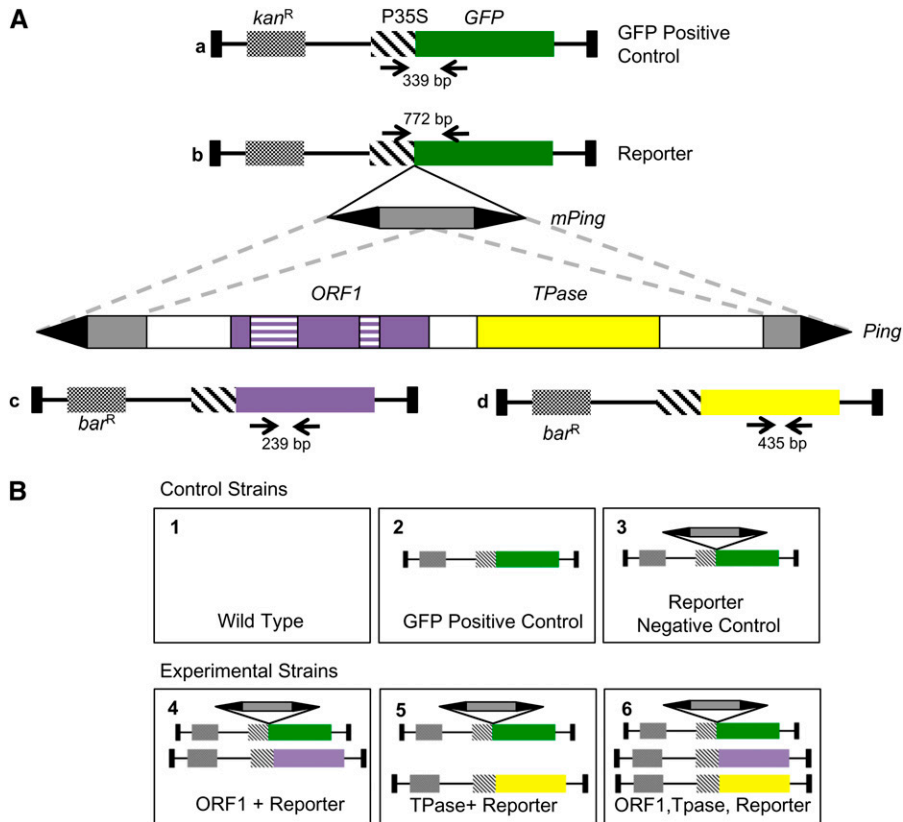


Figure 1 (A) Structure and origin of T-DNA constructs. The two genes of *Ping* are: *Tpase* (yellow) and *ORF1* (purple, introns hashed). DNA shared between *Ping* and *mPing* is in gray and both are flanked by the terminal inverted repeat (black arrowheads). T-DNA constructs contain either a kanamycin resistance gene (*kan^R*) or the glufosinate resistance gene (*bar^R*). *ORF1* T-DNA (c) contains the cDNA for *ORF1* and both *ORF1* (c) and *TPase* (d) are under control of the 35S promoter (hashed boxes). PCR primers used in the experiment described in Figure 2 are shown as arrows below the T-DNA constructs. (B) *Arabidopsis thaliana* strains used in the module. Strain 1 is untransformed wild-type *Arabidopsis*. Strains 2–6 are *Arabidopsis* transformed with the constructs shown.

guided research projects that produce learning gains (Hatfull *et al.* 2006; Call *et al.* 2007; Lopatto *et al.* 2008). Building on these successes we saw the need for a discovery-based plant biology course that was modular in format with short, easily adopted, and inexpensive projects.

We developed the Dynamic Genome (DG) course to test whether freshmen could learn genetic principles by performing experiments derived from a research laboratory and focused on a single biological system. Using this approach, we propose that students will learn a specific suite of basic genetic concepts and laboratory skills that form an enduring tool set to carry into and practice in future courses and research experiences. To this end, the DG course replicates a modern research laboratory as an undergraduate classroom in which students engage in inquiry-based experiences focused on transposable element (TE) biology.

Found in all eukaryotes characterized to date, TEs are mobile genetic elements whose DNA sequences usually comprise the largest component of the data generated by genome sequencing projects. Because TEs are often masked and ignored by most researchers, they represent a tremendous source of untapped raw material for undergraduates to analyze. The structural features of TEs are relatively simple, which makes them a good subject on which to focus a course for beginning students. TEs are divided into two classes based on the mechanism of transposition. Class 1 elements transpose by an RNA intermediate, whereas class 2 elements excise from one chromosomal locus and insert elsewhere in

the genome. Class 2 DNA elements, which are the focus of the DG course, have the characteristic structural feature of terminal inverted repeats (TIRs) that may flank genes necessary for transposition (Figure 1A). Elements that contain a gene that encodes transposase function are capable of moving themselves (autonomous elements), while nonautonomous elements lack functional transposase.

The combination of the abundance and structural simplicity of TEs provides an opportunity for students to address several genetic principles, develop new laboratory skills, and experience the excitement of scientific discovery. The first module of the DG course introduces students to the basic concepts of molecular genetics, experimental design, molecular biological tools, and TE structure and function by reproducing a published experiment on the *PIF/Harbinger* superfamily of TEs (Yang *et al.* 2007). This experiment was chosen specifically because it follows a classic design with both positive and negative controls and clearly illustrates the relationship between genotype and phenotype. The molecular tools of DNA extraction, PCR, agarose gel electrophoresis, and data analysis are also introduced. Following successful completion of this module, students participate in an authentic research project that exploits the extraordinary level of TE insertion site polymorphism among maize strains and the skills acquired in the first module to explore the dynamic nature of genomes.

In this article we describe the module and the project in the context of the concepts and laboratory skills addressed

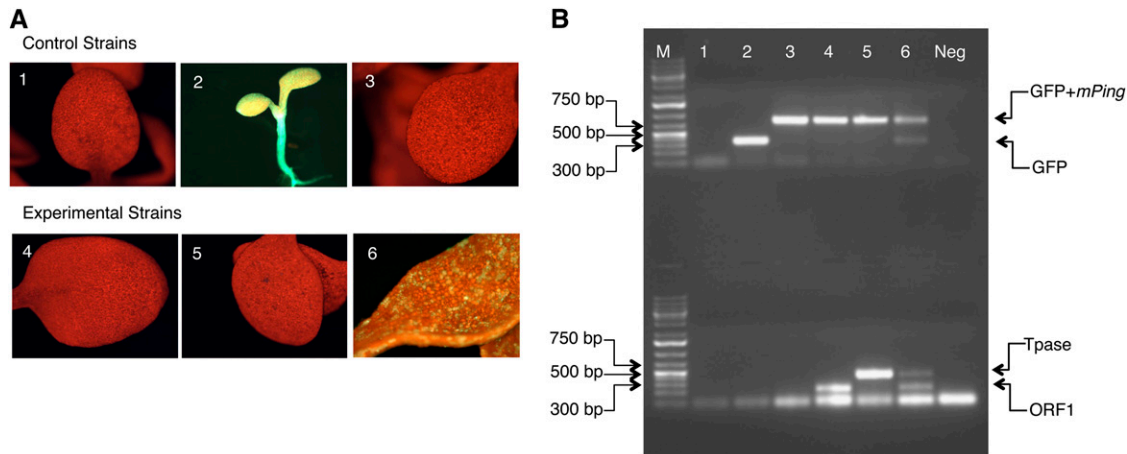


Figure 2 Visualizing the genotype and phenotype of *Arabidopsis* seedlings. (A) Images of *A. thaliana* leaves under blue light. Strain numbers are the same as in Figure 1B. (B) Agarose gel electrophoresis of PCR amplicons. The top half shows products of GFP amplification [772 bp with *mPing* present and 339 bp after *mPing* excision (lane 6 only)] and the bottom half shows the PCR amplicons for *ORF1* (239 bp) and *Tpsase* (435 bp). M: 1-kb Plus O'Gene Ruler (Fermentas). Low molecular weight bands (<75 bp) are primer dimers. Neg, water control.

by each. By their participation in the module, we assessed whether students were able to reproduce a sophisticated experiment. To participate in the research project, students need to apply the skills and knowledge gained in the first module in a completely different experimental setting. Finally we show that the module and project effectively prepare students for future course work and independent research.

Materials and Methods

Plant material and DNA extraction

Seeds of *Arabidopsis thaliana* containing the T-DNA constructs shown in Figure 1 (Yang *et al.* 2007) were sterilized in 20% (v:v) commercial bleach with 0.1% Tween 20 (Fisher), washed three times in sterile water, and incubated at 4° for 2 days. Seeds were germinated on standard MS media plates or in soil and grown for 2 weeks before use in class. Cotyledons were observed using a stereo dissecting microscope under blue light with filters optimized for GFP fluorescence. If a fluorescent stereoscope is unavailable, images are available at http://dynamicgenome.ucr.edu/course_materials.html. DNA was isolated from seedlings as described (Li *et al.* 2010) with modifications (see Supporting Information, File S1 for a detailed protocol).

Thirty-five maize lines [B73, CML247, Harinoso de Ocho (HdO), HP301, Maiz Chapalote, Mo17, Benz, Palomero, Til1, Til14, Til15, Til16, Ki3, CML322, CML333, IL14H, NalTel, Oh43, Ki11, Mag469, M37W, Venezuela 703, Mo37w, Ky21, Mo18W, B37, CML5, Tzi8, C103, IL14H, Oh7B, Tzi9, CML227, B96, and CML248] were selected from the maize diversity panel (Liu *et al.* 2003). Seeds obtained from the National Plant Germplasm System (NPGS) (United States Department of Agriculture, Agricultural Research Service, USDA-ARS) were sterilized in 20% commercial bleach with 0.1% Tween 20 (Sigma), germinated on wet filter paper, and grown in Petri dishes or in soil for 2–3 weeks. Leaf and root tissues of ~5 cm

were ground in a mortar chilled with liquid nitrogen and processed as described in File S1.

Primer design

DNA sequence of the intron and exons flanking the TE was extracted from the Maize Genome Browser (www.maizesequence.org). Students designed PCR primers from exon sequences using the online program Primer3Plus (www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). Primers were synthesized by Integrated DNA Technologies (Coralville, IA).

PCR amplification

GFP, *ORF1*, and *transposase* transgenes were detected using the primers and cycling parameters described in Yang *et al.* (2007). Maize gene 169020 was amplified using primers 169020_ex3_F1 (5'-GGAAGGAGACCAAAGAATCC-3') and 169020_ex5_R1 (5'-AGCGTCATCTGACCTTTTCG-3'). Cycling conditions consisted of initial denaturation at 94° for 5 min followed by 35 cycles of 94° for 30 sec, annealing at 48° for 30 sec, and extension at 72° for 1 min, with a final extension at 72° for 5 min. Maize *Actin 1* gene was amplified using actin1_ex2_F (5'-AAGTACCCGATTGAGCATGG-3') and actin1_ex3_R (5'-CTCAGCGCTCGTAGTGAGG-3'). Cycling conditions consisted of initial denaturation at 94° for 5 min followed by 35 cycles of 94° for 30 sec, annealing at 55° for 30 sec, and extension at 72° for 30 sec, with a final extension at 72° for 5 min. Students used 2X Quick-Load Taq Master Mix (Taq polymerase, dNTPs, gel loading dye; New England Biolabs, Ipswich, MA), to simplify PCR setup to which they added primers and DNA. PCR results were visualized on 1.5% agarose gels (Agarose HS, Denville, South Plainfield, NJ).

DNA sequencing

PCR amplicons were sequenced directly (by Genewiz, www.genewiz.com) after purification by treatment with ExoSAP-IT

exonuclease and shrimp alkaline phosphatase (Affymetrix/USB Molecular, Santa Clara, CA).

CURE evaluation survey

The Classroom Undergraduate Research Experience (CURE) survey was administered after the final exam (Lopatto 2007). L. Jaworski (Grinnell College, Grinnell, IA) provided anonymous, aggregated data as described in Lopatto (2007) and Lopatto *et al.* (2008).

Reproducibility of experiments

Experiments for the module and project were specifically designed for use in any laboratory equipped with micropipettes, a thermocycler, and agarose gel electrophoresis equipment. Seeds for the six *A. thaliana* strains are available at no cost from the *Arabidopsis* Biological Resources Center (<https://abrc.osu.edu/>). Request education kit CS19995. Maize seeds are also available at no cost from the National Plant Germplasm System of the USDA-ARS (<http://www.ars-grin.gov/npgs/orders.html>).

Implementation

The DG course meets twice a week for 3 hr each for laboratory and bioinformatics work. The *Arabidopsis* excision assay requires three laboratory periods. On the first day, students observe seedling phenotypes and begin DNA extraction. On the second day, students finish DNA extraction, assay DNA concentration, and set up PCR. On the final day, they load, run, and analyze agarose gels.

The maize polymorphism research project requires six laboratory sessions, or 3 weeks, to complete. During the first week, students design primers and extract DNA from seedlings. In the second week, primers are tested using DNA from reference strain B73 and then with the other strains. PCR products are sequenced and analyzed in the final week.

Cost

The combined per-student cost of these two modules is approximately \$40 (\$7 for the module, \$35 for the project). However, reducing the number of maize strains to 16 (B73, CML247, Mo17, Ki3, CML322, CML333, NalTel, Oh43, Ki11, M37W, Mo37w, Ky21, Mo18W, CML5, Tzi8, and Oh7B) and having students work in pairs reduces the cost dramatically from \$35 per student to \$8 per pair (see also [File S2](#)).

Results

Assaying virtual sequences

Rationale: In the first module, students replicate a published experiment to master techniques, follow the basic elements of good experimental design, understand the difference between genotype and phenotype, and learn about TE structure. After completing the first module, students will be able to: (1) discuss the relationship between genotype and phenotype, (2) draw conclusions about the requirements for TE excision based on their own data, (3) understand the role

of positive and negative controls, and (4) use a subset of molecular biology techniques in an experimental context.

Experimental design: Previous studies identified the 5-kb autonomous element *Ping* and the 430-bp nonautonomous element *mPing* in the rice genome sequence (Jiang *et al.* 2003; Kikuchi *et al.* 2003). *Ping* encodes two predicted genes: a transposase (*TPase*) and *ORF1*, which has no obvious function (Figure 1A). The purpose of the module was to (1) demonstrate how virtual sequences can be assayed for biological function and (2) determine whether the *TPase* encoded by *Ping* is sufficient for excision or whether *ORF1* is also required.

The module utilizes transgenic *A. thaliana* strains that were previously transformed with rice TE sequences (Yang *et al.* 2007). Three of the strains serve as controls (Figure 1B, strains 1–3), whereas three other strains contain different *Ping* genes (Figure 1B, strains 4–6). In addition, strains 4–6 contain the reporter gene that is the only construct in strain 3. Strain 6 was generated by crossing strains 4 and 5.

Transposase activity is detected as the excision of *mPing* from the *gfp* reporter. Without transposase activity the presence of *mPing* in the *gfp* gene prevents the synthesis of functional GFP protein. However if the necessary proteins are present to excise *mPing*, the *gfp* gene is restored, functional GFP is expressed, and green fluorescent sectors are observed.

Student results: Groups of two students each were given unlabeled 2-week-old seedlings from each of the six strains and they determined seedling phenotypes using fluorescence microscopy. The students recorded the phenotypes for future analysis while the instructor assisted in the use of the microscope. Three distinct phenotypes were possible: uniform red (chlorophyll) fluorescence, green (GFP) fluorescence, or green GFP spots on a red background (Figure 2A where the strains' designations are the same as in Figure 1B). The third phenotype is the result of *mPing* excision from the *gfp* gene and the subsequent expression of functional GFP in those cells. Students recorded the phenotypes of each strain and then isolated DNA for PCR using the primers shown in Figure 1A. An example of student data is shown in Figure 2. Students were able to generate quality data in two to three laboratory periods of 3 hr each.

Phenotypic analysis: Students found that the negative controls (strains 1 and 3) fluoresced only red, while the positive control (strain 2) fluoresced only green (Figure 2A). The positive control validated expression of the *gfp* reporter and that the insertion of *mPing* is sufficient to block *gfp* expression in strain 3. Strains 4 and 5 were phenotypically identical to 3 with uniform red fluorescence, while spots of green fluorescence on a red background were detected in strain 6.

Genotypic analysis: Working in pairs, students extracted DNA from each strain, performed PCR, and resolved the products by gel electrophoresis (see *Materials and Methods*) (Figure 2B). Primers for constructs a and b flank the *mPing*

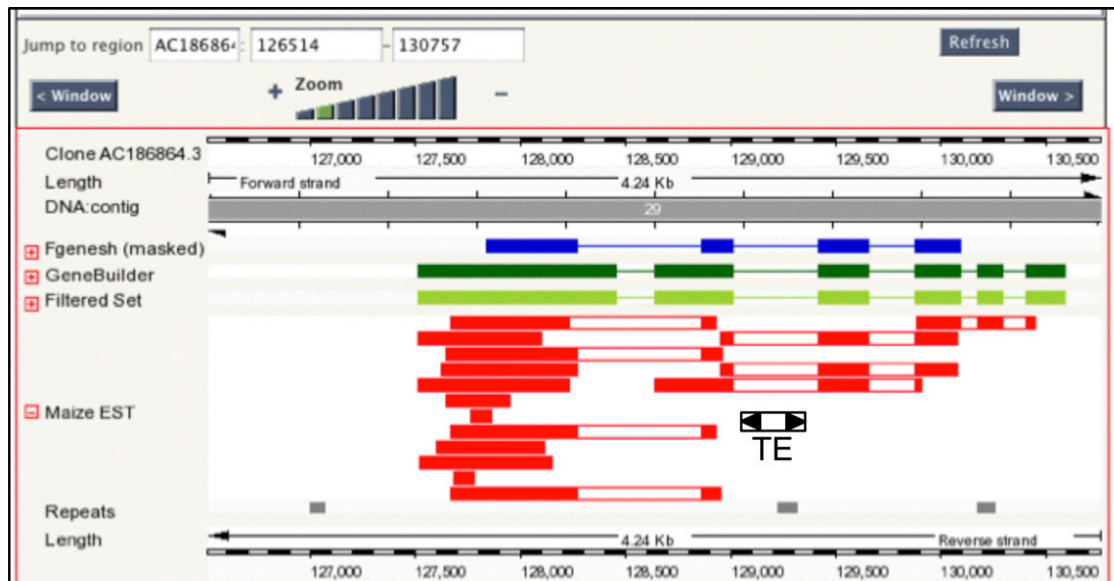


Figure 3 Screen capture of the Maize Genome Browser (maizesequence.org) for predicted gene 169020. The TE insertion fulfills the criteria for use: (1) located in an intron (gene models in blue and green), (2) flanking exons are supported by EST evidence (red), and (3) there are no other repeats in the intron (gray boxes).

insertion site (Figure 1A, arrows), allowing detection of the *gfp* gene and *mPing* excision products in all strains except wild-type (untransformed) *A. thaliana* (Figure 2B, top half of gel). Strains were assayed for the presence of constructs c and d using multiplex PCR (Figure 2B, bottom). One student prepared the samples for the top half of the gel while the other set up the multiplex reactions in the bottom half. We found that this gel format simplified the interpretation of the PCR products, as they were able to determine the genotype of individual strains by scanning down a single lane. Although strains 3 through 6 each contain the reporter (Figure 2A, top, larger band), the excision product was only detected in strain 6 (smaller PCR amplicon, top half of gel), which contained both *Ping* genes (Figure 2B, bottom). No excision products were detected in either strain that contained only one of the two *Ping* genes (construct c or d).

Data analysis: Students combined the PCR results and phenotypic data to answer the original question about the requirements for *mPing* excision. To this end, strategies for analyzing the phenotypic and genotypic data were first discussed in class. Students were then asked to write a short report that included a figure describing the experiment and their conclusions about the requirement for *Ping* genes in *mPing* excision. Students were very successful at generating useful data. Of the 30 groups that completed this module, 19 (63%) obtained interpretable gels. The remaining 11 (37%) borrowed data from other students or from the instructor to participate in data analysis.

The most common sources of experimental error were in the preparation of the PCR master mix and the individual reactions. DNA extractions were analyzed by spectroscopy for concentration and purity allowing students to use their best DNA samples.

The class project: identification of TE insertion polymorphism in maize

The class project served two purposes. First, it gave students the opportunity to generate novel data as they contributed to an ongoing research project. Second, through their participation in the project, students discovered how the skills and genetic concepts mastered in the first module could be applied in a completely different experimental context.

Rationale: In this class project, students discover that individuals within a species harbor variation in TE insertion sites and other sequence polymorphisms. It is specifically designed to challenge the common misconception that genomes mutate in response to changes in the environment (Nehm and Reilly 2007) and that different portions of the genome are under different selection pressures (*i.e.*, exons vs. introns vs. TEs). In addition, because this project utilizes bioinformatics analyses to generate testable hypotheses, students learn how computer databases are central to contemporary genetic research.

Experimental design: The project originated with maize TE sequences that were previously identified by a researcher participating in the National Science Foundation-funded Maize TE Annotation Project as part of the larger maize genome sequencing effort (Baucom *et al.* 2009; Schnable *et al.* 2009). The TE group characterized virtually all of the TEs in the B73 reference genome, and the insertion sites were then cataloged in the Maize Genome Browser (<http://maizesequence.org/>).

In the first offering of the project, seven students were each given a TE sequence that exists as several copies in the maize genome. However, precise insertion sites were not

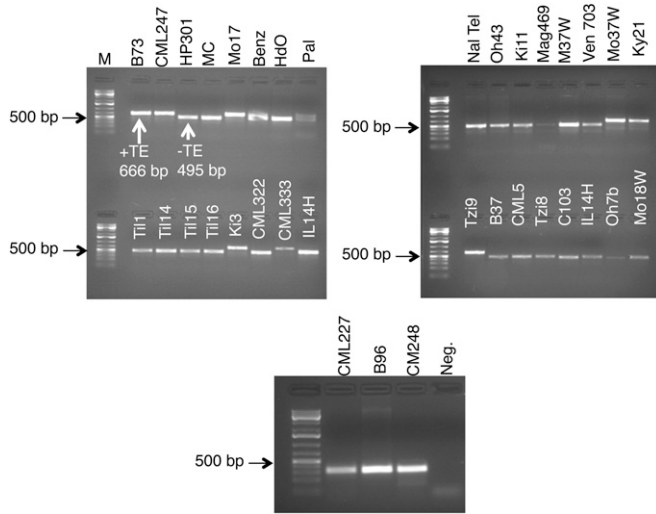


Figure 4 TE polymorphism in two maize introns. Agarose gels showing PCR amplicons from 36 maize strains for locus 169020. The strain names are shown above the lanes with the 36 reactions divided across three gels. The negative control for the group of reactions is shown in the last lane. The amplicon containing the TE is 666 bp while the band lacking it is 495 bp.

known by either students or instructors. The objective was to identify an insertion of a related element (a family member) in an intron of the reference B73 gene and then perform PCR to determine whether this insertion site was polymorphic in a panel of 34 diverse maize inbred strains and landraces.

With their TE sequence in hand, students first used the BLAST feature of the Maize Genome Browser to query the genome. TE insertions meeting the following criteria were chosen for further study: (a) they were located in an intron, (b) the intron was <1 kb, (c) flanking exons had ample EST evidence (Figure 3, in red), and (d) the intron did not contain other TE insertions (as determined by inspecting the repeat track of the browser). Seven students were enrolled in the course that completed the project and each student used a different TE sequence and found one insertion of that TE that fit the above criteria. Students then designed PCR primers from the flanking exon sequence using online software (see *Materials and Methods*).

Upon receipt of their primers, each student performed PCR on B73 DNA, provided by the instructors, to determine whether the primers worked. As a group, the students extracted DNA from the 35 maize inbred strains (B73 included) with each student performing five extractions. Concentrations and purity of DNA samples were determined using UV spectrophotometry and by PCR amplification of the *Actin 1* gene. If a DNA sample did not check out, a replacement was provided by the instructor. Finally, students performed the experiment by using their primers with the 35 DNA samples including B73 to detect polymorphisms.

An example of the PCR results obtained by one student is shown in Figure 4 and the results for the class are tabulated in Table 1. Each student selected four amplicons for DNA

Table 1 Results of polymorphism analysis

Strain	169020	4996	38598	72789	58227	33894	22861
B73 ^a	+	+	+	+	+	+	+
CML247 ^a	+	-	-	-	-	+	-
HP301	-	ND	-	-	-	-	ND
Maiz Chapalote	-	-	-	-	-	+	-
Mo17 ^a	+	-	-	-	+	-	+
Benz	-	+	-	-	-	ND	+
Harisono de Ocho	-	-	-	ND	-	-	ND
Palomero	ND	ND	ND	ND	-	-	+
Til1	-	ND	ND	ND	ND	+	ND
Til14	-	-	-	-	ND	+	-
Til15	-	ND	-	ND	ND	-	+
Til16	-	ND	ND	ND	+	ND	-
Ki3 ^a	+	-	-	-	-	+	-
CML322 ^a	-	-	-	-	+	+	-
CML333 ^a	+	ND	-	ND	-	+	-
IL14H	-	-	-	-	-	-	-
NalTel ^a	-	ND	ND	-	-	-	-
Oh43 ^a	-	-	+	-	+	-	-
Ki11 ^a	-	-	-	-	-	+	ND
Mag469	-	-	+	-	-	-	+
M37W ^a	-	ND	ND	ND	-	ND	ND
Venezuela 703	-	-	-	-	-	ND	-
Mo37W ^a	+	-	+	+	+	+	+
Ky21 ^a	+	-	-	-	-	+	+
Mo18W ^a	-	-	-	-	-	+	-
B37	-	-	-	-	-	-	ND
CML5 ^a	-	-	-	+	-	+	+
Tzi8 ^a	-	-	-	-	-	+	+
C103	-	-	-	-	+	ND	+
IL14H	-	-	-	-	-	ND	+
Oh7B ^a	-	-	-	-	-	+	+
Tzi9	+	-	-	-	-	+	+
CML227	-	-	ND	-	-	-	+
B96	-	-	-	-	-	-	+
CML248	-	-	-	-	+	ND	-

Gene identifiers: GRMZM2G169020, GRMZM2G004996, GRMZM2G038598, GRMZM2G072789, GRMZM2G058227, GRMZM2G033894, and GRMZM2G022861. +, insertion present; -, insertion absent, ND, no PCR amplicon or results ambiguous.

^aThe core 16 strains with the greatest variability.

sequencing: two with and two without the TE insertion (from four strains). The PCR products were purified as described in *Materials and Methods* and sequenced (GeneWiz).

Data analysis: Students downloaded ABI trace files from a file transfer protocol (FTP) site and used a trace-viewing program (4Peaks, <http://www.mekentosj.com/science/4peaks>) to inspect the DNA sequencing results and determine sequence quality. Sequences were extracted from the files using 4Peaks and students used a multiple sequence alignment program (Edgar 2004) to compare to the B73 reference obtained from the Maize Genome Browser. Students were instructed to analyze the sequences and describe how they differ. Each student generated a figure similar to Figure 5 in class and the group discussed the results. Four DNA sequences from PCR bands for the insertion 169020 (Figure 4) are shown in Figure 5. As seen in this alignment, students could identify the TE insertion site and the TIR (Figure 5, blue and yellow, respectively). Furthermore, analysis of each of the multiple sequence alignments revealed additional polymorphisms.

```

B73      ATATTTTTTCGATTCTTCTCTTA-TTGTTTCTCGAGCGCCTAGAGAAGCTCGTTCTTGT
TZI9     ATATTTTTTCGATTCTTCTCTTA-TTGTTTCTCGAGCGCCTAGAGAAGCTCGTTCTTGT
CML247   ATATTTTTTCGATTCTTCTCTTA-TTGTTTCTCGAGCGCCTAGAGAAGCTCGTTCTTGT
IL14H    ATA-TTTTTTCGATTCTTCTCTTA-TTTTTTCTCGAGTGCCTAGAGAAGCTCGTTCTTGT
HdO      ATA-TTTTTTCGATTCTTCTCTTA-TTTTTTCTCGAGTGCCTAGAGAAGCTCGTTCTTGT
***      *****

B73      ACCATCCTTTTCATTTTAGTTGGATGATACTCCCTCCGTCFCATGATATAAGGCGTAACC
TZI9     ACCATCCTTTTCATTTTAGTTGGATGATACTCCCTCCGTCFCATGATATAAGGCGTAACC
CML247   ACCATCCTTTTCATTTTAGTTGGATGATACTCCCTCCGTCFCATGATATAAGGCGTAACC
IL14H    ACCATCCTTTTCATTTTAGTTGGATGA-----
HdO      ACCATCCTTTTCATTTTAGTTGGATGA-----
*****

B73      ACTTTTTATTCCTTGTCGCCGAATATAAGGCGTGCCTCTCTATGCATACGTATATCGATG
TZI9     ACTTTTTATTCCTTGTCGCCGAATATAAGGCGTGCCTCTCTATGCATACGTATATCGATG
CML247   ACTTTTTATTCCTTGTCGCCGAATATAAGGCGTGCCTCTCTATGCATACGTATATCGATG
IL14H    -----
HdO      -----

B73      CAGTGGTATAGAGACAATTAATGTATTTCTTGGTCTTTGAACCAGAGGTGGTTACGCCT
TZI9     CAGTGGTATAGAGACAATTAATGTATTTCTTGGTCTTTGAACCAGAGGTGGTTACGCCT
CML247   CAGTGGTATAGAGACAATTAATGTATTTCTTGGTCTTTGAACCAGAGGTGGTTACGCCT
IL14H    -----
HdO      -----

B73      TATAACTGGGACGGAGGGAGTACTTGTTAAATTTAAGAAATGTCCAGAGAGTTCAGCA
TZI9     TATAACTGGGACGGAGGGAGTACTTGTTAAATTTAAGAAATGTCCAGAGAGTTCAGCA
CML247   TATAACTGGGACGGAGGGAGTACTTGTTAAATTTAAGAAATGTCCAGAGAGTTCAGCA
IL14H    -----TACTTGTTAAA-TTAAGAAATGTCCAGAGAGTTCAGCA
HdO      -----TACTTGTTAAATTTAAGAAATGTCCAGAGAGTTCAGCA
*****

B73      AAGCAATAACATTTGTACTTACCTGTATATGATATATACATGTGGCCTATGAATACTTTT
TZI9     AAGCAATAACATTTGTACTTACCTGTATATGATATATACATGTGGCCTATGAATACTTTT
CML247   AAGCAATAACATTTGTACTTACCTGTATATGATATATACATGTGGCCTATGAATACTTTT
IL14H    AAGCAATAACATCTGTACTTACCTGTATATGATATATACATGTGGCCTATGAATACTTTT
HdO      AAGCAATGACATTTGTACTTACCTGTATATGATATATACATGTGGCCTATGAATACTTTT
*****

```

Figure 5 Multiple sequence alignment showing insertion polymorphisms at locus 169020. Amplicons for strains TZI9, CML247, IL14H, and HdO from the gels in Figure 4 were sequenced and aligned to the B73 sequence. The 172-bp insertion corresponds to the *Stowaway* TE (in blue) and accounts for the difference in amplicon size. The TIR of the TE is highlighted in yellow. Two other types of polymorphism indels and SNPs are highlighted in purple and green, respectively. An asterisk below a column of nucleotides indicates 100% similarity.

For the alignment shown in Figure 5, there were three indels (purple) and five single nucleotide polymorphisms (green). The four novel DNA sequences in Figure 4 (for insertion site 169020) were submitted to GenBank along with four sequences for insertion site 4996 (data not shown).

Course evaluation

A predicted major outcome of the DG course is that students who complete the course are prepared for engagement in scientific research. To test whether this outcome was achieved, we used the CURE survey (Lopatto 2007) and tracked the number of students who engaged in research experiences subsequent to the course. In the CURE survey, which is administered at the beginning and end of each course in the study, students nationwide report their gain in knowledge and abilities on a scale of 1 (no gain) to 5 (very large gain). Twenty DG students from three course sections (including the 7 that participated in the research project) completed the survey and their reported gains were compared with all respondents nationwide in 2010 ($n = 950$). DG students

reported large to very large gains compared to the moderate to large gains reported by all students nationwide in all research skills categories (Figure 6A). Students in the DG courses reported equally large gains in understanding research concepts (Figure 6A).

DG students were surveyed in fall 2010 concerning the amount of time they spent participating in research upon completion of the course. Of the 45 respondents (including students who chose not to complete the CURE survey), 53% participated in a subsequent research project for at least one semester (Figure 6B). The majority of those students, 58%, remained in research positions for at least 1 yr. The comments from students on the survey were overwhelmingly positive. Their response is exemplified by this statement from one student, “The Dynamic Genome class gave me confidence to pursue other research opportunities, because I know (from my experience with our Dynamic Genome final project) that I can develop a question, design an experiment, perform the necessary techniques, and then create an informative poster about my results.”

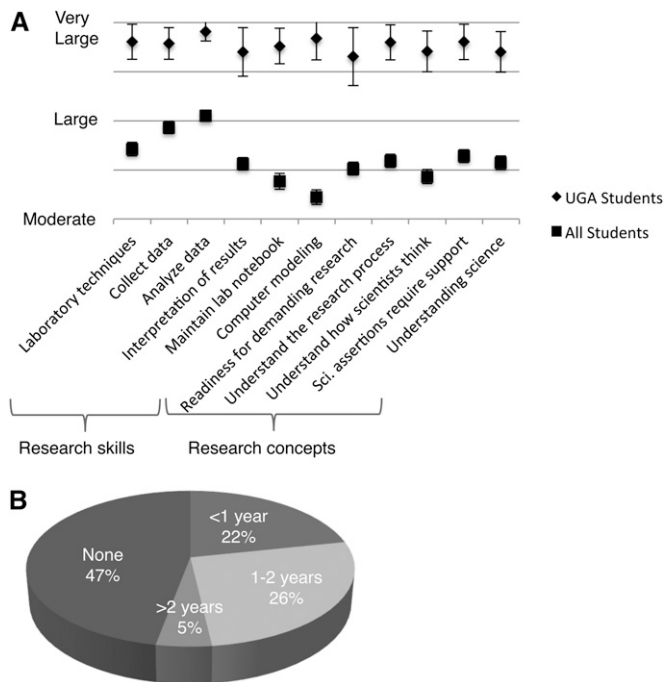


Figure 6 Survey results of students who completed the DG course vs. students nationwide. (A) Student-reported gains on the CURE survey in course elements related to research preparedness (moderate = 3, large = 4, and very large = 5). (B) Participation in undergraduate research after completing a University of Georgia Dynamic Genome course.

Discussion

Many traditional introductory laboratory courses move students through a series of unrelated experiments and as a consequence they learn new skills without the opportunity to reuse prior skills and knowledge. In contrast, the research project of the DG course utilizes the same suite of skills as used in the module and, in this way, builds on previously learned concepts. This approach allows students to move along the levels of Bloom's Taxonomy from rote knowledge to higher levels of cognition (Crowe *et al.* 2008). For example, in terms of developing conceptual understanding, in the module students experience the movement of TEs as they excise from the GFP reporter in transgenic *A. thaliana*. The project then broadens their conceptual understanding of TEs by showing that their movement generates genomic diversity. With regard to mastering methodology, in the module, PCR is used to genotype the transgenic *A. thaliana* plants while in the project it is used to detect TE insertion-site polymorphism. Perhaps most important, in the module students first learn about experimental design, including the significance of positive and negative controls (Figure 1B, strains 2 and 3, respectively). Having gained this knowledge, they are then challenged in the project to design experiments—including all controls—to detect TE insertion-site polymorphism in maize strains. Finally, the experimental design of the module can lead to discussions of cutting edge methodologies that are used broadly in molecular genetics, including model organisms, promoter fusions, reporter

genes, and transgenic organisms. The latter topic can also be a jumping off point for a discussion of the social implication of transgenic crops.

The class project also provides opportunities for class discussions that can lead to a deeper understanding of biological concepts that are often difficult for beginning students to grasp in a lecture course. The analysis of TE sequences, which is part of the class project, can furnish a dramatic illustration of genetic selection. Specifically, students can perform multiple alignments of TEs and cellular genes and see that mutations accumulate in the former but not in the latter. When asked to explain this situation, most students replied that TEs must be more susceptible to mutation than cellular genes. This exposed a common misunderstanding of the mutational process and provided an opportunity to more fully explain the concepts of random mutation and purifying selection.

A model for other courses

The DG course can serve as a model for other research-intensive courses to train undergraduates in the life sciences. Replicating a research laboratory as an undergraduate classroom allows for synergy rather than competition between a professor's teaching and research programs. This synergy is manifest in many ways. A professor interested in developing such a course would first need to identify a biological problem related to his or her research interests. By focusing on their area of expertise, a professor is better positioned to identify key experiments in the discipline that illustrate the principles of experimental design and then develop cutting-edge research projects that originate in the research laboratory. Similarly, graduate students from the research laboratory are ideally suited to serve as course teaching assistants. They benefit the students because they are intimately familiar with the biology and methodology of the course material. They also derive benefit by their involvement in a positive teaching experience that is more meaningful than, for example, grading papers. Finally, the DG class model provides an ideal outlet for the broader impact requirements of research grant proposals. As mentioned previously, the maize TE polymorphism project described in the results had its origin in a research proposal to identify TEs in the newly sequenced maize B73 genome (Schnable *et al.* 2009). In addition to strengthening the research proposal, grant funds can then be used to offset some of the extra costs associated with running a cutting-edge laboratory course.

Converting "old" projects into new modules

Because the class project is tied to the ongoing research in the Wessler research laboratory, it needs to be periodically updated to reflect new research directions. With the successful completion of the maize TE annotation project, a new project focusing on using high throughput sequences to identify TEs in rice was initiated. However, rather than discarding the maize TE polymorphism project, it was recast as a module that now precedes the *A. thaliana* module in

current DG courses. This new module retains much of the pedagogy of the original project but requires only two 3-hr class periods (see [File S2](#)). In this way the class project serves two purposes. First, it is a platform for authentic scientific discovery for students. Second, it provides opportunities for instructors to devise and test new bioinformatics and/or experimental methodologies that can be converted into modules to update the tool kit taught during the first half of the course.

Acknowledgments

We thank Matthew Collin and Richard Cardullo for helpful comments on the manuscript. Leslie Jaworski of Grinnell College collected and analyzed the CURE data. This work was supported by a Howard Hughes Medical Institute Professor award (52005731) and National Science Foundation grants (IOS-1027542 and IOS-1118550) to S.R.W. as well as the University of Georgia Research Foundation and the University of California, Riverside. Institutional Review Board Project Number 2009-10835-0.

Literature Cited

- American Association for the Advancement of Science, 2011 Vision and Change in Undergraduate Biology Education: A Call to Action. Washington DC: American Association for the Advancement of Science. Available at: <http://visionandchange.org/files/2011/02/Vision-and-Change-low-res.pdf>. Accessed: March 16, 2012.
- Baucom, R. S., J. C. Estill, C. Chaparro, N. Upshaw, A. Jogi *et al.*, 2009 Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. *PLoS Genet.* 5: e1000732.
- Call, G. B., J. M. Olson, J. Chen, N. Villarasa, K. T. Ngo *et al.*, 2007 Genomewide clonal analysis of lethal mutations in the *Drosophila melanogaster* eye: comparison of the X chromosome and autosomes. *Genetics* 177: 689–697.
- Crowe, A., C. Dirks, and M. Wenderoth, 2008 Biology in bloom: implementing Bloom's taxonomy to enhance student learning in biology. *CBE Life Sci. Educ.* 7: 368.
- Derting, T. L., and D. Ebert-May, 2010 Learner-centered inquiry in undergraduate biology: positive relationships with long-term student achievement. *CBE Life Sci. Educ.* 9: 462–472.
- Edgar, R. C., 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792–1797.
- Handelsman, J., D. Ebert-May, R. Beichner, P. Bruns, A. Chang *et al.*, 2004 Education. Scientific teaching. *Science* 304: 521–522.
- Harrison, M., D. Dunbar, L. Ratmansky, K. Boyd, and D. Lopatto, 2011 Classroom-based science research at the introductory level: changes in career choices and attitude. *CBE Life Sci. Educ.* 10: 279–286.
- Hatfull, G. F., M. L. Pedulla, D. Jacobs-Sera, P. M. Cichon, A. Foley *et al.*, 2006 Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet.* 2: e92.
- Jiang, N., Z. Bao, X. Zhang, H. Hirochika, S. R. Eddy *et al.*, 2003 An active DNA transposon family in rice. *Nature* 421: 163–167.
- Kikuchi, K., K. Terauchi, M. Wada, and H.-Y. Hirano, 2003 The plant MITE mPing is mobilized in anther culture. *Nature* 421: 167–170.
- Li, H., M. Freeling, and D. Lisch, 2010 Epigenetic reprogramming during vegetative phase change in maize. *Proc. Natl. Acad. Sci. USA* 107: 22184–22189.
- Liu, K., M. Goodman, S. Muse, J. S. Smith, E. Buckler *et al.*, 2003 Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics* 165: 2117–2128.
- Lopatto, D., 2007 Undergraduate research experiences support science career decisions and active learning. *CBE Life Sci. Educ.* 6: 297–306.
- Lopatto, D., C. Alvarez, D. Barnard, C. Chandrasekaran, H. Chung *et al.*, 2008 UNDERGRADUATE RESEARCH: Genomics Education Partnership. *Science* 322: 684.
- Nehm, R. H., and L. Reilly, 2007 Biology majors' knowledge and misconceptions of natural selection. *Bioscience* 57(3): 263–272.
- Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. Wei *et al.*, 2009 The B73 maize genome: complexity, diversity, and dynamics. *Science (New York, NY)* 326: 1112–1115.
- Seymour, E., A.-B. Hunter, S. L. Laursen, and T. DeAntoni, 2004 Establishing the benefits of research experiences for undergraduates in the sciences: first findings from a three-year study. *Sci. Educ.* 88: 493–534.
- Yang, G., F. Zhang, C. N. Hancock, and S. R. Wessler, 2007 Transposition of the rice miniature inverted repeat transposable element mPing in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 104: 10962–10967.

Communicating editor: A. P. Mitchell

GENETICS

Supporting Information

<http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.112.147355/-/DC1>

Transposing from the Laboratory to the Classroom to Generate Authentic Research Experiences for Undergraduates

James M. Burnette, III, and Susan R. Wessler

File S1
DNA Extraction

This protocol is based on one described by Li *et al.* (2010) and has been modified to work in a college laboratory setting. Students need approximately 1.5 hours to extract DNA from up to five samples and the protocol can be carried out over several class periods by stopping at steps 7 and 9.

Materials list:

Extraction Buffer (100 mM Tris, pH 8.0, 50 mM EDTA and 500 mM NaCl)
10% SDS (sodium dodecyl sulfate)
5M KOAc (Potassium Acetate)
1.5 cm by 5 cm piece of Miracloth (Calbiochem, La Jolla, CA)
100% Isopropanol
70% Ethanol
Sterile water
Ice
Liquid nitrogen
65°C heating block
Sterile 1.5 ml tubes (2 for each prep)
Mortar and pestle

All chemicals were purchased from Fisher Scientific.

Method:

1. Label one tube for each plant.
2. Harvest 2-3 seedlings and place in a mortar. Fill with about 50 ml of liquid nitrogen. Grind tissue with pestle.
3. Add 1 ml of extraction buffer to the tube.
4. Add 120 μ l of 10% SDS. Mix by inverting.
If preparing more than one sample, prepare each sample to this step and place on ice.
5. Incubate tube(s) at 65°C for 20 minutes.
6. Add 300 μ l 5M KOAc. Mix well by inverting several times (important!), then place on ice 5 minutes. *Stopping point: Samples can be frozen for a future class period. Thaw samples before starting with step 7.*
7. Centrifuge for 5 minutes at >12,000 rpm. Label a second tube.
8. Pass 700 μ l of the supernatant through a miracloth funnel into the second tube.
9. Add 600 μ l of isopropanol. Mix the contents thoroughly by inverting. *Stopping point: Samples can be frozen for a future class period.*
10. Spin for 5 minutes at 14,000 rpm.
11. Carefully pour off and discard the supernatant. Use a P20 set to 20 μ l to remove the remaining drops of liquid without disturbing the DNA pellet.
12. Add 500 μ l of 70% ethanol and flick the tube until the pellet comes off the bottom.
13. Spin 5 minutes.
14. Pour off the ethanol. Use a P20 set to 20 μ l to remove the remaining drops without disturbing the pellet.
15. Leave the tube open on the bench to air dry for 5-10 minutes.
16. Resuspend the DNA in 50 μ l TE and incubate at room temperature for 5 minutes for complete resuspension. Samples should be frozen for storage.

File S2
Module version of the Polymorphism Project

The polymorphism project was used as the starting material for a new module that is now a permanent part of the first half of the Dynamic Genome course. The module version takes only 2 3-hour class periods (compared with XX periods for the project) because students do not order PCR primers, extract genomic DNA or sequence PCR bands. Instead, the instructors provide these materials. This module was used at UGA in three course offerings and one offering at UCR as well as in a high school course for four years. The cost is around \$8 per student.

The module starts with the students determining the location of the 169020 insertion site in the Maize Genome Browser. After a discussion of the meaning of a reference genome sequences (in this case B73), students are asked to consider the potential of sequence variation between different strains of maize. In the laboratory, students are provided with purified DNA from 16 strains and a primer pair for the 169020 insertion site. The 16 maize inbred strains are a subset of the 34 strains analyzed in the project (Table 1) so that half contain the insertion and half do not. Each student selects three strains for analysis and the reference strain B73. PCR amplicons are resolved by gel electrophoresis and, after gel analysis, students are provided with computer files of the amplicon sequences (available on the webpage dynamicgenome.ucr.edu/course_materials). They compare the sequences to B73 using multiple sequence alignment programs to reveal the different types of polymorphism, which are discussed in class.