Using computational molecular modeling software to demonstrate how DNA mutations cause phenotypes

Tara L. Phelps-Durr

Department of Biology, Radford University, Radford VA

Abstract

Students require a deep understanding of the central dogma before they can understand complex topics such as evolution and biochemical disorders. However, getting undergraduate biology students to apply higher-order thinking skills to the central dogma is a challenge. Students remember and regurgitate the molecular details of transcription and translation but if asked to apply these details, such as how a DNA mutation might affect phenotype, it becomes clear that most students do not deeply understand the central dogma. This lesson is a five-week series of laboratory activities designed to help students transition from applying lower order thinking skills to the central dogma to applying higher-order thinking skills. Over five weeks, students explore the phenotype of *Arabidopsis asymmetric leaves 1* (*as1*) and *as2* mutants. Students isolate DNA from wild-type and mutant plants and determine the sequence of the *AS1* and *AS2* alleles. Students use the DNA sequence data to determine the mutant protein amino acid sequences. They submit the mutant and wild-type protein sequences to a free online server and obtain three-dimensional (3-D) models of the wild-type and mutant proteins. They use free software to analyze and compare the 3-D models to determine the structural differences between the wild-type and mutant proteins. These computer-generated models can be 3-D printed allowing students to better visualize the protein structure. The overall goal is to use student-centered laboratory activities to demonstrate the relationship between DNA sequence, protein structure/function, and phenotype.

Supporting Materials: S1. Computational Molecular Modeling- Materials and Equipment, S2. Computational Molecular Modeling- Growing Arabidopsis, S3. Computational Molecular Modeling- AS1 AS2 Background Information, S4. Computational Molecular Modeling- Pre-Laboratory Assessment, S5. Computational Molecular Modeling- Week 1. Phenotypic Analysis of as1 and as2 mutants, S6. Computational Molecular Modeling- Week 2. DNA isolation and PCR, S7. Computational Molecular Modeling- Week 3. Gel electrophoresis and Bioinformatics, S8. Computational Molecular Modeling- Week 4. Analysis of PCR Sequences and Submission to I-TASSER, S9. Computational Molecular Modeling- Week 4. Mutation Homework, S10. Computational Molecular Modeling- Week 4. PCR Product Sequences, S11. Computational Molecular Modeling- Week 5. 3-D Protein Analysis, S12. Computational Molecular Modeling- Week 5. ICM-Browser Tips, S13. Computational Molecular Modeling- AS1 and AS2 Sequences and Primers, S14. Computational Molecular Modeling- Laboratory Report Guidelines and Rubric, S15. Computational Molecular Modeling- Sequence Alignments, S16. Computational Molecular Modeling- Petiole Length Raw Data, S17. Computational Molecular Modeling- AS2 normal_mutant ICM-Browser file, S18. Computational Molecular Modeling- AS1 normal_mutant ICM-Browser file, S19. Computational Molecular Modeling- AS2 normal.stl and S20. Computational Molecular Modeling- AS1 normal.stl.

*Correspondence to: 801 E Main St, Radford, VA 24142   E-mail: tphelpsdu@radford.edu

Learning Goal(s)

Students will understand the relationship between DNA sequence, protein structure, protein function, and phenotype.

Learning Objective(s)

Students successfully completing this lesson will:

- Practice basic molecular biology laboratory skills such as DNA isolation, PCR, and gel electrophoresis.
- Gather and analyze quantitative and qualitative scientific data and present it in figures.
- Use bioinformatics to analyze DNA sequences and obtain protein sequences for molecular modeling.
- Make and analyze three-dimensional (3-D) protein models using molecular modeling software.
- Write a laboratory report using the collected data to explain how mutations in the DNA cause changes in protein structure/function which lead to mutant phenotypes.

INTRODUCTION

The Vision and Change Report (V&C; www.visionandchange.org) is one of the more recent national calls to improve undergraduate biology education. The report lists five core concepts and six core competencies that all undergraduate biology majors should master. In addition, the report recommends designing courses so that students use biological facts in a way that promotes a deeper understanding of biology.

Designing courses that encourage students to use biological facts to promote a deeper understanding of biology is no trivial task. It is often easy for students to apply Lower Order Thinking Skills (LOTS), which include the remembering, understanding, and applying levels of Bloom's Taxonomy. However, it is more difficult for students to apply Higher-Order Thinking Skills (HOTS), which include the analyzing, evaluating, and creating levels of Bloom's Taxonomy (1). For example, students can often recite the central dogma (DNA>RNA>protein) but have very little understanding of what the arrows represent (2). Students often quickly memorize the molecular mechanisms of transcription and translation and confidently state that DNA sequence influences protein structure, which in turn influences protein function. However, when students encounter a new DNA mutation and must evaluate and clearly articulate how this DNA mutation might influence protein structure and function, it becomes clear that most students do not have a deep understanding of the relationship between DNA and protein.

I present a five-week series of laboratory activities designed to help students understand how DNA sequence influences protein structure/function and phenotype. The students first qualitatively and quantitatively analyze *Arabidopsis* phenotypes that result from mutations in either the ASYMMETRIC LEAVES1 (*AS1*) or *AS2* genes*. They then isolate DNA from wild-type and mutant plants and determine the sequence of the *AS1* and *AS2* genes by sequencing PCR products. The three-dimensional (3-D) structures of the wild-type and mutant AS1 and AS2 proteins are determined by combining bioinformatics with computational molecular modeling. At the end of the five-week activities, the students write a laboratory report to assemble their data and articulate the relationship between DNA sequence, protein structure/function, and phenotype. These activities are student centered, hands-on, and designed to help students transition from using LOTS to using HOTS when thinking about the central dogma.

Students must be able to apply HOTS to the central dogma if they are to understand complex topics including evolutionary processes and human disorders. One factor that makes it difficult for students to apply HOTS to the central dogma is that molecules and molecular processes cannot directly observe. Genetics textbooks often have good diagrams that compare wild-type and mutant DNA sequences followed by a picture that allows one to compare the structure of the wild-type and mutant proteins. However, these pictures are two-dimensional depictions of 3-D molecules that students passively observe. There are also many wonderful animations are (such as the DNA Learning Center 3-D animation library https://www.dnalc.org/resources/3d/) available that show the processes of transcription and translation. While most of these animations clearly show 3-D protein structure, students are still merely passive observers of these videos. To help students think in 3-D about protein structure, the laboratory activities presented here combine phenotype and genotype analysis with computational molecular modeling of proteins. The

activities enable students to discover the changes in DNA sequence and then to build comparative 3-D structures of the wild-type and mutant proteins.

These hands-on activities provide students the opportunity to transition from memorizing and regurgitating the central dogma to evaluating how the 3-D shape of proteins changes in response to DNA mutations. By the end of these activities, students gain a deeper understanding of the central dogma and are prepared to understand subsequent topics that are more complex.

> * The following nomenclature is used in this article. Italicized and all caps *AS1* and *AS2* refer to the wild-type sequence of the genes. All caps (not italicized) AS1 and AS2 refer to the protein sequence encoded by the genes. Italicized lowercase *as1* and *as2* refer to mutant alleles.

Intended Audience

Sophomore and Junior level Biology majors.

Learning Time

I teach this five-week series of activities in a course that meets on Monday, Wednesday and Fridays for fifty minutes of lecture and once each week for a required three-hour teaching laboratory. All the activities are performed in the teaching laboratory portion of the course, I cover the necessary background information, such as, transcription, translation, PCR, mutation, protein structure, and primary literature readings in the lecture portion of the course. As written, the first three weeks of the activities are molecular biology experiments and the last two weeks are computer laboratory activities.

Pre-requisite student knowledge

Students will need a basic understanding of DNA structure, mutation, gene function, DNA replication/PCR, protein structure, and the central dogma.

SCIENTIFIC TEACHING THEMES

Active learning

The pre-laboratory activities require the students to watch videos and to complete background reading or problem solving activities that will prepare them for the in-laboratory activities. At the beginning of the laboratory, the instructor should ask the students to share what they found with their partner and then the class as a whole should discuss the necessary background information. While the students are discussing with their partners, the instructor may consider checking that all students have completed the pre-laboratory activities. During the laboratory, students work in pairs to complete the molecular biology experiments but individually to complete the computational molecular modeling activities. To ensure that all students are actively engaged in the data collection, the instructor should interact with each team and ensure that all students actively participate in the experiments. At the end of the five-week series of activities, students individually write a laboratory report, which helps reinforce active participation by all students.

Assessment

The pre-assessment (Supporting File S4) is given to students prior to beginning any laboratory exercises. In the first part of the pre-assessment, the students use a set of terms to write

a paragraph that explains how mutations in DNA cause an abnormal phenotype. In the second part of the pre-assessment, students self-assess their previous experiences and comfort level with computer software. The purpose of the second part of the pre-assessment is to help the instructor determine how much guidance the students may need during the bioinformatics and computational molecular modeling activities. At the end of the five-week series of activities, the students individually write a laboratory report where they present and analyze the data collected throughout the activities. This assessment will help an instructor evaluate if a student can use HOTS in the process of explaining how mutations in the *AS1* and *AS2* genes change protein structure/function and cause a mutant phenotype. The laboratory report guidelines with a rubric are available to the instructor (Supporting File S14).

Inclusive teaching

Students work in pairs as they complete the molecular biology activities and individually during the computational biology activities. During each weekly activity, the instructor should ensure that the students are collaborating and working together as opposed to each student individually completing separate portions and then combining their efforts to complete the activity quickly. To help ensure that every student understands and contributes to each portion of the activity, I strongly suggest that each student individually writes the final assessment (i.e. the laboratory report).

Spatial skills have been correlated to success in STEM fields (3). Unfortunately, studies have also shown that there is a gender gap in performance on spatial skills assessments (4). Spatial skills competency can increase with practice (5). The activities presented here give all students a chance to practice their spatial skills by exploring 3-D structure at their own pace. It should also be noted that the molecular models of the proteins can be 3-D printed, which further benefits students who struggle with spatial perception and those students who are visually impaired.

LESSON PLAN

Please see the "Teaching Timeline" table at the end of this article for a description of the teaching timeline for the Lesson.

Supporting File S1 includes a list of materials and equipment required for each week's activities. Prior to beginning these activities, the instructor will need to grow the *Arabidopsis* plants required for the week one activities. The instructor can find complete instructions for growing Arabidopsis in Supporting File S2; plants will be ready for analysis approximately three weeks after sowing the seeds. Supporting File S3 provides instructors and students the required background information they need to understand the function of the *AS1* and *AS2* genes. Prior to starting the five-week series of laboratories, the instructor should give the students the pre-laboratory assessment (Supporting file S4). The detailed instructions for week's one, two and three activities are available in Supporting files S5, S6, and S7. The complete week four laboratory activities are described in Supporting file S8. Prior to attending the week 4 laboratory, the students should complete the homework assignment (Supporting file S9) and be given the sequences of the PCR products (Supporting file S10). The complete week five laboratory instructions is provided in Supporting file S11 and extra tips for navigating the molecular modeling software, ICM-Browser, is provided in Supporting file S12. The

detailed weekly laboratory instructions (Supporting files S5, S6, S7, S8 and 12) include both the student and instructor pre-laboratory preparations for each week. Supporting Files S5, S6, S7, S8 and S12 also contain a student laboratory handout that students fill out and turn in at the end of each weekly laboratory session. In addition to growing the plants for the molecular biology activities, the instructor must also become familiar with the computational modeling program, ICM-Browser (http://www.molsoft.com/icm_browser.html), used to analyze the 3-D structure of proteins in these activities. Prior to the start of this five-week series, the instructor should (at a minimum) complete the protein modeling using the provided "Easy 3-D Protein Modeling Instructions," which are available in Supplemental file S11.

Week 1:

The students examine and document the phenotype of wild-type and mutant *Arabidopsis* plants. The mutant *Arabidopsis* lines have a mutation in either the *AS1* or the *AS2* gene. Here, I will provide a brief description of these genes; however, the instructor should review Supporting File S3 for a detailed description. *AS1* and *AS2* encode proteins that physically interact and are required to suppress the expression of *KNOTTED1-like* homeobox (*KNOX*) genes during leaf development (6,7). In wild-type *Arabidopsis* plants, the KNOX genes are expressed only in the undifferentiated meristem cells (8). Expression of the KNOX genes in the differentiating leaf cells is repressed because the *AS1* and *AS2* genes are expressed in the leaf tissue (6,7,8). In *as1* and *as2* mutants, *KNOX* gene expression is not completely repressed in the developing leaves resulting in abnormal leaf development (6,7,8,9,10). During the week one activities, the students take pictures of the wild-type and mutant plants and make a figure that clearly shows the mutant phenotypes. In addition, since one of the most notable phenotypes is the variation in petiole length, the students measure the petioles of the wild-type and mutant plants and graphically display the class results.

Weeks 2-4:

Once the phenotypes have been analyzed during week one, the students spend three weeks determining and analyzing the sequence of the *AS1* and *AS2* genes isolated from wild-type, *as1*-mutant, and *as2*-mutant plants. During week two, the students isolate DNA and PCR amplify the regions of the *AS1* and *AS2* genes that contain the mutations. In week three, students analyze the PCR products by gel electrophoresis. While the students are waiting for the gel electrophoresis, they use bioinformatics tools to obtain the mRNA and protein sequences of the wild-type *AS1* and *AS2* genes. At the end of the week three activities, the class submits the wild-type AS1 and AS2 protein sequences to the protein modeling server I-TASSER (11; http://zhanglab.ccmb.med.umich.edu/I-TASSER/) to obtain 3-D structures of the wild-type proteins. In week four, the students analyze the sequences of the PCR products from week two. As part of the sequence analysis, the students compare the sequence of the *AS1* and *AS2* PCR products amplified from DNA isolated from wild type, and as1 and as2 mutant plants. The students compare the PCR sequences by using a free online sequence alignment resource (CLUSTALW; http://www.genome.jp/tools/clustalw/). When aligned with the wild-type *AS1* allele, the *as1* allele is a single base pair deletion at position 1227 and the *as2* mutation is a 13 bp deletion beginning at position 1365 (See highlighted sequences in Supplemental files S13 and S15).

The students analyze the sequence alignment information and change the sequence of the wild-type *AS1* and *AS2* mRNAs that they retrieved from the *Arabidopsis* Information Resource website (https://www.arabidopsis.org/) in week three into the sequences of the mutant mRNAs. The instructions in Supporting file S8 explain how to identify the start codon and submit the mutant mRNA sequences to the online resource ExPasy (http://web.expasy.org/translate/), which translates the mutant mRNA into the amino acid sequence of the mutant AS1 and AS2 proteins. At the end of the week-four activities, the class submits the mutant AS1 and AS2 protein sequences to I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) to obtain 3-D structures of the mutant proteins.

Week 5:

The students analyze the 3-D models of the wild-type and mutant AS1 and AS2 proteins. I have given the instructor two options for completing the computational molecular modeling activities (Supporting File S11). The easy modeling instructions allow the students to use ICM-Browser modeling files (Supporting Files S17 and S18) that have pre-loaded 3-D protein structures from I-TASSER. The advantage to using these files is that the modeling instructions in Supporting File S11 were written to precisely fit the wild-type and mutant AS1 and AS2 protein models pre-loaded into the ICM-Browser files (Supporting Files S17 and S18). The disadvantage to using the easy instructions is that the activities are less inquiry-based. If the instructor chooses to use the easy instructions, students will not analyze the 3-D protein models they generated in I-TASSER in weeks three and four because the easy instructions use ICM-Browser files (Supporting Files S17 and S18) that contain pre-loaded 3-D models of the AS1 and AS2 wild-type and mutant proteins. For instructors using the easy instructions, I still recommend submitting the amino acid sequences to I-TASSER in weeks three and four because the activity shows the students how the 3-D protein models in Supporting files S17 and S18 were made. For the most inquiry-based approach, the instructor should use the "Advanced 3-D Protein Modeling Instructions" that require the instructor and students to analyze the protein models that the class retrieved from I-TASSER. However, given the constant improvement of molecular modeling programs, it is likely that future models of AS1 and AS2 generated in I-TASSER will differ from the models in the provided files (Supporting Files S17 and S18). Therefore, I do not provide detailed instructions for how to analyze future 3-D models. Instructors who choose to carry out the advanced activity with students should first complete the easy activity on their own, to familiarize them with the modeling software. Both the easy and advanced modeling instructions guide the students through assigning secondary structure, visualizing the 3-D models in different representations (ribbon vs. ball and stick vs. space-filling), and displaying the wild-type and mutant proteins in different colors. Throughout the modeling activities, the students collect images to record their analyses of the 3-D models. The students also analyze the amino acid sequence alignments of the wild-type vs. mutant proteins. For the easy instructions, the sequence alignment is already pre-loaded in Supporting Files S17 and 18. The instructions for obtaining and importing the amino acid sequence alignment into the molecular modeling software are included with the advanced protein modeling instructions.

Supporting File S15 provides the sequence alignments of the wild-type and mutant AS1 and AS2 sequences with the mutations highlighted in yellow. The *as1* mutant allele contains
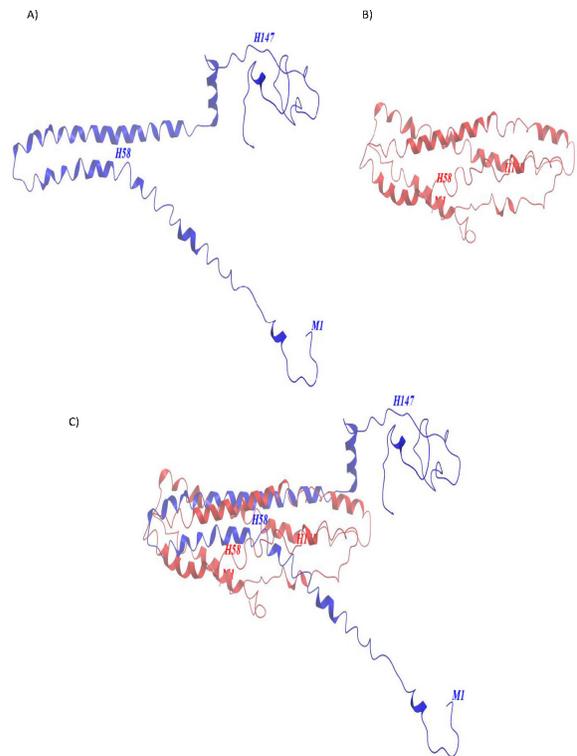

Figure 1. Examples of 3-D structures of the wild type AS2 (A), mutant AS2 (B) and superimposed wild type and mutant AS2 proteins (C) made in ICM-Browser

a single bp deletion and the *as2* mutant allele contains a 13 bp deletion, both of which result in a frameshift during translation. The wild-type and mutant AS1 and AS2 proteins are identical prior to the DNA deletion and dramatically different in sequence after the deletion. Students analyze the AS1 and AS2 protein sequence alignments and predict what regions of the wild-type and mutant proteins will have the same 3-D shape. Students typically assume that the wild-type and mutant proteins will have the same 3-D shape in the region where the wild-type and mutant proteins are identical in sequence but a different 3-D shape where the mutant sequence differs due to the frameshift. After analyzing the protein sequence alignments, the instructions guide students through the process of labelling an amino acid located at the beginning, middle, and end of the wild-type and mutant proteins. Students then superimpose the wild-type and mutant 3-D protein structures to visualize which, if any, regions of the wild-type and mutant proteins have the same 3-D structure, as well as what regions differ (Figure 1).

After completing the week five activities, the students are ready to assemble the data into a laboratory report. I give the students one week to complete the laboratory report. I designed these activities for a sophomore level genetics course and many of the students have never written a laboratory report prior to taking the genetics course. Therefore, I provide students a detailed laboratory report guidelines and rubric (Supporting File 14) to guide them through the writing process. To further help students complete the final report, I embedded the data collection and figure making in the weekly activities. For example, as part of the week one activities, the students make a figure that displays the wild-type and mutant phenotypes of

the plant and they make a graph of the petiole length data. The students are required to include both the figure and the graph in the laboratory report. If the students have properly completed the weekly activities, they should have all of the images and figures required to complete the laboratory report. I use the laboratory report as the final assessment to determine how well the students understand that mutations at the DNA level cause changes in protein structure and function, which in turn results in the observed *as1* and *as2* mutant phenotypes.

Options to Extend the Lesson

Extension 1: The Evolution of Compound Leaves

The suppression of *KNOX* genes by AS1 and AS2 is not a regulatory relationship conserved in all plants. The *KNOX* genes are expressed in the meristem tissue of *Arabidopsis* and maintain the undifferentiated meristem cell population. As cells are recruited from the meristem cell population and differentiate into leaves, the expression of the *KNOX* genes must be suppressed by AS1 and AS2 in order for wild-type leaves to form. Thus in *Arabidopsis*, a simple leaf plant, AS1 and AS2 are expressed in the leaf tissue and the *KNOX* genes are expressed in the meristem. In plants with compound leaves, such as tomato, the *AS1* and *KNOX* gene orthologues are both expressed in the leaf tissue (12). Thus, the regulation of *KNOX* gene expression is likely a factor that contributed to the evolution of compound leaves (13). The instructor could incorporate this information into the lesson by having the students read the review article by Bharathan and Sinha (13) and having an in-laboratory discussion that introduces students to the concept that changes in gene regulation lead to developmental variation and the evolution of new species.

Extension 2: Identifying and Understanding the Function of Protein Domains

AS1 is a myb domain transcription factor and *AS2* contains a leucine zipper-like domain (see Supporting File S3 for more detail). The activities I have presented do not focus on the location or function of either of these domains. The instructor could have the students BLAST the amino acid sequence of the normal AS1 and AS2 proteins. During the BLAST process, an image of the conserved protein domains will appear. Clicking on the superfamily name, in the case of AS1 the SANT superfamily, will result in a list of domain hits that will open in a new window. In the new window, the location of the domain and a description of the domain function is provided. The instructor could use this information to have the students label the domains in the AS1 and AS2 protein models as they complete the computational modeling activities (Supporting File S11). The instructor could then explain how protein domains define protein function.

Extension 3: 3-D Printing the Protein Models

The molecular modeling software, ICM-Browser, allows students to rotate the 3-D protein models, providing students the opportunity to experience the 3-D structure of the proteins. Despite the excellent ICM-Browser graphics, some students will still struggle to visualize the 3-D structure of the proteins on the computer screen. The instructor could 3-D print the normal AS1 and AS2 proteins to help students grasp the 3-D structure of these proteins. Supporting File S19 is a .stl file of

Table 1 Learning goals from the activities presented here aligned with Bloom's Taxonomy.

Two level classification	Bloom's Taxonomy	Specific learning goals
LOTS	Remember	Recite central dogma
	Understand	Explain differences between DNA, RNA and proteins
	Apply[1]	Determine amino acid sequence of a protein from a DNA sequence
HOTS	Analyze	Compare the wild type and mutant DNA and protein sequences
	Evaluate	Predict how changes in amino acid Sequence affects the 3-D structure of a protein
	Create	Write a laboratory report that clearly presents and describes the data

[1]Apply can fit in the LOTS or HOTS category depending on the complexity of the task requiring application.

the normal AS2 protein and S20 is a .stl file of the normal AS1 protein.

TEACHING DISCUSSION

The overall goal of this series of hands-on activities is to help students understand the relationship between DNA sequence, protein structure/function, and phenotype and to help students transition from using lower order thinking skills to using higher order thinking skills when thinking about the central dogma. Table 1 summarizes how the specific learning goals in these activities fit into Bloom's Taxonomy.

In terms of Bloom's Taxonomy, LOTS are defined as a student's ability to remember, understand and apply a topic (14). In terms of the classroom activities presented here, students who can recite the central dogma, explain the differences between DNA, RNA and proteins, and provide a basic explanation of the underlying molecular processes of the central dogma are exhibiting LOTS. In terms of Bloom's Taxonomy, HOTS are defined as a student's ability to analyze, evaluate and create using basic knowledge (14). In terms of the classroom activities presented here, students who can determine the amino acid sequence of a protein when given a DNA sequence are applying the central dogma. Students who can compare a wild-type and mutant DNA sequence and the resulting wild-type and mutant proteins are using basic ideas of the central dogma to analyze. Students who can predict how changes in the amino acid sequence of the mutant will affect the three-dimensional structure and function of the protein are using the central dogma to evaluate. Finally, students who can write a good laboratory report and articulate how mutations in the DNA ultimately cause a mutant phenotype are using data to create.

Students benefit the most from these laboratory activities if the lecture and laboratory portion of the course are aligned so that as students complete the laboratory activities they are simultaneously learning the about the central dogma and the structural differences between DNA, RNA and proteins during the lecture portion of the course. Aligning the lecture and laboratory also allows the instructor to assess central dogma-related concepts at different levels. For example, I assess LOTS during the lecture portion of the course as the students are covering basic content by giving quizzes and in-class practice problems and I assess HOTS by having the student complete the laboratory report at the end of the laboratory activities.

During the laboratory activities, the students will apply their basic understanding of the central dogma to determine how

the amino acid sequence of a protein will change because of a particular mutation in the DNA sequence. This task is challenging for the students for two reasons. First, the students must compare the sequences of PCR products they generate in the laboratory and discover the mutations in the *AS1* and *AS2* genes. In my experience, students struggle with this task simply because they have to complete and analyze the results of several different types of experiments before they discover the mutations. Second, the mutations in the *AS1* and *AS2* genes are a single base pair (bp) and a 13 bp deletion, respectively, that result in frameshift mutations. Most students can work together and quickly grasp how the single bp mutation causes a frameshift; however, students often have difficulties understanding how the 13 bp deletion causes a frameshift.

Where the students really start to struggle in these activities is when they have to start thinking in 3-D. Before students use the computational modeling program to examine the 3-D structure of the wild-type and mutant AS1 and AS2 proteins, they predict whether the wild-type and mutant proteins will have regions that fold into the same 3-D structure. Almost all students predict that the mutant proteins will differ structurally from the wild-type proteins only in regions where the amino acid sequence differs between the mutant and wild-type proteins. This prediction indicates that students do not understand how amino acid side chains from various regions of the protein interact and specify the 3-D shape of a protein. Because students have misconceptions about how proteins acquire their 3-D structures, they cannot clearly articulate how a change in 3-D structure affects protein function and how this change causes a mutant phenotype.

Using the 3-D computational modeling software, ICM-Browser, helps students understand the 3-D structure of proteins. Within the software, students can rotate, label, color, zoom in and out, and change the representation of the protein (ball and stick, protein backbone, or space-filling). As the students manipulate the 3-D structure of the protein, they acquire a better understanding of the protein's structure and it becomes easier for them to understand how structure relates to function. In the laboratory activities presented here, the students use ICM-Browser to superimpose and compare the wild-type and mutant 3-D structures of the AS1 and AS2 proteins. Once the students see the 3-D differences between the wild-type and mutant proteins, they gain a better understanding of why the 3-D structure of a protein directly relates to its function. By the time the students analyze the 3-D structure of the proteins they have already analyzed the phenotype and genotype of the *as1* and *as2* mutants which helps them make the connection between DNA sequence, protein structure/function, and phenotype.

After the students have completed the laboratory activities, they individually write a laboratory report. Detailed laboratory report guidelines are provided in the Supporting File S14. If the students have not had much experience with scientific writing, they may struggle with writing the laboratory report. However, if completed as written, the laboratory exercises (Supporting Files S5, S6, S7, S8 and S11) and the laboratory report guidelines (Supporting File S14) walk students through the process of making all of the figures required for their laboratory report. Despite having made the figures throughout the five-week series of activities, students may find it challenging to analyze and present the figures in a written report. One way to prevent overwhelming the students with the writing of the laboratory report is to have the students write the laboratory report one section at a time and provide the students with feedback before they complete the next section. Part of the week one activities requires the students to make and explain figures of the *as1* and *as2* phenotypes and the petiole length data. I provide a lot of feedback on the presentation of these figures so that my writing expectations are clear early in the project. If, based on my assessment of the week one figures, the students' writing skills seem weak I will have them write the introduction to the laboratory report after they have completed the week two activities. This gives me another opportunity to provide students feedback on their writing but it also helps me verify that the students have an understanding of the basic background information (Supporting File S3). Finally, the instructor could have the students write the majority of the materials and methods section of the laboratory report once the students complete the week four activities.

When presenting these laboratory exercises to students, a few logistical problems will arise that an instructor will need to address. First, these activities are likely to be the first time that students, and perhaps the instructor, have used a molecular modeling program such as ICM-Browser. This program is one of the more user-friendly molecular modeling programs and is not overly difficult to learn. However, the instructor will need to encourage students to be patient, use the help features, and follow the instructions carefully while they are learning how to use the program. Second, when students are performing computer-centered activities, it is possible for the students to click through and complete the exercises without meaningfully immersing themselves into the activity. The instructor can prevent "click through" by interacting with each pair of students intermittently in a way that assesses the students' understanding of and progression through the activity.

Finally, as they complete this sequence of activities, students are practicing the skills required to master three of the V&C Core Competencies: Process of Science; Quantitative Reasoning; and Modeling/Simulation. These activities also align with two V&C Core Concepts: Information Flow and Structure/Function. If an instructor chooses to do so, they could incorporate a third V&C Core Concept, Evolution, by having students read articles that discuss how the regulation of KNOX gene expression has contributed to the evolution of compound leaves in plants (6,7,8,9,10,15,13).

SUPPORTING MATERIALS

- S1. Computational Molecular Modeling- Materials and Equipment
- S2. Computational Molecular Modeling- Growing *Arabidopsis*
- S3. Computational Molecular Modeling- AS1 AS2 Background Information
- S4. Computational Molecular Modeling- Pre-Laboratory Assessment
- S5. Computational Molecular Modeling- Week 1. Phenotypic Analysis of as1 and as2 mutants
- S6. Computational Molecular Modeling- Week 2. DNA isolation and PCR
- S7. Computational Molecular Modeling- Week 3. Gel electrophoresis and Bioinformatics
- S8. Computational Molecular Modeling- Week 4. Analysis of PCR Sequences and Submission to I-TASSER
- S9. Computational Molecular Modeling- Week 4. Mutation Homework
- S10. Computational Molecular Modeling- Week 4. PCR

Product Sequences
- S11. Computational Molecular Modeling- Week 5. 3-D Protein Analysis
- S12. Computational Molecular Modeling- Week 5. ICM-Browser Tips
- S13. Computational Molecular Modeling- AS1 and AS2 Sequences and Primers
- S14. Computational Molecular Modeling- Laboratory Report Guidelines and Rubric
- S15. Computational Molecular Modeling- Sequence Alignments
- S16. Computational Molecular Modeling- Petiole Length Raw Data
- S17. Computational Molecular Modeling- AS2 normal_mutant ICM-Browser file
- S18. Computational Molecular Modeling- AS1 normal_mutant ICM-Browser file
- S19. Computational Molecular Modeling- AS2 normal.stl
- S20. Computational Molecular Modeling- AS1 normal.stl

REFERENCES

1. Wilson LO, Leslie C. 2013. Anderson and Krathwohl - Understanding the New Version of Bloom's Taxonomy 1-7.
2. Wright LK, Fisk JN, Newman DL. 2014. DNA -> RNA: What Do Students Think the Arrow Means? Cell Biol Educ 13:338-348.
3. Uttal DH, Cohen C a. 2012. Spatial Thinking and STEM Education. When, Why, and How?Psychology of Learning and Motivation - Advances in Research and Theory. Elsevier Inc.
4. Sorby S a. 1990. Developing 3-D Spatial Visualization Skills SherylA . Sorby. Eng Des Graph J 63:21-32.
5. Uttal DH, Meadow NG, Tipton E, Hand LL, Alden AR, Warren C, Newcombe NS. 2012. The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. Psychol Bull 139:352-402.
6. Xu L, Xu Y, Dong A, Sun Y, Pi L, Xu Y, Huang H. 2003. Novel as1 and as2 defects in leaf adaxial-abaxial polarity reveal the requirement for ASYMMETRIC LEAVES1 and 2 and ERECTA functions in specifying leaf adaxial identity. Development 130:4097-4107.
7. Guo M, Thomas J, Collins G, Timmermans MCP. 2008. Direct repression of KNOX loci by the ASYMMETRIC LEAVES1 complex of Arabidopsis. Plant Cell 20:48-58.
8. Ori N, Eshed Y, Chuck G, Bowman JL, Hake S. 2000. Mechanisms that control knox gene expression in the Arabidopsis shoot. Development 127:5523-5532.
9. Byrne ME, Barley R, Curtis M, Arroyo JM, Dunham M, Hudson a, Martienssen R a. 2000. Asymmetric leaves1 mediates leaf patterning and stem cell function in Arabidopsis. Nature 408:967-971.
10. Lin W, Shuai B, Springer PS. 2003. The Arabidopsis LATERAL ORGAN BOUNDARIES-domain gene ASYMMETRIC LEAVES2 functions in the repression of KNOX gene expression and in adaxial-abaxial patterning. Plant Cell 15:2241-2252.
11. Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5:725-738.
12. Kim M, Pham T, Hamidi A, McCormick S, Kuzoff RK, Sinha N. 2003. Reduced leaf complexity in tomato wiry mutants suggests a role for PHAN and KNOX genes in generating compound leaves. Development 130:4405-4415.
13. Bharathan G, Sinha NR. 2001. Update on leaf development the regulation of compound leaf development. Plant Physiol 127:1533-1538.
14. Costa ' s and Bloom ' s Levels of Thinking : Comparison Chart.
15. Bharathan G, Goliber TE, Moore C, Kessler S, Pham T, Sinha NR. 2002. Homologies in leaf form inferred from KNOXI gene expression during development. Science 296:1858-1860.

Table 2. Computational Molecular Modeling-Teaching Timeline

Activity	Description	Estimated Time	Notes
Preparation for Five Week Laboratory Series			
Order supplies	Review the list of materials and methods and order supplies.	1 hour	A list of the materials and equipment needed is available in Supporting file S1.
Plant *Arabidopsis*	Plant seeds approximately 3 weeks prior to meeting for the week 1 laboratory.	1 hour	Detailed instructions for planting *Arabidopsis* are provided in Supporting file S2.
Download ICM-Browser	ICM-Browser is the computational molecular modeling program used in these activities to analyze 3-D models of proteins.	30 minutes	ICM-Browser is free and can be downloaded from: http://www.molsoft.com/icm_browser.html
Review computational molecular modeling activity.	Instructors who are unfamiliar with computational molecular modeling should complete the easy modeling instructions available in Supporting file S11.	2 hours	The complete easy modeling instructions are available in Supporting file S11. The pre-made ICM-Browser files needed to complete the easy modeling instructions are Supporting files S17 and S18.

Activity	Description	Estimated Time	Notes
Pre-laboratory Assessment			
Assign the students the pre-laboratory assessment approximately one week prior to starting the week 1 activities.	This assessment determines how well the students understand the relationship between DNA sequence and protein structure prior to completing the activities. I allow the students to complete the assessment outside of the laboratory.	10 minutes	The pre-laboratory assessment is provided as Supporting file S4.
Preparation for Week 1			
Give students the pre-laboratory preparation instructions.	The student pre-laboratory assignment asks the students to read the AS1_2 background information document and find images of plant phenotypes in the scientific literature.	10 minutes	The student pre-laboratory instructions are embedded in Supporting file S5. The AS1_2 background information is Supporting file S3.
Prepare for the in-laboratory discussion.	Prepare to discuss good figure making techniques and basic leaf structure.	30 minutes	The instructor notes for the in-laboratory discussion are available in Supporting file S5.
Setup the laboratory.	Students will need: 1. An electronic version of the student laboratory protocol. 2. Wild type, as1 and as2 mutant Arabidopsis plants. 3. Rulers 4. Forceps 5. Razor blades 6. Computer with Word and Excel 7. Camera (or cell phone)	1 hour	Detailed instructor pre-laboratory notes are provided in Supporting file S5. The electronic version of the student laboratory protocol is available in Supporting file S5.
Complete in-laboratory activity.	Phenotype analysis of *as1* and *as2* mutants.	2 hours	If the instructor cannot grow *Arabidopsis*, raw petiole length data for students to analyze is provided in Supporting file S16.
Preparation for Week 2			
Give students the pre-laboratory preparation instructions and the genomic sequences of the AS1 and AS2 genes and PCR primers.	The student pre-laboratory assignment asks the students to watch videos that demonstrate PCR and DNA replication. They will also use the AS1 and AS2 sequences to predict the size of the PCR products. Finally, students need to identify a scientific article that describes PCR in the materials and methods section of the paper.	10 minutes	The student pre-laboratory instructions are embedded in Supporting file S6. The AS1 and AS2 genomic sequences and PCR primer sequences are provided in Supporting file S13.
Prepare for the in-laboratory discussion.	Prepare to discuss how PCR works. If necessary help students determine the predicted PCR product sizes. Discuss how authors write materials and methods sections.	30 minutes	The instructor notes for the in-laboratory discussion are available in Supporting file S6.
Setup the laboratory.	Students will need: 1. An electronic version of the student laboratory protocol. 2. Wild type, as1 and as2 mutant Arabidopsis plants. 3. DNA isolation reagents. 4. Pipettors and tips. 5. Tubes. 6. PCR reagents.	1 hour	Detailed pre-laboratory instructor notes are provided in Supporting file S6. The electronic version of the student laboratory protocol is available in Supporting file S6.
Complete in-laboratory activity.	DNA isolation and PCR amplification of AS1 and AS2.	3 hours	

Activity	Description	Estimated Time	Notes
Preparation for Week 3			
Give students the pre-laboratory preparation instructions.	The student pre-laboratory instructions asks the students to watch videos that demonstrate gel electrophoresis.	10 minutes	The student pre-laboratory instructions are embedded in Supporting file S7.
Prepare for the in-laboratory discussion.	Prepare to discuss how gel electrophoresis works and have a practice gel prepared to demonstrate loading DNA samples.	30 minutes	The instructor notes for the in-laboratory discussion are available in Supporting file S7.
Setup the laboratory.	Students will need: 1. An electronic version of the student laboratory protocol. 2. The PCR samples from the previous week. 3. Loading buffer. 4. A DNA molecular weight marker. 5. Pipettors and tips. 6. A computer with internet access.	1 hour	Detailed pre-laboratory instructor notes are provided in Supporting file S7. The electronic version of the student laboratory protocol is available in Supporting file S7.
Complete in-laboratory activity.	Gel electrophoresis and bioinformatics.	3 hours	

Activity	Description	Estimated Time	Notes
Preparation for Week 4			
Give students the pre-laboratory preparation instruction and the mutation homework problem.	The student pre-laboratory instructions asks the students to watch a video that demonstrates how DNA sequencing works.	10 minutes	The student pre-laboratory instructions are embedded in Supporting file S8. The mutation homework is provided in Supporting file S9.
Prepare for the in-laboratory discussion.	Prepare to discuss how DNA sequencing works and go over the homework problem.	30 minutes	The instructor notes for the in-laboratory discussion are available in Supporting file S8.
Setup the laboratory.	The students will need: 1. An electronic version of the student laboratory protocol. 2. A computer with internet access. 3. The AS1 and AS2 PCR products amplified from wild type, as1 and as2 mutant plants.	10 minutes	Detailed pre-laboratory instructor notes are provided in Supporting file S8. The electronic version of the student laboratory protocol is available in Supporting file S8. The AS1 and AS2 PCR sequences are provided in Supporting file S10.
Complete in-laboratory activity.	Analysis of PCR sequences.	3 hours	Samples of the sequence alignments that the students make during this laboratory session are provided in Supporting file S15.
Preparation for Week 5			
Give students the pre-laboratory preparation instructions and the document with the top 10 tips for using ICM-Browser. 3-D print the AS1 and AS2 normal proteins (not required).	The student pre-laboratory instructions asks the students to review the top 10 tips document and perform an internet search to learn about myb and leucine zipper protein domains. The students need to download ICM-Browser prior to the laboratory.	10 minutes	The student pre-laboratory instructions are embedded in Supporting file S11. The Top 10 General ICM-Browser Tips document is provided in Supporting file S12. ICM-Browser is free and can be downloaded from: http://www.molsoft.com/icm_browser.html Supporting files S19 and S20 are .stl files of the AS1 and AS2 normal proteins. These files can be 3-D printed.
Prepare for the in-laboratory discussion.	Prepare to discuss myb and leucine zipper protein domains and demonstrate the ICM-Browser software.	1 hour	The instructor notes for the in-laboratory discussion are available in Supporting file S11.
Setup the laboratory.	The students will need: 1. An electronic version of the student laboratory protocol. 2. A computer with internet access and ICM-Browser. 3. The top 10 tips for using ICM-Browser. 4. The ICM-Browser files (easy modeling instructions) OR the .pdb files of the AS1 and AS2 proteins created in weeks 3 and 4.	10 minutes	Detailed pre-laboratory instructor notes are provided in Supporting file S11. The ICM-Browser files are provided as Supporting files S17 and S18.
Complete in-laboratory activity.	3-D analysis of proteins.	3 hours	

Activity	Description	Estimated Time	Notes
Preparation for Student Laboratory Report			
Give students the laboratory report guidelines at least one week in advance of the laboratory report due date.	Review the laboratory report guidelines and give students a format for the references.	30 minutes	The laboratory report guidelines are provided as Supporting file S14.