JOURNAL OF MICROBIOLOGY & BIOLOGY EDUCATION
DOI: https://doi.org/10.1128/jmbe.v20i2.1650

JMBE Journal of Microbiology & Biology Education

Supplemental Materials

for

Undergraduates Phenotyping *Arabidopsis* Knockouts in a Course-Based Undergraduate Research Experience: Exploring Plant Fitness and Vigor Using Quantitative Phenotyping Methods

Courtney J. Murren[1], Michael J. Wolyniak[2], Matthew T. Rutter[1], April M. Bisner[1], Hilary S. Callahan[3], Allan E. Strand[1], and Lisa A. Corwin[4*]*

[1]Department of Biology, College of Charleston, Charleston, SC 29424; [2]Department of Biology, Hampden-Sydney College, Hampden-Sydney, VA 23943; [3]Department of Biological Sciences, Barnard College, Columbia University, New York, NY 10027; [4]Department of Ecology and Evolutionary Biology, University of Colorado at Boulder, Boulder, CO 80309

Table of Contents

(Total pages 59)

Appendix 1: Materials list for unPAK CURE, pre-CURE setup, and phenotype protocols
Appendix 2: Student activity 1
Appendix 3: Datasheet 1
Appendix 4: Student activity 2
Appendix 5: Datasheet 2
Appendix 6: Student activity 3
Appendix 7: Writing assignment 1
Appendix 8: Student activity 4
Appendix 9: Student activity 5
Appendix 10: Instructor-completed examples for SA1, SA2, and SA3
Appendix 11: Supplemental table—Faculty feedback

*Corresponding authors. Mailing address: Courtney J. Murren, College of Charleston, Department of Biology, 66 George Street, Charleston, SC 29424. Phone: 843-953-8077. E-mail: murrenc@cofc.edu.
Lisa A. Corwin, Mailing address: Department of Ecology and Evolutionary Biology, University of Colorado at Boulder, 1900 Pleasant St. 334 UCB, Boulder, CO 80309. Phone: 303-735-5213. E-mail: lisa.corwin@colorado.edu.
Received: 3 July 2018, Accepted: 20 December 2018, Published: 28 June 2019.

unPAK CURE

Supplemental Materials Table of Contents

Appendix 1: Materials List unPAK CURE, PreCURE set up (contact information, seed information, and experimental design steps), **Phenotype Protocols** (S2-S9)

Appendix 2: Student Activity 1 (SA1). Background on Arabidopsis: the ecology and genetics in response to climate change (S10-S16)

a. Natural populations group (S12)
b. Mutant group (S13)

Appendix 3: Datasheet 1 (D1). Rosette traits and metadata (S17-S18)

Appendix 4: Student Activity 2 (SA2). Literature and Database Background

c. Natural populations group (S19-S21)
d. Mutant group (S22-S24)

Appendix 5: Datasheet 2 (D2). Fruit number and metadata (S25-S26)

Appendix 6: Student Activity 3 (SA3). Hypotheses, graphing, and statistics: *Arabidopsis thaliana* (S27-S29)

Appendix 7: Written Assignment 1 (WA1). Specific instructions including rubric. (S30-S37)

Appendix 8: Student Activity 4 (SA4). Peer review (S38-S41)

Appendix 9: Student Activity 5 (SA5). Student data curation example (S42-S45).

Appendix 10: Instructor completed examples for SA1, SA2 and SA3 (S46-S57)

Appendix 11: Supplemental Table Faculty Feedback (S58)

Appendix 1: Materials List, PreCURE set up, & Phenotype Protocols

Supplemental Table 1: Materials list for unPAK CURE. See PreCURE set up and experimental design steps below to identify the appropriate number of plants per student group.

Item General Name	**Item Specific Name**	**Vendor**	**Item #**	**Items for**
Filter Paper (for sowing seed)	Filter Paper Qualitative Size 7.5cm	VWR	28310-026	For instructor setting up whole project
Petri dishes (Round)	100 x 15 mm Stackable Petri Dish	VWR	25384-342	For instructor setting up whole project
Beakers	50mL Beakers	VWR	10754-946	For instructor setting up whole project
Thin Paint Brushes		Any Arts and Crafts Store		For instructor setting up whole project
Wash Bottles	VWR Wash Bottles 125 mL	VWR	414004-225	For instructor setting up whole project
Plastic Wrap	Saran Wrap or other plastic wrap	Grocery Store		For instructor setting up whole project
Pot tag labels	Weatherproof Polyester Laser 1.5” x 0.5” SC- 100 Labels / Sheet	Online Labels.com	OL1100LP	For each plant
Pot Tags (Plastic labels)	Preferably 4” so they fit under domes	Hummert.com		For each plant
2.5 inch pots	02.50 Kord Trad SQ	Myerslawnandgarden.com	SQN02500	For each plant
Trays	Perma-Nest plant trays 22”x 11”x 2.5”	Hummert.com	65-6963-9	For each tray of plants
Potting Medium	ProMix BX	Varied sources but type of potting mix matters garden center or amazon		Whole project
Tray Covers (domes)	Perma-Nest dome 22x11x3	Hummert.com	65-6964-9	For each tray of plants
Laboratory Tape (for labeling trays)	VWR Laboratory Tape Pressure Sensitive	VWR	89097-920	Whole project
Calipers		VWR	*36934-152*	*Ideally 1 per pair of students*
Rulers	Standard cm ruler	Office supply store or Dollar Store or scientific catalog		*Ideally 1 per pair of students*

Counters	Hand Tally Counter	VWR	47005-420	*Ideally 1 per pair of students*
Scissors	Standard scissors	Office Supply Store		For instructor setting up whole project
Tweezers	Standard Tweezers	Office Supply Store		For instructor setting up whole project
Clipboards		Office Supply store		*Ideally 1 per pair of students*
Swizzles also known as drink stir swizzle stick or stir stix	Plastic colored toothpicks also work	Soodhalter Plastics or Dollar Store or Grocery Store		Per plant
Ethanol		VWR	BDH1156-1LP	For instructor setting up whole project
Seeds	Specific mutants	From unPAK– pre-screened for insert	Email project manager for specific seeds	For instructor setting up whole project
Seeds	Specific controls	From unPAK stock center – specific natural accessions and parental lines for controls	Email project manager for specific seeds	For instructor setting up whole project
Growth chamber or greenhouse	Growing area	Percival AR41L2 growth chamber ideal, but others work		For instructor setting up whole project
Sticky pads	Growing seed dispersal safety	VWR	VWR 89097-510	For classroom door, at door of growing spaces

unPAK Project Contact Information:

PI College of Charleston, Matthew Rutter rutterm@cofc.edu
- Inquiries about experimental design
- Inquiries about the ecology and evolution of Arabidopsis

CoPI Courtney Murren murrenc@cofc.edu
- Inquiries about CUREs
- Inquiries about experimental design

Co PI Allan Strand stranda@cofc.edu
- Inquiries about the unPAK database

Project Manager April Bisner bisneram@cofc.edu
- Program general information and protocols

unPAK Stock center: unPAKstockcenter@gmail.com
Visit arabidopsisunpak.org for more information

PreCURE set up and experimental design steps:

Scale of the experiment (number of mutants, treatments and replicates) will depend on growing space, number of students enrolled, number of plants instructed to measure, and whether a laboratory coordinator/TA or an instructor will be maintaining the plants. These activities have been successfully run in courses with and without TA support. For instructors in their first semester of unPAK CURE implementation and are learning to grow *Arabidopsis* for the first time, we recommend the use of only a single common environment. Instructors who are more experienced with growing *Arabidopsis* may choose to include treatments during their first semester. Depending on access to growth facilities by students outside of class hours, students may participate in plant maintenance (e.g. watering). Additionally, depending on semester length, course level, class size and other activities in a course, students have also successfully participated in sowing seeds during class time. While these components add additional experiences for students, neither are required for students to achieve learning gains and participate productively. Guidance on experimental design and potential modifications for class size and growth space are discussed below. In addition to this guidance, mentoring on experimental design modifications is available from Courtney Murren or Matt Rutter upon request (contact information above).

Technical note regarding Salk T-DNA lines and controls: The unPAK stocks are derived from Salk lines available at ABRC (arabidopsis.org). UnPAK seeds have gone through two processes: 1) seeds have been bulked in a common environment, and 2) seeds have been screened for insertions. In the process of T-DNA knockout at the Salk institute, resulting lines may have more than one insertion mutation (details of the insertion process and the number of inserts can be found here http://signal.salk.edu/tdna_FAQs.html). As part of the unPAK suite of protocols before seed distribution, unPAK instructors have screened and identified whether a Salk T-DNA line has a single insert mutation or more (see Rutter et al. 2017). Only lines with a single insert are circulated from unPAK - screened stocks for phenotyping. The Salk T-DNA lines are built from a common genetic background, Col-0 (a.k.a. Columbia) was the first *Arabidopsis* line to have its genome sequenced. The unPAK project also provides seeds for a) specific controls for this project, including the parental wild-type Col-0 line without any insertions (denoted COL 70000), and b) 10 natural accessions selected by unPAK administrators as controls (see below). The 10 natural accessions are used as phytometer controls common to all experiments, enabling researchers and students to compare standardized data across growing spaces (see database at arabidopsisunpak.org) and to

contextualize data from Col-0 to natural genetic variation in this species. Each seed shipment from the unPAK stock center will include these controls and the appropriate number of mutant lines for the available growing space.

Phytometer (plant meters) controls measure variation in growing conditions across institutions and experimental runs. Selected phytometer lines are part of the 1001 genomes project (1001genomes.org) and represent natural variation in flowering time of winter annuals. Thus, we request that instructors use the same set of 10 phytometer lines (natural accessions: CS22596; CS22617; CS22618; CS22633, CS22635; CS22636; CS22647; CS22652; CS22658; CS22659) in all unPAK experiments, such that comparison of data collected across locations is possible. Seed stocks of these 10 natural accession lines are shipped alongside any screened Salk T-DNA unPAK lines.

The number of plants grown can be modified to meet the growing space and plant maintenance support will be available to instructors. Given that the number of plants grown can be modified, we provide details of both preferred and minimum experimental size as starting point for developing appropriate replication for particular instruction needs and opportunities. Modifications for semesters which are shorter than the 14 week model (Fig 1) are possible, and would result instructors planting even earlier preCURE than the example so that both rosette and fruiting/flowering stages can be investigated during the course. Faculty with multiple lab meetings per week may also choose to engage students in routine plant maintenance during class time.

Minimum experiment size (randomized complete block design):

Lines needed: 1 parental control (Col-0), 10 unPAK selected natural accessions, 1 unPAK screened Salk-TDNA mutant.

Replication per treatment: at least 3 replicates of control, mutant line and natural accession per treatment. This extremely minimal experiment does not allow for over-replication of Col-0 (see below).

Minimum treatments: 2

Total experiment size: 72 plants.

Needed growth space: 3 trays (30cm*57cm trays), Approximately 5.5 ft^2 (0.51 m^2)

Preferred experiment size to achieve pedagogical goals:

The preferred experiment size includes more mutant lines and over-replicated parental line (Col-0). This is the typical randomized block design experiment conducted at the College of Charleston.

Lines needed: 1 parental control (Col-0), 10 unPAK selected natural accessions, 10 unPAK screened Salk-T-DNA mutants.

Replication per treatments: 6 replicates of each mutant and natural accession per treatment. Col-0 is replicated at a higher level such that each student group has 6 replicates per treatment to measure (144 plants split evenly between two treatments) and ensure sufficient parental lines are included for statistical examination of distribution of parental effects. Minimum 2 controlled growth environments are needed (e.g., 20C and 24C).

Minimum treatments: 2

Total experiment size: 384 plants

Needed growth space: 12 trays, Approximately 13.6m ft^2, 1.3m^2 per growth chamber and 2 chambers

Additional notes on designing experiment size to achieve pedagogical and individual instructor research goals in one semester:

Additional replication per mutant line and additional treatments may be required for instructors to conduct a complete and publishable research project in a single class offering. Note that this may be accomplished over multiple class offerings (in the same or subsequent semesters) using the approaches described above. For example, replication is needed in an additional pair of growth chambers or across semesters to ensure independence between growth chamber and treatment prior to publication of work from the temperature manipulations. Whether or not publication of work is expected after one semester, contribution of data to the unPAK database is a valuable educational and research outcome.

Guidance for line (seed) selection: Lines can be selected based on seeds screened, bulked and availability from the unPAK stock center at the replication levels needed for an experiment (determined by the instructor). Lines may be selected at "random" (e.g., a set of lines previously not phenotyped under any environments or under the environmental treatment selected by the instructor). Even when line selection is random and does not involve specific hypotheses, the phenotypic data collected will still contribute to unPAK's goals of phenotyping all available single insert lines. A random line selection approach leads to opportunities for novel discovery for students. An alternative to random selection is to select lines of specific interest to the instructor based on their research interests or other course goals (e.g., the study of trichome structure in a plant science course).

Suggested treatments: Example worksheets (Appendix 2) are designed for temperature treatments (20C and 24C), which has a central focus on plant response to climate change. If other treatments

are selected these segments of the handouts will require modification. Ecologically important and relevant environments that unPAK seeks to understand across mutant lines while being easy to manipulate in the context of a class project are outlined in the table below.

Environment	**Recommended Levels**	**Suggested method**	**Ecological Context**
Temperature	20C, 24C	Alteration of environmental chamber settings	Climate change and latitudinal variation in temperature
Salinity	0μM, 20μM, 40μM NaCl	Use of VWR BHD8014 in watering	Soil salinization with land use change
Competition	1 vs. 3 plants vs 5 plants per 3" (7.62) inch pot	Representation of natural density variation in pots	Competition, density
Nutrients	No, low 10%, 100% of Hoagland's Basal Salt mixture (2ml 1/wk)	Use of different levels of Hoagland's Basal Salts phytoteclab.com	Nutrient availability
Daylength	12hr, 16hr	Alteration of environmental chamber settings	Latitudinal variation

Guidance for trait measurement: Two traits are required for contribution to the unPAK database, rosette diameter and total fruit production. Rosette diameter is measured with a caliper (in mm) or ruler at or after plant bolting (see protocol page S7). Fruits are counted as 'good fruit' when siliques (a type of fruit from the mustard family with two valves and where fruit length is at least three times the width) are present, contain seeds and at least 5 mm in length.

Additional traits can be measured or counted easily in a classroom setting following specific unPAK protocols (see arabidopsisunpak.org for these additional protocols). The following traits can be included in the database: inflorescence height at maturity, fruit length, leaf number, branching, and number of aborted fruits. Measuring additional traits could also be used as means for students to learn to creatively develop replicated measurement protocols of their own. However, if they do not follow unPAK protocols, they cannot be included in the research database. Other traits which may be of interest to students but are not currently curated for the database include cauline leaf length, peduncle length, and petiole length among others.

CURE unPAK phenotype protocol

Protocols for measurement of rosette diameter and fruit number. *Note: unPAK specific protocols for additional traits can be found at arabidopsisunpak.org.*

Rosette diameter:

diameter.at.bolt: During early growth in Arabidopsis, leaves are arranged in a basal rosette. This circular rosette of leaves is typically flush to the potting mix surface. At onset of bolting, we measure rosette diameter as the maximum diameter. Using a digital caliper set to millimeters, we measure across the widest point in the rosette from the edge of the tip of one leaf to the opposite edge of the rosette. To confirm that the widest point is selected, we rotate the calipers until no additional diameter is a longer distance and retain the longest measurement.

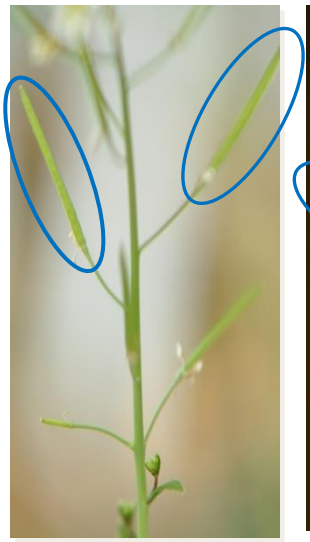
Photo credit: A. Bisner

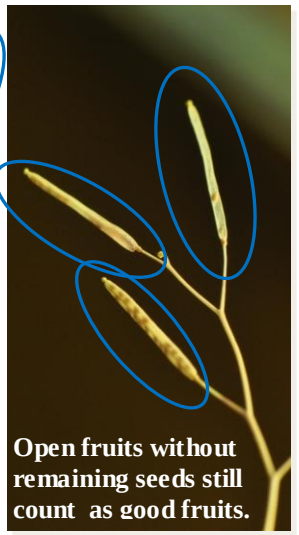
To measure rosette diameter at a consistent point during development of the plant, measurement occurs at onset of bolting (i.e. when the inflorescence has begun to elongate). We consider date of onset of bolting as when the main inflorescence has reached a height of 5 mm.

rdiam.snapshot: If rosette diameter is not measured at the onset of bolting, but the widest point measured is consistent with the protocol above, the trait is considered a "snapshot" of plant development. Date of measurement (time since germination) must also be recorded in this case.

Fruit (fitness) measurements:

fruitnum: Total number of fruit (siliques) containing seeds are counted when the main stem has stopped growing. This developmental point is when the terminal bud has senesced or developed into a fruit. When the plant has stopped growing on the main stem, fruit are counted on the whole plant including all branches (basal branches and branches on the main stem). To be counted as a fruit, the length must be at least 5 mm, with evidence of seeds within the fruit (i.e., bumps from developed seeds are visible or can be felt). If only the replum (the separation between the two valves of the silique) remains (as the valves have dehisced), it must be at least 5 mm and there must be evidence of multiple seeds produced (indentations on the replum; or seeds as in picture below).





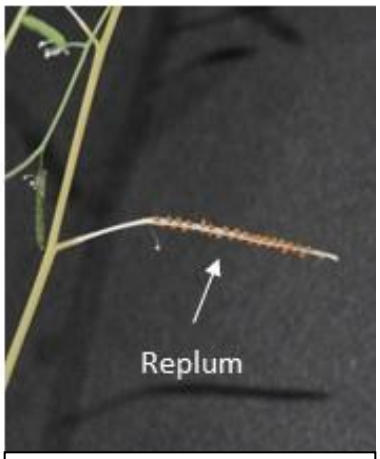


Photo credit: A. Bisner

fruitnum.snapshot: If fruit are counted at a different developmental point than described for fruitnum above, then the date of measurement must be included and the trait is considered a "snapshot".

aborted.fruits: This trait is not typically measured during unPAK CUREs. It is included here for clarity and completeness. When fruitnum are counted, not all of the reproductive output results in a fully developed silique. This includes flowers, buds or fruit that are <5 mm and are likely to contain only one or no seeds. Since the end of watering is standardized to the date of maturity/senescence of the main stem, not all flowers will develop into fruit. The total remaining flowers, buds and fruit <5 mm on all branches and the main stem are considered aborted or unfilled fruits.

Appendix 2: Student Activity 1 (SA1)

Background *Arabidopsis*: The ecology and genetics of responses to warming temperatures

Welcome to the unPAK CURE project.

Each of these projects is going to be spread across several weeks and will produce a final product, which will be the culmination of your grade on that project.

We will be building hypotheses, collecting data, testing our hypotheses using our data, implementing statistical tests on our data, and presenting our results graphically and in a written format. *The data you collect will be new to science, and will be shared with other scientists around the world through the unPAK project database!* **There are opportunities to document biological phenomena that have never been described before….**

Arabidopsis thaliana and climate change

From your online background reading (arabidopsisunpak.org/?page_id=965) , you should have some familiarity with the basics of *Arabidopsis* biology and ecology. In this project, we will work with:

- Plants that have been grown in growth chambers at two temperatures.
- Plants from a classic lab line known as "Columbia",
- And depending on your choice of group you will study EITHER:
  - ***Natural populations group:*** plants collected from two natural populations from Europe

OR

  - ***Mutants group:*** plants from the original lab line, the first to have its genome fully sequenced (goes by multiple names: Columbia Col-0 or CS70000), *and* "knockout" mutant lines that have one gene in the genome made nonfunctional with an insertion mutation (see O'Malley and Ecker 2010 for details).

You will be taking phenotypic data at two points in time (life history stages): early in development (this week) and later in development (in a few weeks). You will construct and test hypotheses about the effects of temperature and your plant groups. However, you will not know which individual plant belongs to which group until after you have collected all your data. We do this to minimize researcher bias. At the end of the project, data that pass peer mentoring and project quality control (your data!) will be shared in a public database! (www.arabidopsisunpak.org).

Today, you will work with a lab partner to accomplish the following:

- Choose which plants you want to work with (***Natural populations*** or ***Mutants***)
- Think about the effects of genetics, adaptation and history on phenotypes then develop working hypotheses for response to temperature and phenotypic plasticity
- Select phenotypes to measure with measuring tools available
- Measure phenotypes
- Enter your data in an Excel Spreadsheet (following the example)

Experiment set-up (*what's been going on behind the scenes*):

- On (add instructor specific date) all seeds for this class were set up by your instructors to be cold treated on wet filter paper for one week. This cold period synchronizes the germination of seeds.
- Following cold treatment on (instructor specific date), seeds were transferred onto potting mix and into a randomized design of 12 trays in 2 growth chambers by your instructors. Seedlings were transferred to 2.5 inch pots on standard ProMixBx potting mix. Plants were bottom watered in their trays two times per week, such that potting mix remained moist.
- Your plants have been grown at two temperatures (20°C and 24°C) in Percival growth chambers. Plants were grown in a growth chamber with 16 hours of light per 24 hour period under those temperatures.
- Half of your plants have been grown at 20°C. *This temperature corresponds to a warm spring day within the normal range of native Arabidopsis populations*.
- The other half have been grown at 24°C. *This temperature reflects what average spring temperatures might look like in 2100 under the predictions of the Intergovernmental Panel on Climate Change (IPCC).*
- Plants are grown in trays (green trays) that hold 32 plants each. Plants with different genotypes are being grown in a randomized block design.
- Each group will have two different accessions (Natural Population A & Natural Population B – **or** – Columbia 70000 parental line & Salk-TDNA Mutant). You will measure approximately 6-9 replicate plants per accession per treatment (between 24 and 35 plants total) based on germination and early survivorship.

  Note: We use the words *line, accession, genotypes,* and *maternal line* as interchangeable descriptions of the seed material with the same genetic heritage.

Work with your lab partner and come up with 5 possible effects that warmer temperatures might have on plant growth and development.

Natural populations group: Some of you will work with plants that have been collected from natural habitats throughout Europe. If you are in this group, proceed by answering the following questions.

What kind of habitat characteristics might vary between the ecological locations you would find *Arabidopsis* in?

Identify 3 plant characteristics that might differ because of adaptations to local environmental conditions.

Why might different populations respond to our temperature increase in different ways? Are there any reasons that there might not be a response to temperature?

Mutants group: Other groups will work with two mutant lines, each of which has one of the 27000 Arabidopsis genes disabled. The disabled gene has been "knocked out" by insertion of a "nonsense" sequence by an infectious bacterium, *Agrobacterium tumefaciens*. This set of lines are called Salk T-DNA lines. Subsequent molecular biology research by the folks at Salk Institute determined just where that insertion occurred. See the article: O'Malley, RC., and JR. Ecker. 2010. Linking genotype to phenotype using the Arabidopsis unimutant collection. The Plant Journal 61: 928-940.

Any phenotypic effects that you discover may be new to science! However, in most types of experiments it isn't possible to discern the effects of a single mutation. If you are in this group, proceed by answering the following questions.

Why might it be hard to detect the consequences of knocking out a gene?

What kind of phenotypes do you think would be interesting to examine to determine the effects of a mutation?

Now that you have thought a bit about the types of phenotypes that could be associated with temperature, natural populations with a known origin or mutants with single knocked-out genes, it's time to get some data.

In Class To Do List:

1) Discuss with your partner and sign up on the board to be in either the mutant or natural population group. Both group's data will lead to very interesting findings and reports.
2) Get a list of plants to be studied from your instructor (you'll have information about the position in the experimental design, but not genotype or treatment ID).
3) Find your plants based on their plant number. (They will be in the 12 trays of plants in pots laid out on the tables). Take your set of pots back to your table.
4) Work through the background handout.
5) Discuss with your partner the potential traits to measure.
   a) Each pair of students will choose a quantitative phenotypic trait of interest and confirm the selection with your instructor.
   b) Possible all-class phenotype: In a few minutes, we will discuss the all-class phenotype and vote as a class on the phenotype we want to measure. Then we will describe how we measure this trait in our metadata.
6) Required trait: Measure the maximum rosette diameter on each plant and record this measurement. Your instructor will describe specifically how to measure and use of calipers.
   (a) Why might we be interested in the whole class measuring the same trait?
7) Once consensus has been made for the class trait, *use the paper data sheet provided.* This includes the plant tag information and space for writing your phenotypes and space on the flip side for your meta-data. Each student will have between 24 and 36 plants to measure.
   a) Take notes and pictures along the way of how you measured the class trait, group trait, and rosette diameter.
8) Return your plants to their original positions in the trays using the tag number information.
9) Wipe down your tables as needed.
10) Record your data in an excel file using your own laptop. Make sure both you and your lab partner have an e-copy of your excel file before you leave discussion today. *Include meta-data materials in a separate tab of the Excel file with the quantitative data*. See details below.

Data Management

What do I do with my data and what the heck are "meta-data"?

One of the most important skills you need to have as a biologist is to manage your data in a way that is organized and can be passed on to others who will work with the data. Your Arabidopsis data will be shared with others in the class and with the scientific community, so it needs to be clear for interpretation!

Here are some rules of thumb:

1. Use a spreadsheet program to house your data. This could be an Excel or a Google spreadsheet. We will eventually be doing some analysis in Excel, so you will need your data in a format that can be copy-and-pasted into Excel regardless.

2. Follow the header format from your handout to enter your data into a spreadsheet. This format will make the next steps run more smoothly. It is good practice to write in pencil on paper when working with wet and live materials. When you have completed your data collection, transcribe your data into an electronic spreadsheet.

3. Each plant has a unique identifier from its original position -- the plant number.

4. Each individual plant's information will be on one row, with the first column of that row specifying its physical location in a growth chamber.

5. The next 2 columns should indicate "Temperature" and "Genotype". Leave the cells in these two columns blank for the moment. When you've collected all your data from the plants in a few weeks, I will reveal both the treatments and genotype information for analysis. This way you are collecting your data while being "blind" to the effects that you are measuring.

6. The next columns will reflect the phenotypes you measure. The first should be Rosette_Size_mm. Always put your units of measurement in the column title cell, but don't put the units in the data cells! (e.g. record 10, 20, 18 …) Use the meta-data form on the back of the sheet to record the units information.

7. As you measure your data, fill in the cells in the appropriate row for the plant you are measuring and the column corresponding to the data.

8. Save frequently, and make sure you pass a copy to yourself and your lab partner! Name your e-file in a way that will make sense in a few weeks. e.g. last names of lab partners, project and date: Murren_Rutter_Arabidopsis_First_measurement_26Jan2016.xls

Sample paper datasheet:

Plant_Number	Temperature	Genotype	Rosette_diam_mm	Leaf_number	Hairy
44			24	8	YES
12			31	12	NO
256			15	22	YES

Metadata are data about the data. In our case, these descriptive data includes (but not limited to) when the plants were measured, details of how additional traits were measured, and who conducted the measurements.

Sample meta-data:

Measurement date: 1/26/2016

Partners: Courtney Murren & Matt Rutter (murrenc@cofc.edu; rutterm@cofc.edu)

"Rosette size" is measured in mm across the widest part of the leaves. "Leaf number" includes a count of all leaves but does not include cotyledons. "Hairy" is categorical data on whether the topmost leaf (most recently developed) looked fuzzy.

Photographs: rosette26Jan2016.jpg was taken with Matt's phone and indicated how we measured rosette diameter.

Note meta-data can be typed into a second tab of an Excel sheet that you rename "meta-data" OR can be saved as a Word document file (easy if you want to paste photos directly in the metadata). However, make sure to put a note in the Excel file where to find the meta-data (e.g. second sheet) and name the file the same as the xls file with the word meta-data added (e.g. Murren_Rutter_Arabidopsis_First_measurement_metadata_27Jan2015.doc)

Homework assignment for this project for next class:

1. ***Make sure your data and metadata are in order and that both you and your lab partner have a complete copy in an electronic spreadsheet. (e.g. email each other)***

Appendix 3: Datasheet 1 (D1).

Datasheet 1.

Plant_Number	Tray	Treatment	Genotype	Rosette_diam_mm	Class_trait	Your_trait

Datasheet 1. Metadata

Names:

Dates of measurement:

Measurements and units:

How were measurements made:

Appendix 4: Student Activity 2 (SA2)

Instructor note: At CofC, library faculty will visit the classroom to introduce students to search techniques in Web of Science, Science Direct, JSTOR and google scholar for finding scholarly articles. We highly recommend collaborating with a library faculty member on this section of the curriculum if it is available.
For those without this option, many library search tools provide online tutorials and tips for performing informative searches on their platform. For example, Science Direct provides trainings through their online support center and via FAQs, and JSTOR offers a "How to Use JSTOR (for students)" page (https://guides.jstor.org/how-to-use-jstor). These resources can be used to help students explore the scholarly literature and find what they are looking for.

Links at arabidopsisunpak.org aid in instruction for conducting searches at TAIR and T-DNA express by following the links under the Education/CURE Resources tab http://arabidopsisunpak.org/?page_id=965.

Literature and Database Background: Natural Populations Groups

Last week, you took data on *Arabidopsis thaliana* phenotypes looking at plants from two genetic backgrounds (two natural populations) grown in two temperatures (20 & 24C).

Your assignments today are to:

1. **Learn about the locations that your natural populations came from**. Work on this section in collaboration with your lab partner.
2. **Locate 5 primary literature sources** relevant to your populations or the study and read the abstracts associated with these papers. This will result in a total of 10 citations (5 per partner —it will be helpful for the final assignment if you turn up different papers!)
    a. Type up these references using the bibliography format below (see last page of the assignment). It is *not acceptable* to copy and paste from Web of Science *without* reformatting. **Bring this list for next week including a copy for your instructor and a copy for your lab partner.**

About your populations

Your instructor will give you information on the accession numbers of your two populations (given as CS#####). From there you'll find out if there are additional codes for your populations. The first step is to figure out where those populations are from. In this handout population, line, accession, and ecotype are all synonyms that describe a set of seeds that was collected from a particular natural population and has been cataloged in the Arabidopsis database called The Arabidopsis Information Resource (TAIR). TAIR is the current central clearinghouse for much Arabidopsis information.

1) To do this you are going to use TAIR.
1. Go to the TAIR website: www.arabidopsis.org
2. Your first step is to find your populations, also known in this database as ecotypes.

3. Type the name of your population into the search bar at the top. Make sure you switch your search to the "germplasm" or "seed stock" database, as the default is "gene".
4. You should find your ecotype on a list. Clicking on the link will take you to a page with various data about the population, including the city / region and country where it was originally collected.
5. An example using TAIR to find information about your lines can be found here: http://arabidopsisunpak.org/?page_id=220

Where are your populations from?
Population 1:
Population 2:

2) You might want to look around this page a bit, as it may contain some additional information about phenotypes or even publications associated with your population. Write any relevant information here.

3) Now it is time to use Google. First, figure out the latitude and longitude of your locations. Talk to your instructor to find out if your populations have "unusual" names.

Population 1:

Population 2:

4) You may also want to use Google-maps or other tools such as https://www.nhc.noaa.gov/gccalc.shtml to estimate distance between populations.

5) Can you find anything out about the climate at your sites of origin? One site to look at is wunderground.com

- Type in the name of the location that your population is from.
- Scroll down the page to the Almanac section.
- Click on the link for September Calendar View.

Most of the time, Arabidopsis grows primarily in the spring between the end of February and the end of April. You will need to change the month in the calendar view to each of February, March and April. What are the average temperatures like at your sites during those time periods?

Population 1:

Population 2:

What is precipitation like at your sites at those times?

Population 1:

Population 2:

Now that you know a bit about where your populations are from, do you think they would differ in their response to the temperature treatment? Why or why not?

6) Your next step is to look for some literature relevant to the study. You may wish to use the literature databases and a search engine (e.g. Web of Science, Google-Scholar etc.) to find out if anyone has studied your populations before. What did they find out? What do we know about Arabidopsis growth, size or temperature? Has anyone else previously studied your phenotypes? What is known about the phenotype rosette diameter, your group trait or the class trait?

Find 5 primary literature references that are relevant to your study. Email either the abstracts or papers to yourself, and record the citation information – so that you can clean up the information into proper format for a bibliography. It would be quite useful to find a general paper on *Arabidopsis* ecology, several about plasticity or performance in response to temperature, and if possible, a paper on population variation for your group or class phenotype.

7) Feel free to record any interesting background information about your populations that might be useful in a lab report against which you may want to contrast your results.

What is due next week?
1) A list of 5 journal sources (different for each person in pair) using the reference format below.

Example reference format. Copy the formatting and punctuation exactly!

Duncan AB, and Little TJ. 2007. Parasite-driven genetic change in a natural *Daphnia* population. Evolution 61: 796-803.

Conner JK, Davis R, and Rush S. 1995. The effect of wild radish floral morphology on pollination efficiency by four taxa of pollinators. Oecologia 105: 509-516.

Literature and Database Background: Mutant Groups

Last week, you took data on *Arabidopsis thaliana* phenotypes looking at plants from two genetic lines (Columbia and a knockout mutant) grown at two temperatures (20° and 24°C).

Your assignments today are to:

1. **Learn about the genethat is knocked out in your mutant lines**. Work on this section in collaboration with your lab partner.
2. **Learn about the parental strain Col-0.** Work on this section in collaboration with your lab partner.
3. **Locate 5 primary literature sources** relevant to your mutants or the study and read the abstracts associated with these papers. This will result in a total of 10 citations (5 per partner —it will be helpful for the final assignment if you turn up different papers!
    a. Type up these references using the bibliography format below (see last page of the assignment). It is *not acceptable* to copy and paste from Web of Science *without* reformatting. **Bring this list for next week including a copy for your instructor and a copy for your lab partner.**

Your mutants

Your instructor will give you information on the Salk mutant line that you are investigating. You are also studying Col-0 (aka Columbia) – the original line to have its genome fully sequenced and the parent of the mutant line. The first step is to figure out what locus/loci are affected by those mutations and what we may (or may not) know about the line.

1) To do this you are going to use The Arabidopsis Information Resource (TAIR). TAIR is the current central clearinghouse for all Arabidopsis information.

1. Go to the TAIR website: www.arabidopsis.org
2. Your first step is to find your mutant lines...
3. Type the name of your line exactly (as in "SALK_138583C") into the search bar at the top. **VERY IMPORTANT:** Make sure you use the pull-down menu to switch to the ***germplasm*** database (the default is gene).
4. You should find your line listed on a search results page. You want to click on the **locus** that has been affected by the mutation by clicking under the gene ID (**this will start with the letters AT**) under Polymorphism / Locus.
5. More details about using TAIR to find information about your lines can be found here: http://arabidopsisunpak.org/?page_id=230

2) What locus/loci are affected by your mutation? Give the AT # and any other names you may find for the gene.

Mutant:

3) Give some basics about this line from the Description field.

Mutant:

4) What kinds of plant tissues is the gene normally expressed in? What kinds of functions does it have? Some of this is summarized in the Annotations section. One possible option is no function is currently known or described.

Mutant:

5) You might want to look around this page a bit, as it contains information about any previously studied phenotypes or publications which also studied this line. (hint: these may be helpful citations for your paper). The site **araport.org** can also be useful for searching for your gene (starts with AT), as it is another collection of information about Arabidopsis genetics.

6) Another site to look at is: signal.salk.edu; click on the T-DNA express link at the top of the page. More information about using this webpage can be found at: http://arabidopsisunpak.org/?page_id=230 (note—do not include the "C" at the end of the Salk line as shown in the image). The part of the gene where the insertion occurs is usually indicated by **LOCN:**
What part of the gene is the insertion located in? Common options include exon, intron, UTR (untranscribed region), and promoter.

Mutant:

Why might the insertion location matter?

7) Another site to look at is: http://arabidopsis.info/CollectionInfo?id=94 You will find background on the "Columbia" line. You may also choose to do an ecotype search on "Columbia" or Col-0. The control line we are using is CS70000.

8) Your next step is to look for some literature relevant to the study. You may wish to use the literature databases and a search engine to find out if anyone has studied your locus before. What did they find out? What do we know about Arabidopsis growth, size, and temperature? Or your group or class phenotypes? Or your Salk T-DNA line?

Find 5 primary literature references that are relevant to your study. Email either the abstracts with full citation information (see below) or PDFs of the papers to yourself. Additional references that will be helpful are general *Arabidopsis thaliana* ecology, a reference describing the Salk T-DNA mutants, Arabidopsis plasticity of rosette diameter or group or class trait, and Arabidopsis response to temperature. REVIEW papers about Arabidopsis ecology or about T-DNA insertion phenotypes that might be especially helpful.

9) Feel free to record any interesting background information about your loci that might be useful in a lab report once you have your results.

What is due next week?
1) A list of 5 sources using the reference format below (5 that are different from your partner).

Example reference format. Copy the formatting and punctuation exactly!

Duncan AB, and Little TJ. 2007. Parasite-driven genetic change in a natural *Daphnia* population. Evolution 61: 796-803.

Conner JK, Davis R, and Rush S. 1995. The effect of wild radish floral morphology on pollination efficiency by four taxa of pollinators. Oecologia 105: 509-516.

Appendix 5: Datasheet 2 (D2)

Plant_Number	Tray	Temperature_ Treatment	Genotype	Fruit_number	Class_trait	Your_trait

D2 Datasheet 2. Metadata

Student names & contact info:

Dates of measurement:

Measurements and units:

How were the measurements made: (photo filenames and location of photographs)

Filename of saved data

Appendix 6: Student Activity 3 (SA3)

Instructor notes: At CofC students have a pre- or co-requisite of introduction to statistics in the math department to enroll in the course which uses this module. Students have conducted at least one statistical test prior to the activity described here. Additionally, throughout the lecture portion of the course and in discussion of primary literature, evaluation of graphs and interpretation of statistical outcomes are emphasized. We use Excel for data management and implementation of statistics as all students have access to this tool on our campus. At Arabidopsisunpak.org there is a video tutorial for use of general instructions. This is also available as a hand-out on the website (education / CURE resources / working with excel tab is at http://arabidopsisunpak.org/?page_id=285).

For the activity below, students first watch the video tutorial (YouTube at link above) and then complete a set of exercises on a classic dataset (Bumpus), which students then submit. Feedback on this assignment is provided before class. This allows the students to practice the skills necessary for their project. When students arrive in class, they implement these skills with their own dataset. Depending on the level of the course, learning objectives, and computer lab facilities, the statistical tests described below could also be implemented in other statistical environments such as R or SAS.

Hypotheses, graphing and statistics: *Arabidopsis thaliana*

You have watched the Excel YouTube video and completed and submitted the graphing and statistics homework on a classic ecology dataset (working with Excel homework http://arabidopsisunpak.org/?page_id=965). Your goal today is to graph and calculate statistics on your own Arabidopsis data that will contribute to discussing your biological argument in your lab report. First, recall which traits you had measured and what you had learned about your plant lines (mutant or natural populations). Refine your earlier overarching hypothesis about differences or relationships among your traits or lines from what you have learned in your further primary literature readings.

*Your instructor will give you information about which plants were from which genotype, and which plants were from which temperature treatment. Add this data to your data file.

* What traits did you measure? Were these traits continuous or categorical? What else do you know about your plants? Are these factors continuous or categorical?

* Your lines:

* General hypothesis: (Example: Temperature affects plant growth, but the effects differ between genetic lines.)

List 2-3 refined hypotheses based on the ones that you developed in SA1. Draw on your newly obtained knowledge from scholarly sources, background information from databases (developed in handout SA2), and additional information we have examined in the lecture component of this course in relation to plant biology, ecology and genetics.

Identify 4 specific tests you will perform. **PLEASE PERFORM AT LEAST ONE CORRELATION AND AT LEAST ONE T-TEST. It is hard to imagine that you will not want to contrast phenotypes of your different types of plants (populations, or mutant vs. control) and to contrast plants under different treatments.**

Outline a way in which your 4 tests will allow you to tell a coherent story about the biology of your plants.

Now generate the appropriate statistical test and a figure to go with each specific hypothesis.
***Note: You cannot do a correlation analysis using "temperature treatment" or "genotype" as a variable – these are both categorical.**

Provide a one sentence description of your results, including the statistics. Please show the sentences to your instructor before finishing today.

Test 1:

Test 2:

Test 3:

Test 4:

YOU MUST TURN IN YOUR DATA TO GET A GRADE ON YOUR LAB REPORT – see instructions

Appendix 7: Written Assignment 1 (WA1).

unPAK CURE WRITTEN ASSIGNMENT

Each student will write/type (in their own words), and hand in their own report. Graphs can be the same between students from the same group.

This written assignment will be a lab report. This is an opportunity to gain some expertise in the style of writing, which sections information goes into, and to obtain feedback on writing. We are happy to meet with you if you have questions about your writing.

For THIS assignment you will organize your paper into the typical sections of a journal article –see notes below on specifics on what to include in particular for this assignment.

1) **Introduction:** In this section, you should give background on *Arabidopsis thaliana,* and your mutants or your natural populations. The first paragraph is a general overview of the project (including introduction of the temperature treatment) written as if it was the first paragraph of a journal article. In the second and third paragraphs you should introduce your two mutants or natural populations. **Use the literature! Citations for general concepts and Arabidopsis info are needed.** Follow the parenthetical style of citations. The fourth paragraph should introduce the specific biological questions or predictions for your study.

THE INTRODUCTION SHOULD MAKE THE CASE FOR YOUR QUESTION

2) **Materials and Methods:** A short paragraph describing the origins of the plant material, your measurements, how you made the measurements and your use of Excel as your main software to evaluate your questions using graphical and statistical methods. **The Materials and Methods section should be written in sufficient detail such that someone else could replicate all of the experiments described.**

3) **Results:** A few paragraphs describing your results in words. You will specifically refer to and describe the patterns in the graphs (labeled as Figures 1-4), report your test statistics including df and p-values. Each figure should have a figure legend, be named figure 1-4, and the figure legend should be able to stand alone. **AVOID GRIDLINES IN YOUR FIGURES.** Take a peek back at papers we have discussed as a class for style. Each student will have four figures and at least one scatterplot and one bar graph. Note that **YOU WILL BE GRADED ON DESCRIBING YOUR STATISTICS CORRECTLY.**

4) **Discussion:** Two paragraphs describing the conclusions that you draw based on your statistical analyses and interpretations of your graphs. Do these results offer support for (or refute) your biological hypotheses/predictions? How so? What are the broader interpretations of your findings? Again, use the primary literature to provide a framework for your discussion. **DO NOT JUST SUMMARIZE YOUR RESULTS, INTERPRET THEM.**

5) **Abstract:** This should be the first section the reader sees, but it should be the last section written. The abstract must include: 1) the research conducted, including the rationale, 2) methods, 3) key results, and 4) the main conclusion, including key points of discussion. Approximately 80-150 words for this section.

FORMATTING GUIDELINES:

All scientific journals have a set of guidelines that must be followed for a paper to be considered for publication. Below find the "Instructions for Authors".

- The written portion of your report should be double-spaced and use 12-point font.
- *Title, name & affiliation of investigator, and name of partner* should be placed at the top of the first page.
- The sections (Abstract, Introduction, etc.) should be clearly labeled.
- *Figures or tables* may be placed in line with the text or following the text on separate sheets. Refer to all diagrams, graphs, and photographs as 'Figures'.
- *Figure legends should be included.* Key information describing each figure should be in the first sentence of the legend because this is the text that will be immediately visible. The rest of the legend should be a self-contained full explanation of the figure with all abbreviations defined.
- *Literature cited* should immediately follow the text. Give the full reference for all sources in the text, and make sure all sources referenced are in the text. Use the format shown below.

GENERAL TIPS (aka – how to avoid common mistakes):

1) **Do not use the word "prove"** (proven etc.). In the data analyses we are doing here, we are *finding support* in favor of a particular prediction, or *failing to find support*.

2) **Data are plural.** For example: *The data were analyzed using Excel*. Data set is singular if you want to mix your language usage a bit.

3) **Format presentation of statistical results in a sentence format**. For example: *There was a significant positive relationship between shell length and shell width ($r = 0.92$, $df = 43$, $p<0.001$).*

4) The **word species** is both plural AND singular.

5) **Refer to your figures in your text.** For example: *We found a positive relationship between herbivory and leaf size (Figure 2).*

6) If the species is in very common usage then the common name will suffice, although the scientific name should still be given at first mention (e.g., soybean (*Glycine max*). For

subsequent uses, abbreviate genera to their initial letters (e.g., *G. max*) except where this could result in confusion between species or if the species name is the first word of the sentence.

7) **Make sure scientific names are in *italics* with only the genus name capitalized.** For example: *Moehringia macrophylla*. Be sure to check the literature cited as well. Common names should NOT be capitalized. **YOU WILL LOSE POINTS FOR DOING THIS INCORRECTLY.**

8) Page numbers, and author's name (i.e. your name) should be in either a header or footer on each page of the document.

CITATION FORMAT – YOU WILL BE GRADED ON THIS.

***Citations in the text*:** Each time an article is cited, the authors' last names and the year should be placed in parentheses. If the authors' name(s) are the subject of the sentence, only the year goes in parentheses. When there are three or more authors, all but the first author are abbreviated as "et al." (meaning "and others").

Here is an example involving one single-authored paper, one co-authored paper, and one with more than two authors:

> Some biologists, like Rutter et al. (2012), have said one thing, while others (Strand and Sotka 2003, Callahan et al., 2004) have said another.

***References in the literature cited section*:** Use a typical format as shown below.

- Journal article:
  - Duncan AB, and Little TJ. 2007. Parasite-driven genetic change in a natural *Daphnia* population. Evolution 61: 796-803.
  - Conner JK, Davis R, and Rush S. 1995. The effect of wild radish floral morphology on pollination efficiency by four taxa of pollinators. Oecologia 105: 509-516.
- Book Chapter:
  - Sancho G and de Buron I. 2002. Citation styles in European Journals. In *International Guidelines for Citation*, ed. Southgate A. CofC Press: Charleston.
- Book:
  - Wiseman R. 1999. Stickler for proper citation: a memoir. CofC Press: Charleston.
- Website:
  - *General format (include whatever information is available):* Editor, author, or compiler name. Name of Site. Url (web address). Date of access.
  - Felluga, Dino. Guide to Literary and Critical Theory. http://literarytheory.com/felluga.html. 10 May 2006.

AVOIDING PLAGIARISM

The bottom line: *Plagiarism is grounds for receiving an F on the assignment, or an F in the class. Consider this your warning!* You must write in your own words using your own sentence construction based on your own understanding. Do not rearrange parts of sentences written by other authors (including your partner) or simply change, remove, or add words--these all constitute plagiarism. Avoid quoting authors verbatim, even if you put the text in quote marks and cite them (unless you need to comment on their exact words and the meaning would be lost unless you quoted them exactly). Cases of suspected plagiarism will be handled according to the College of Charleston honor code. CofC plagiarism violations range from Class 1 (majority of submitted assignment is intentionally plagiarized) to Class 3 (unintentional plagiarism). Prior to the assignment being due, office hours are a great time to ask questions and gain assistance in writing.

Next week in class, your paper will be peer reviewed using a checklist and written comments. You will revise your paper, respond to reviewers' comments, and hand in the original, checklist, and revised paper.

Grading Rubric (example):

Paper section	**No credit -** includes plagiarized or entirely missing elements	**Partial credit -** some or all elements are missing	**Full credit** - all described elements are included
Title and address	Title and address are missing.	Either title is not descriptive or address is missing.	Title is descriptive, and authors names and scholarly addresses are included.
	0 points	1 point	2 points
Abstract	Abstract is missing	Some elements of abstract are missing or vague	Abstract includes a rationale of research conducted, general statement of methods, key findings, and a main conclusion
	0 points	2 points per intro, method, results, and discussion elements	8 points
Introduction	Introduction is missing citations or background information on the study	Introduction has incomplete overview, background on the study organism, background on specific lines or treatments, or missing specific questions or hypotheses	Introduction includes 1) a general overview of the problem, 2) background information of Arabidopsis and response to specific environmental treatment, 3) background information on the mutant or natural lines studied, 4) citation of the literature, 5) a final paragraph introducing the questions or hypotheses being tested (i.e., their hypothesis of how the environmental treatments will affect their specific genotypes)
	0 points	1-24 points. 5 points earned for each of the elements.	25 points
Methods	Methods is missing	Methods section is missing a subset of elements of experimental design, how measurements were made, or tools	Methods includes 1) information on plant lines used, 2) size of pots, 3) replication, 4) how treatments were implemented, 5) measurements including

		used for statistics and graphing	how and when measurements were made, and 6) methods (or tools used) for statistics and graphing. Photos of plants are a bonus.
	0 points	4 points earned for describing experimental design, lines used, measurements, and graphing and statistics procedures	20 points for all elements included
Results text	Results section is missing or does not follow handout on how to communicate statistical results.	Results are incomplete, for example, there is missing text associated with one of the figures.	Results include general findings and patterns and refers to figures in text (in order).
	0 points	2 points for text associated with each figure, 1 point for overall structure	9 points
Results Figures (Figures 1-4)	Figure is/are missing or nearly incomplete.	Figure include Excel generated headers, which should not be present, but is missing axis labels or the figure legend.	All required figures (4) are included, with at least one bar and one scatterplot in the set. All figures must have complete legends that note all relevant components. Bar graphs must include appropriate error bars, and bars must be labelled with categories. Scatterplots must not include the best fit line of excel. Axes must be properly labeled.
		5 points earned per figure, 1 point each for axes, error bars, denoting groups with legend, text legend with all components	20 points (5 per figure)
Results statistics	Statistical test results missing for all figures.	Figure(s) are missing accompanying statistical test(s). A common mistake is including R^2 generated by the best fit line in	The reporting of statistical findings includes the test statistic, degrees of freedom and *p* value. There is evidence of implementation of the

		Excel instead of the correlation coefficient to accompany scatterplot.	appropriate test for the data type. The degrees of freedom match the experimental design. Statistics are used to support descriptions of patterns in phenotypic responses of their genotypes.
		3 points for each test, with all components reported and concise sentence	12 points (3 points per test)
Discussion	Discussion is missing	The discussion is incomplete and does not include primary literature or answer the biological questions. The discussion simply summarizes results without interpretation.	Includes at least two paragraphs describing the conclusions based on interpretation of graphs and statistical analyses. Connects data to questions/hypotheses in the introduction. Clearly describes how questions are answered. Results are interpreted and connections between genotype, phenotype, and environment are clearly made. Primary literature is included in the interpretation of results and includes broader implications of the findings.
	0 points	1-9 points earned for completeness including each hypothesis described and linked to primary literature	10 points
References	References are missing.	References include inappropriate non-scholarly articles (e.g. Wikipedia). Literature cited section contains numbered citations or citations not in alphabetical order.	All referenced scholarly works follow handout for format of in-text references and of the literature cited. References in literature cited are journal articles and listed in alphabetical order.

	0 points	1 point lost if not in appropriate format, non-scholarly source, or not referenced in text.	5 points
Grammar etc.	The document includes the use of bulleted lists or otherwise non-standard writing styles.	There are numerous spelling or grammar errors. Scientific names do not follow the appropriate format.	Scientific names have the genus capitalized and both genus and species in italics. Scientific name format is also included in literature cited. Use of language throughout the document is clear with paragraphs that communicate ideas easily and succinctly.
	0 points	2 points each for grammar, format of scientific names, in text reference format, clear logic in communication, spelling and other technical aspect of writing	11 points
Implementation of reviewers' comments and response to reviewers' comments	Final assignment does not include revisions in response to reviewers' comments and lacks a letter response to reviewers' comments.	The letter in response to reviewers' comments is missing or modification of the document only partially addresses concerns raised by peer reviewers.	Final document incorporates changed in response to peer review comments and includes letter of response to reviewers that succinctly describes how reviewers' comments were addressed.
	0 points	4 points in revision of the document in response to each of the reviewers	8 points
Total points			130

Appendix 8: Student Activity 4 (SA4).

Peer review

Author of Manuscript_____

Reviewer #1_____, Reviewer #2_____

Section_____ Date_____

Arabidopsis project Peer Review Checklist and paragraph–

- Author: To be turned in with final paper including a short letter (1-2 paragraph) to the editors (your instructor/s) specifying how these changes were incorporated into the final version.
- Reviewer: Include Y or N on each line below. Add brief descriptors when needed. Sometimes the correct answer is N. Feel free to indicate spelling errors or other topics. This is a minimum list - getting a Y from reviewers on each of these does not indicate that the sections have met the professor's rubric - it is a preliminary check. Help your peers - let them know if you think you can help them get a better paper!

___ 1. Is there a descriptive title? (e.g., "Arabidopsis project" is not descriptive)

___ 2. Is the author's name and affiliation included?

Introduction:

___3. Does the introduction start with a broad topic in biology (such as mutation, climate change, adaptation, or phenotypic plasticity)?

___4. Does the introduction end with specific questions or hypotheses?

___5. Does the introduction discuss environmental variation (e.g., regarding temperature)?

___6. Does the introduction discuss genetic variation (either between populations or between a parental and mutant line)?

___7. Does the author cite primary literature?

___8. Are the in-text citations formatted properly (Smith 2014) or (Rutter and Strand 2014) or (Murren et al. 2016)?

___9. Are there sentences that seem to need additional citations? (Mark them directly on the manuscript>> REF)

___10. Is the introduction telling an interesting and coherent story?

Methods:

___11. Does the author describe their genetic lines or natural populations (e.g., where they are from or which gene the mutation is in)?

___12. Is the basic biology of Arabidopsis described (e.g. relevant stages of the life cycle)?

___13. Is basic biology cited using the primary literature and are online databases or other general references included?

___14. Are the temperature treatments described?

___15. Are the rosette measurements described?

___16. Does the author use the first person (I or we)?

Correct answer: active first person is preferred in evolutionary ecology writing.

___17. Are units of measurement described?

___18. Are the second set of measurements described, and do they include units?

___19. Are the number of plants measured included?

___20. If figures (photographs) are included on how traits were measured, do these have figure legends and are they referred to in order in the text?

___21. Are there 1 or 2 sentences describing what statistical tests were performed including what program was used to implement these tests?

___22. In general, do the methods seem sufficient for you to replicate this experiment (seeds sown, randomized, maintained, measured, etc.)?

Results:

___23. Are there text paragraphs describing the results?

___24. Does each figure have a figure legend below the figure?

___25. Does the figure legend include a description of all pattern or color choices (if multiple were used)?

___26. Are there at least 4 figures and at least one scatterplot and one bar graph?

___27. Do the figures have gridlines (final versions should not have gridlines)?

___28. Do the figures have an Excel generated title (there should be no title, only a descriptive figure legend)?

___29. Is the font size of the figures readable (10 – 12-point font)?

___30. Are the axes labelled with:

___ units (as appropriate),

___ a descriptor (categories), and

___ values (quantitative)?

___31. Do quantitative axes start at zero?

___32. For bar graphs: Do the bar graphs have error bars?

___33. Do figure legends state if error bars are based on standard deviation?

___34. Does the scatterplot have a trend line?

___35. Does the scatterplot have an equation or R^2 generated by Excel (It should not have either, this is not the correlation coefficient)?

___36. Does the results text refer to each figure?

___Are the figures referred to in order?

___37. Is each statistical result presented using a parenthetical statement that includes the test statistic, degrees of freedom and *p* value?

___38. Are the statistical results described in terms of the support for a biological pattern?

___i.e. Does the author describe null hypotheses of statistics? If so, these should be recast.

___39. Are there at least 4 statistical tests?

___At least one t-test corresponding to a bar graph?

___At least one correlation corresponding to a scatterplot?

Discussion: → not yet completed for draft

Literature cited (if included):

___40. Does the literature cited format exactly follow the handout? *NOTE: EVEN online journals follow this format. DOI can replace page numbers for e-journals.*

Conner JK, Davis R, and Rush S. 1995. The effect of wild radish floral morphology on pollination efficiency by four taxa of pollinators. Oecologia 105: 509-516.

Duncan AB, and Little TJ. 2007. Parasite-driven genetic change in a natural *Daphnia* population. Evolution 61: 796-803.

___41. Are there at least four references to the primary literature?

___42. Is the Literature Cited organized alphabetically by first author's last name?

Overall:

___43. Does the author include additional graphs/stats to tell the story?

___44. Is *Arabidopsis thaliana* in italics? With only the genus capitalized? Or after first use as *A. thaliana*?

___45. General language use

___Is use of the word "data" plural?

___Does the author use the word "prove" (they shouldn't)?

___Is the word "species' used as both singular and plural?

___Is "specie" used (it shouldn't be, it is not a biology word)?

___46. Each reviewer please include a brief paragraph of comments for your colleague to help improve the grade on their paper and their writing as a scientist. Please use professional constructive language.

Appendix 9: Student Activity 5 (SA5).

Student Data Curation EXAMPLE

Instructor note: The data file template is available at http://arabidopsisunpak.org/?page_id=965 and updated regularly.

Each *group* must create a datafile in Excel that contains all of your data in A VERY SPECIFIC format for the unPAK database manager to upload to Arabidopsisunpak.org where your data will be authored and shared with the community. After you submit your data file, it will be peer reviewed and you will have the opportunity to make revisions. Upon submission, you may also be required to do further revisions per your instructor. After acceptance of your data file, both partners will receive points.

How do you make a properly formatted datafile?

1. Download the template data file (S16BIO211CURE_TEMPLATE.xls) from Arabidopsisunpak.org
2. Fill in the file with your data
    a. BEFORE YOU ENTER ANY DATA YOU MUST FILL IN YOUR METADATA SHEET – **ALL CELLS HIGHLIGHTED IN YELLOW ARE REQUIRED**. The first worksheet in the Excel file is a metadata sheet where you must fill in additional information about your measurement.
        i. The top two tables specify the two traits to be submitted. One of these was measured at the first timepoint (rosette diameter) and the other at the second timepoint (fruit number). Only include those two traits.

        The chart below includes codes of BIO 211 students that will be included this semester. Note that you may have measured in cm; if so, convert measurements to mm.

        ii. Examples:

TRAIT DESCRIPTION	TRAIT CODE	UNITS DESCRIPTION	UNIT CODE
Rosette diameter on date measured	rdiam.snapshot	Millimeters	mm
Number of fruits on date measured	fruitnum.snapshot	individual fruits	individual fruits

        iii. Below this you'll see step 3 of metadata entry. Some of this information can be found on the syllabus. Format dates as MONTH/DAY/YEAR (e.g., 03/12/15 or 11/11/15).

    b. Next, at the bottom of the document select the tab for the worksheet titled "DATA."

c. **Each row should represent a single data point** – in other words, one row represents one trait measured for one plant. (*See the example spreadsheet on the next page*)
    i. **Note if you had a missing plant, or the plant died between measurement 1 and measurement 2, then don't mark anything in the value cell (leave the cell blank).**

d. Below are explanations of what to enter into the cells under each column heading. For many columns, what is entered will be the same in every row. An * indicates information that your instructor will provide.
    i. *Experiment: A code describing the experiment name, every row will be the same.
    ii. *Institution: A code for the institution where the project was conducted.
    iii. *Facility: The location within that institution where the plants were grown. i.e. Growth chamber.
    iv. *Treatment: Which treatment the plant being measured was subjected to.
    v. *Accession: The name of the plant line (SALK_XXXXXC or CSXXXXX) -- be sure to include the C at the end of the Salk line name and CS in front of the natural accession.
    vi. *Plant ID: The plant # code/growing location code.
    vii. Trait: A code that describes the trait that was measured *(see directions on page 1)*. Clicking on a cell in this column will reveal an icon with two arrows. Click on it to reveal a drop-down list of options to choose from.
    viii. Value: The trait value measured (e.g., enter "99" if 99 fruits were counted for that plant).
    ix. Units: The units the traits were measured in. Choose the correct units from the dropdown menu.
    x. Date measured: The date the traits were measured **(MUST *enter in format mm/dd/yy)*.**
    xi. Name: The first and last names of both group members separated by a semicolon.
    xii. Comment: Any relevant comments (e.g. notes about damaged or missing plants).

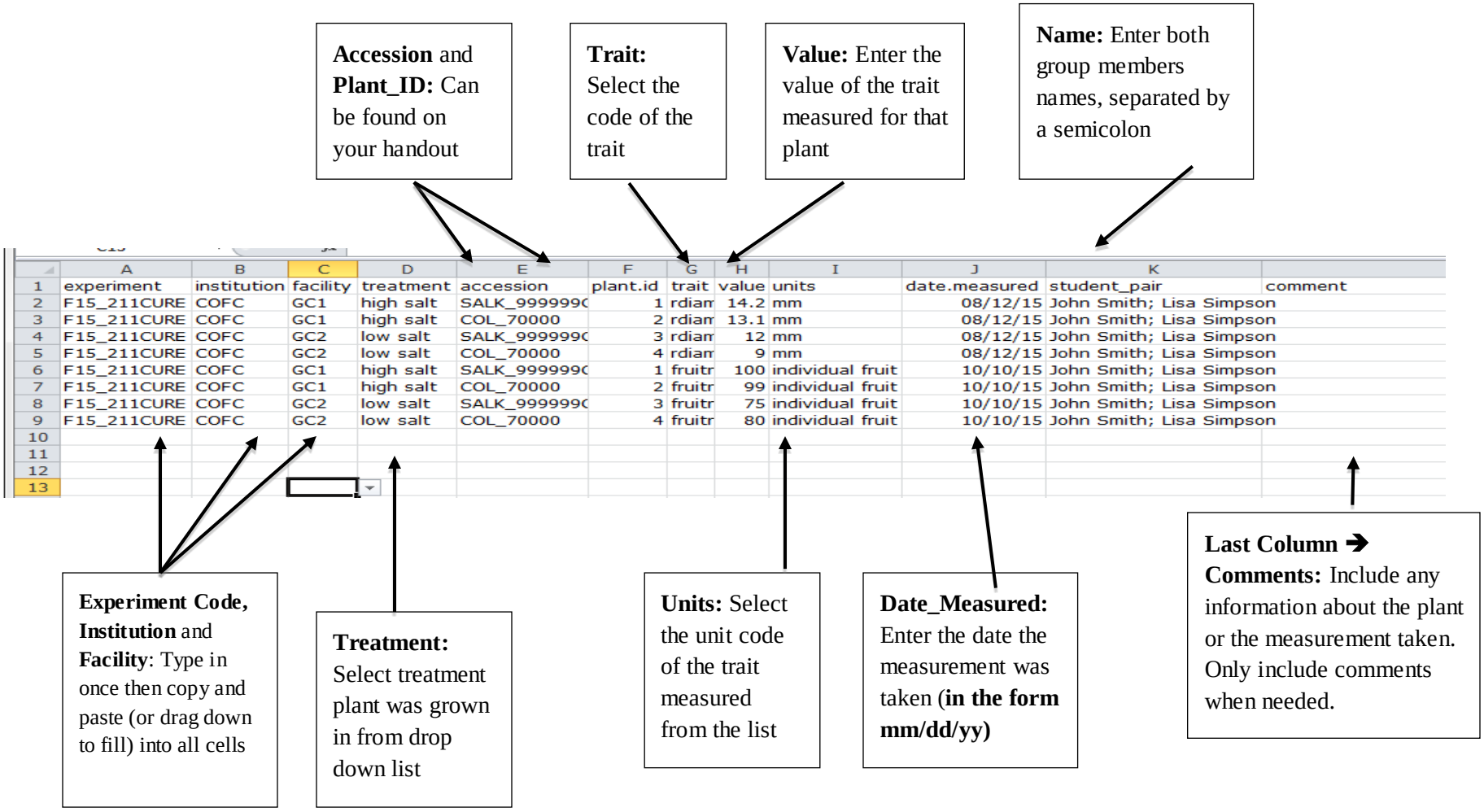
e. ***Spring 2016 CURE Information:**
    i. Experiment ID: S16_BIOL211CURE
    ii. Institution: (USE unPAK designated 4 letter code)
    iii. Facility: GC5 (short for "Growth Chamber 5" which was set to 20C) or GC6 (Growth chamber 6, which was set to 24C). Trays 1-6 were in GC5 and Trays 7-12 were in GC6.
    iv. Treatment: Either 20 or 24
    v. Accession and Plant ID can be found on your earlier handouts

f. There is a third spreadsheet in the Excel document named "Codes" – do not change anything on that sheet. It is used to generate the drop-down lists.

3. Save your datasheet as a new file
    a. Save the file as "S16_BIO211CURE_Student1LastName_StudentTwoLastName.xls"
        i. Note: be sure to save your file as a .xls file, **not .xlsx**

4. Submit your data to your instructor, who will share the data with the unPAK Project Manager.

Example Spreadsheet (treatments from a different semester):
Below is an example spreadsheet showing two traits (rosette diameter and fruit number) measured for four plants (Plant 1 – SALK_99999C grown in high salt; Plant 2- SALK_888888C grown in high salt; Plant 3 – SALK_99999C grown in tap water; Plant 4 – SALK_888888C grown in tap water).



	A	B	C	D	E	F	G	H	I	J	K	
1	experiment	institution	facility	treatment	accession	plant.id	trait	value	units	date.measured	student_pair	comment
2	F15_211CURE	COFC	GC1	high salt	SALK_999999C	1	rdiam	14.2	mm	08/12/15	John Smith; Lisa Simpson	
3	F15_211CURE	COFC	GC1	high salt	COL_70000	2	rdiam	13.1	mm	08/12/15	John Smith; Lisa Simpson	
4	F15_211CURE	COFC	GC2	low salt	SALK_999999C	3	rdiam	12	mm	08/12/15	John Smith; Lisa Simpson	
5	F15_211CURE	COFC	GC2	low salt	COL_70000	4	rdiam	9	mm	08/12/15	John Smith; Lisa Simpson	
6	F15_211CURE	COFC	GC1	high salt	SALK_999999C	1	fruitn	100	individual fruit	10/10/15	John Smith; Lisa Simpson	
7	F15_211CURE	COFC	GC1	high salt	COL_70000	2	fruitn	99	individual fruit	10/10/15	John Smith; Lisa Simpson	
8	F15_211CURE	COFC	GC2	low salt	SALK_999999C	3	fruitn	75	individual fruit	10/10/15	John Smith; Lisa Simpson	
9	F15_211CURE	COFC	GC2	low salt	COL_70000	4	fruitn	80	individual fruit	10/10/15	John Smith; Lisa Simpson	
10												
11												
12												
13												

Appendix 10: Instructor completed examples for SA1, SA2 and SA3

ANSWER KEY AND HIGHLIGHTED SECTIONS WHERE INSTRUCTOR MODIFICATION IS REQUIRED FOR OTHER TREATMENTS

INSTRUCTOR GUIDE TO APPENDIX 2

Student Activity 1 (SA1)

Background *Arabidopsis*: the ecology and genetics of responses to warming temperatures

Welcome to the unPAK CURE project.

Each of these projects is going to be spread across several weeks and will produce a final product, which will be the culmination of your grade on that project.

We will be building hypotheses , collecting data, testing our hypotheses using our data and implementing statistical tests, and presenting our results graphically and in a written format. *The data you collect will be new to science and will be shared with other scientists around the world through the unPAK project database!* **There are opportunities to document biological phenomena that have never been described before….**

Arabidopsis thaliana and climate change

From your online background reading (arabidopsisunpak.org/?page_id=965) , you should have some familiarity with the basics of *Arabidopsis* biology and ecology. In this project, we will work with:

- plants that have been grown in growth chambers at two temperatures.
- plants from a classic lab line (known as "Columbia"),
- and depending on your choice of group you will study EITHER:
  - ***Natural populations group:*** plants collected from two natural populations from Europe
  - OR
  - ***Mutants group:*** plants from the original lab line, the first to have its genome fully sequenced (goes by multiple names: Columbia Col-0 or CS70000), *and* "knockout" mutant lines that have one gene in the genome made nonfunctional with an insertion mutation (see O'Malley and Ecker 2010 for some details).

You will be taking phenotypic data at two points in time (life history stages): early in development (this week) and later in development (in a few weeks). You will construct and test hypotheses about the effects of temperature and your plant groups. However, you won't know which individual plant belongs to which group until after you have collected all your data. We do this to minimize researcher bias. At the end of the project, data that pass peer mentoring and project quality control will be shared in a public database! (www.arabidopsisunpak.org).

Today, we are going to work with a lab partner to accomplish the following:

- Choose which plants you want to work with (***Natural populations*** or ***Mutants***)
- Think about effects of genetics, adaptation and history on phenotypes and develop working hypotheses for response to temperature, and phenotypic plasticity
- Select phenotypes to measure with measuring tools available
- Measure phenotypes
- Enter your data in an Excel spreadsheet (following the example).

Experiment set-up (*what's been going on behind the scenes*):

- On (add instructor specific date) all seeds for this class were set up by your instructors to be cold treated on wet filter paper for one week. This cold period synchronizes the germination of seeds.
- Following cold treatment on (instructor specific date), seeds were transferred onto potting mix and into a randomized design into 12 trays in 2 growth chambers by your instructors. Seedlings were transferred to 2.5 inch pots on standard ProMixBx potting mix. Plants were bottom watered in their trays two times per week, such that potting mix remained moist.
- Your plants have been grown at two temperatures (20°C and 24°C OR INSTRUCTOR SPECIFIC TREATMENTS) in Percival growth chambers. Plants were grown in a growth chamber with 16 hours of light per 24 hour period under constant temperature.
- Half of your plants have been grown at 20°C. *This temperature would correspond to a warm spring day within the normal range of native Arabidopsis populations.*
- The other half has been grown at 24°C. *This temperature reflects what average spring temperatures might look like in 2100 under the predictions of the Intergovernmental Panel on Climate Change (IPCC).*
- Plants are grown in trays (green trays) that hold 32 plants each. Plants of different genotypes are being grown in a randomized design.
- Each group will have two different accessions (Natural Population A & Natural Population B – **or** – Columbia 70,000 parental line & Salk-TDNA Mutant). You will measure approximately 6-9 replicate plants per accession per treatment (between 24 and 35 plants total) based on germination and early survivorship.
  note: We use the words line, accession, genotypes, maternal line as interchangeable descriptions of the seed material with the same genetic heritage.

Work with your lab partner and come up with 5 possible effects that warmer temperatures might have on plant growth and development.

Natural populations group: Some of you will work with plants that have been collected from natural habitats throughout Europe. If you are in this group, proceed by answering the following questions.

What kind of habitat characteristics might vary in important ways between the ecological locations you would find *Arabidopsis* in?

Answers will vary but include temperature, altitude, day length, soil nutrient characteristics, soil texture characteristics, other plant species, disturbance regime, or proximity to human modified landscape.

Identify 3 plant characteristics that might differ because of adaptations to local environmental conditions.

Answers will vary but can include, height, flower production, seed set, and root size among others.

Why might different populations respond to our temperature increase in different ways? Are there any reasons there might not be a response to temperature?

Answers to the first question will vary but could include:

- Plants can show local adaptation to certain geographic locations.
- Plants may have physiological responses that maintain the morphological phenotypes across environments.

Answers to the second question will vary but could include:

- Natural accessions may not be responsive to differences of only 4°C.
- Reponses may be due to differences in light, soil or water but temperature may not affect photosynthesis or metabolism.

Mutants group: Other groups will work with two mutant lines, each of which has one of the 27000 Arabidopsis genes disabled. The disabled gene has been "knocked out" by insertion of a "nonsense" sequence by an infectious bacterium, *Agrobacterium tumefaciens*. This set of lines are called Salk T-DNA lines. Subsequent molecular biology research by the folks at Salk Institute determined just where that insertion occurred. See the article: O'Malley, RC. and JR. Ecker. 2010. Linking genotype to phenotype using the Arabidopsis unimutant collection. The Plant Journal 61: 928-940.

Any phenotypic effects that you discover may be new to science! However, in most types of experiments it isn't possible to discern the effects of a single mutation. If you are in this group, proceed by answering the following questions.

Why might it be hard to detect the consequences of knocking out a gene?

Answers can vary and may include:

- Genes may influence physiological characteristics, which may not be apparent when employing morphological measurement techniques used in this class.
- Insertion mutations may still allow for a gene to partially function.
- Single genes may have very small effects that are not observable in comparison to the contributions of the remainder of the genome.

What kind of phenotypes do you think it would be interesting to examine to determine effects of mutation?

Answers will vary and may include such examples as any leaf traits, which influence plants' capacity to photosynthesize; growth rates or characteristics, which may indicate developmental differences among lines; or fruit production, which is an important fitness trait.

Now that you have thought a bit about the types of phenotypes that could be associated with temperature, natural populations with known origin, or mutants with single genes knocked-out, it's time to get some data.

EXAMPLE DATA USED BELOW

INSTRUCTOR GUIDE TO APPENDIX4 FOR NATURAL ACCESSIONS

Student Activity 2 (SA2)

Literature and Database Background: Natural Populations Groups

Last week, you took data on *Arabidopsis thaliana* phenotypes looking at plants from two genetic backgrounds (two natural populations) grown in two temperatures (20 & 24C).

Your assignments today are to:

3. **Learn about the locations that your natural populations came from**. Work on this section in collaboration with your lab partner.
4. **Locate 5 primary literature sources** relevant to your populations or the study and read the abstracts associated with these papers. This will result in a total of 10 citations (5 per partner —it will be helpful for the final assignment if you turn up different papers!
    a. Type up these references using the bibliography format below (see last page of the assignment). It is *not acceptable* to copy and paste from Web of Science *without* reformatting. **Bring this list for next week including a copy for your instructor and a copy for your lab partner.**

About your populations

Your instructor will give you information on the accession numbers of your two populations (given as CS#####). From there you'll find out if there are additional codes for your populations. The first step is to figure out where those populations are from. In this handout population, line, accession and ecotype are all synonyms that describe a set of seeds that was collected from a particular natural population and has been cataloged in the Arabidopsis database called The Arabidopsis Information Resource (TAIR).

1) To do this you are going to use TAIR. TAIR is the current central clearinghouse for most Arabidopsis information.

6. Go to the TAIR website: www.arabidopsis.org
7. Your first step is to find your populations, also known in this database as ecotypes.
8. Type the name of your population into the search bar at the top. Make sure you switch your search to the "germplasm" or "seed stock" database, as the default is "gene".
9. You should find your ecotype on a list. Clicking on the link will take you to a page with various data about the population, including the city / region and country where it was originally collected.
10. An example using TAIR to find information about your lines can be found here: http://arabidopsisunpak.org/?page_id=220

Where are your populations from?

Population: Gückingen, Gu-0, Glueckingen Germany
Population: Tossa del Mar, Spain Ts-1

2) You might want to look around this page a bit, as it may contain some additional information about phenotypes or even publications associated with your population. Write any relevant information here.

3) Now it's time to use Google. First, figure out the latitude and longitude of your locations. Talk to your instructor to find out if your populations have "unusual" names.

Population: Gu-0 50.3, 8

Population: Ts-1 41.7194; 2.93

4) You may also want to use Google-maps or other tools such as https://www.nhc.noaa.gov/gccalc.shtml to estimate distance between populations.
1030 km

5) Can you find anything out about the climate at your site of origins? One site to look at is wunderground.com

- Type in the name of the location that your population is from.
- Scroll down the page to the Almanac section.
- Click on the link for September Calendar View.

Most of the time, Arabidopsis grows primarily in the spring between the end of February and the end of April. You will need to change the month in the calendar view to each of February, March and April. What are the average temperatures like at your sites during those time periods?

Population: Gu-0, high/low 40°F/32°F (4.4°C/0°C) February, 48°F/37°F (8.89°C/2.78°C) March, 55°F/41°F (12.78°C/5°C) April

Population: Ts-1 high/low 59°F/36°F (9.4°C/2.2°C) February, 63°F/40°F (17.22°C/4.4°C) March, 67°F/44°F (19.4°C/6.67°C)April

What is precipitation like at your sites at those times?

Population: 3.2-2.7 inches per month

Population: 1.8-2.5 inches per month

Now that you know a bit about where your populations are from, do you think they would differ in their response to the temperature treatment? Why or why not?

The population from Germany may be more sensitive to a warming climate than the one from Spain due to the typical climate regime present in Germany (cooler with more precipitation) than the one in Spain (warmer with less precipitation).

6) Your next step is to look for some literature relevant to the study. You may wish to use the literature databases and a search engine (e.g. Web of Science and Google-Scholar) to find out if anyone has studied your populations before. What did they find out? What do we know about Arabidopsis growth or size and temperature? Has anyone else previously studied your phenotypes? What is known about the phenotype rosette diameter, your group trait or the class trait?

Some examples: Ferguson L, Sancho G, Rutter MT, and Murren CJ. 2016. Root architecture, plant size and soil nutrient variation in natural populations of *Arabidopsis thaliana*. Evolutionary Ecology. 30:155-171.

Notes about this paper

Huang XY and Salt DE. 2016. Plant ionomics: from elemental profiling to environmental adaptation. Molecular Plant. 9:787-797.

Notes about this paper

Find 5 primary literature references that are relevant to your study. Email either the abstracts or papers to yourself and record the citation information so that you can clean up the information into proper format for a bibliography. It would be quite useful to find a general paper on Arabidopsis ecology, several about plasticity or performance in response to temperature, and if possible population variation for your group or class phenotype.

7) Feel free to record any interesting background information about your populations that might be useful in a lab report against which you may want to contrast your results.

Numerous various answers here. The Tossa del Mar population (Ts-1 41.7194; 2.93) is Na+ tolerant (salt tolerant). Citation: Rus A, Baxter I, Muthukumar B, Gustin J, Lahner B, Yakubova E, Salt DE. 2006. Natural variatns fo AtHKT1 enhance Na+ accumulation in two wild populations of Arabidopsis. PLoS Genetics. 1:e210.

What is due next week?

1) A list of 5 journal sources (different for each person in pair) using the reference format below.

Example reference format. Copy the formatting and punctuation exactly!

Duncan AB, and Little TJ. 2007. Parasite-driven genetic change in a natural *Daphnia* population. Evolution 61: 796-803.

Conner JK, Davis R, and Rush S. 1995. The effect of wild radish floral morphology on pollination efficiency by four taxa of pollinators. Oecologia 105: 509-516.

EXAMPLE DATA USED BELOW

INSTRUCTOR GUIDE TO APPENDIX4 FOR MUTANT GROUPS

Student Activity 2 (SA2)

Literature and Database Background: Mutant Groups

Last week, you took data on *Arabidopsis thaliana* phenotypes looking at plants from two genetic lines (Columbia and a knockout mutant) grown at two temperatures (20, 24C).

Your assignments today are to:

4. **Learn about the gene / genes that were knocked out in your mutant lines**. Work on this section in collaboration with your lab partner.
5. **Learn about the parental strain Col-0.** Work on this section in collaboration with your lab partner.
6. **Locate 5 primary literature sources** relevant to your mutants or the study and read the abstracts associated with these papers. This will result in a total of 10 citations (5 per partner —it will be helpful for the final assignment if you turn up different papers!)
    1. Type up these references using the bibliography format below (see last page of the assignment). It is *not acceptable* to copy and paste from Web of Science *without* reformatting. **Bring this list for next week including a copy for your instructor and a copy for your lab partner.**

Your mutants

Your instructor will give you information on the Salk mutant line that you are investigating. You are also studying Col-0 (a.k.a. Columbia) – the original line to have its genome fully sequenced and the parent of the mutant line. The first step is to figure out what locus/loci are affected by those mutations and what we may (or may not) know about the line.

1) To do this you are going to use The Arabidopsis Information Resource (TAIR). TAIR is the current central clearinghouse for all Arabidopsis information.

6. Go to the TAIR website: www.arabidopsis.org
7. Your first step is to find your mutant lines...
8. Type the name of your line exactly (as in "SALK_138583C") into the search bar at the top. **VERY IMPORTANT:** Make sure you use the pull-down menu to switch to the ***germplasm*** database (the default is gene).
9. You should find your line listed on a search results page. You want to click on the **locus** that has been affected by the mutation by clicking under the gene ID (**this will start with the letters AT**) under Polymorphism / Locus.
10. More details about using TAIR to find information about your lines can be found here: http://arabidopsisunpak.org/?page_id=230

2) What locus/loci are affected by your mutation? Give the AT # and any other names you may find for the gene.
Mutant: SALK_010647 AT2G47790 GIGANTUS1 or GTS1 or GTS

3) Give some basics about this line from the Description field.

Mutant:
"Encodes GIGANTUS1 (GTS1), a member of Transducin/WD40 protein superfamily. This gene controls seed germination, growth and biomass accumulation."

4) What kinds of plant tissues is the gene normally expressed in? What kinds of functions does it have? Some of this is summarized in the Annotations section. One possible option is no function is currently known or described.
Mutant:
This gene is present in leaf tissues and in tissues associated with flowering. It might function in determining leaf structure, flowering structure, or inflorescence-related traits.

5) You might want to look around this page a bit, as it contains information about any previously studied phenotypes or publications which also studied this line. (hint: these may be helpful citations for your paper). The site **araport.org** can also be useful for searching for your gene (starts with AT), as it is another collection of information about Arabidopsis genetics.

6) Another site to look at is: signal.salk.edu; click on the T-DNA express link at the top of the page. More information about using this webpage can be found at: http://arabidopsisunpak.org/?page_id=230 (note—do not include the "C" at the end of the Salk line as shown in the image). The part of the gene where the insertion occurs is usually indicated by **LOCN:**
What part of the gene is the insertion located in? Common options include exon, intron, UTR (untranscribed region), promoter.

Mutant: SALK_010647 exon

Why might the insertion location matter?
The protein may or may *not* be rendered non-functional with the insertion mutation.
7) Another site to look at is: http://arabidopsis.info/CollectionInfo?id=94 You will find background on the "Columbia" line. You may also choose to do an ecotype search on "Columbia" or Col-0. The control line we are using is CS70000.

8) Your next step is to look for some literature relevant to the study. You may wish to use the literature databases and a search engine to find out if anyone has studied your locus before. What did they find out? What do we know about Arabidopsis growth, size and temperature? Or your

group or class phenotypes? Or your Salk T-DNA line?

Growth and size influenced by GTS1

Gachomo, E. W., Jimenez-Lopez, J. C., Baptiste, L. J., & Kotchoni, S. O. (2014). GIGANTUS1 (GTS1), a member of Transducin/WD40 protein superfamily, controls seed germination, growth and biomass accumulation through ribosome-biogenesis protein interactions in Arabidopsis thaliana. *BMC plant biology*, *14*(1), 37.

Notes on paper

Find 5 primary literature references that are relevant to your study. Email either the abstracts or papers to yourself and record the citation information (see below). Additional references that will be helpful are general *Arabidopsis thaliana* ecology, a reference describing the Salk T-DNA mutants, Arabidopsis plasticity of rosette diameter or group or class trait, and Arabidopsis response to temperature. REVIEW papers about Arabidopsis ecology or about T-DNA insertion phenotypes that might be especially helpful.

9) Feel free to record any interesting background information about your loci that might be useful in a lab report once you have your results.

The paper about the mutant suggests it grows quickly.

What is due next week?
1) A list of 5 sources using the reference format below (5 that are different from your partner).

Example reference format. Copy the formatting and punctuation exactly!
Duncan AB, and Little TJ. 2007. Parasite-driven genetic change in a natural *Daphnia* population. Evolution 61: 796-803.
Conner JK, Davis R, and Rush S. 1995. The effect of wild radish floral morphology on pollination efficiency by four taxa of pollinators. Oecologia 105: 509-516.

EXAMPLE DATA USED BELOW

INSTRUCTOR GUIDE TO APPENDIX3

Student Activity 3 (SA3)

Hypotheses, graphing and statistics: *Arabidopsis thaliana*

You have watched the Excel YouTube video and completed and submitted the graphing and statistics homework on a classic ecology dataset (working with Excel homework http://arabidopsisunpak.org/?page_id=965). Your goal today is to graph and calculate statistics on your own Arabidopsis data that will contribute to discussing your biological argument in your lab report. First, recall what traits you measured and what you learned about your plant lines (mutant or natural populations). Refine your earlier overarching hypothesis about differences or relationships among your traits or lines from what you have learned additionally about these particular lines and treatments from your further literature readings.

*Your instructor will give you information about which plants were from which genotype, and which plants were from which temperature treatment. Add this data to your data file.

* What traits did you measure? Were these traits continuous or categorical? What else do you know about your plants? Are these factors continuous or categorical?

Rosette diameter, fruit number, *other traits selected by the student*

The measured traits of rosette diameter and fruit number were continuous characters. *The other traits could be categorical or continuous depending on what the students choose.*

The temperature treatment and genotype are categorical factors.

* Your lines: SALK_010647C

* General hypothesis: (Example: Temperature affects plant growth, but the effects differ between genetic lines.) Initially (SA1), we coach students to build general hypotheses based on material introduced to that point in the course, from the prerequisite courses, and from the information in the pre-lab reading material. From there, students are guided to engage further in the scholarly literature and database information, and thus further refine their hypotheses.

List 2-3 refined hypotheses based on the ones that you developed in SA1. Draw on your newly obtained knowledge from scholarly sources, background information from databases (developed in handout SA2) and additional information we have examined in the lecture component of this course in relation to plant biology, ecology and genetics.

Mutant line 1 has smaller rosettes in high temperature than in low temperature. Rosette size is positively correlated with cauline leaf number. Thus, it is unsurprising that mutant line 1 has fewer cauline leaves in high temperature than in low temperature.

Identify 4 specific tests you will perform. **PLEASE PERFORM AT LEAST ONE CORRELATION AND AT LEAST ONE T-TEST. It is hard to imagine that you will not want to contrast your different types of plants (populations, or mutant vs. control) and to contrast plants in the different treatments.**

1. T-test to compare rosette diameter of Col-0 and SALK_010647C
2. T-test to compare rosette diameter between 20C and 24C
3. Correlation to evaluate the relationship between rosette diameter and fruit number
4. Correlation to evaluate the relationship between rosette size and cauline leaf number

Outline a way in which your 4 tests will allow you to tell a coherent story about the biology of your plants.

These tests will allow me to examine the relationships between photosynthetic structures and fruit production.

Now generate the appropriate statistical test and a figure to go with each specific hypothesis. ***Note: You cannot do a correlation analysis using "temperature treatment" or "genotype" as a variable – these are both categorical.**

Provide a one sentence description of your results, including the statistics. Please show the sentences to your instructor before finishing today.

Test 1: We detected a significant difference between Col-0 and GTS in rosette diameter (df=11, t= 4.4 p<0.05).

Test 2: We detected a significant difference between temperature treatments (df=11, t=6.2, p<0.05).

Test 3:

We detected a significant positive correlation between rosette diameter and fruit production (r=0.88 p<0.001, df=30).

Test 4:

We detected a significant positive correlation between rosette diameter and cauline leaf number (r=0.62, p<0.001, df = 30).

Appendix 11: Supplemental Table Faculty Feedback

Supplemental Table Faculty Feedback. Included are select quotes from faculty from seven different institutions on their positive experience in implementing the unPAK CURE. These are in response to these questions: *Why did you choose to implement a CURE? What goals did this module accomplish that a non-research based course might not have accomplished? What was the most important outcome of this module for your course?*

Why instructors chose to implement:
"I was looking for a way for my introductory biology students to gain perspective on the excitement, frustrations, and difficulty of basic scientific research."
"I chose CURE because I wanted to give students hands on experiences in collecting ecological data. Giving the students hands-on experiences allowed them to much better acquaint themselves with the course material in general."
"I was looking for an open-ended exercise that would engage students in a project for which the outcome could not be predicted in advance."
"I liked the idea of a small project that fit into a larger project, because it allowed me to explain why certain protocols had to be followed in order to maintain cross-institutional consistency."
"The AIBS Vision and Change recommendation puts equal emphasis on core concepts and core competencies. Reading over the list of competencies, it is hard to escape the conclusion that one of the best ways to integrate them into a course is through a CURE"
Outcomes of implementing the module:
"It was invaluable for the students to see that no dataset is perfect, and that sometimes "negative" results are an advance in scientific understanding."
"The most important outcome of this module was the ability to get students excited about science, data collection and really getting an in-depth understanding of ecological concepts."
"It was great to be able to tell them that their work also contributes to the wider understanding of environmental pressures, by combining their data with those of others."
"For the first year, the most important outcome was that students felt ownership of a project. They were invested in it in a different way than a typical lab exercise."
"That students learn about the process of science with special emphasis on scientific communication, specifically writing scientific papers."
"Sometimes a student will tell me that they've never done a 'real' experiment before. Every experiment they've done until my CURE has been something with an outcome that was easily predictable. It means a lot to me that students get an authentic experience."