Table of Contents

In-Lab Resources:

The protocols for cloning the CRISPR-Cas9 constructs are adapted from **Methods S2: Golden Gate cloning method for the assembly of one or two gRNAs** by Xing et al., 2014 *BMC Plant Biology* (https://doi.org/10.1186/s12870-014-0327-y).

The pipeline can be completed over the course of 9 lab periods. Day 3 and Day 6 may require students to leave lab early while PCRs are running and return later to run gels. Pre-made PCR cocktails can streamline the setup process and make this unnecessary. Because *Agrobacterium tumefaciens* has a longer recovery time than *E. coli*, it may be better to perform transformations first and determine whether constructs are correct during the recovery period or shorten the recovery period to ~1+ h before plating.

Appendix 1: Summary of student demographics

The table below shows the ethnicity of students that participated in the in-person CURE 2018-2019 and the remote-learning CURE 2019-2020.

	2018-2019 in person		2019-2020 remote		Cumulative	
	Number	**Percent**	**Number**	**Percent**	**Number**	**Percent**
Black	11	2.24	7	1.36	18	1.79
Hispanic	127	25.81	120	23.39	247	24.58
Asian	251	51.02	296	57.70	547	54.43
White	58	11.79	60	11.70	118	11.74
Other	45	9.15	2	0.39	47	4.68
Multiple	--*	--	28	5.46	28	2.79
Total	492	100	513	100	1005	100

* Multiple category not included in the 2018-2019 survey.

Appendix 2. Learning outcomes for each week of the remote-learning and in-person CURES. Outcomes specific to the remote and in-lab versions are indicated. Numbers are used for alignment of outcomes with exam rubrics (Appendix 29).

Week	Learning Outcomes
1	Students will be able to: ● Explain the overall goal of the project and identify their specific role. ● (1) Use online resources and databases to research *A. thaliana* gene information. ● Create a detailed lab notebook entry. Remote: ● (2) Design PCR primers with the aid of online tools and databases to amplify their assigned *A. thaliana* gene. In-lab: ● (3) Identify homologs of their gene and design guide RNAs targeting their gene and any close homologs.
2	Students will be able to: Remote: ● (4) Define homology and identify gene homologs through DNA database searches. ● (5) Discuss the evolutionary relationships between paralogs and orthologs. ● (6) Create a phylogenetic tree for a gene family using online tools and identify paralogs and orthologs for their assigned gene. In-lab: • Incorporate their guide RNAs into a DNA cassette using PCR. • Set up a Golden Gate cloning reaction to incorporate their guide RNA cassette into a CRISPR-Cas9 vector.
3	Students will be able to: ● (7) Discuss the mechanism of CRISPR-Cas9 mutagenesis. ● (8) Form a hypothesis about the potential consequences of mutating a gene and what that suggests about the gene's function. Remote: ● (9) Design guide RNAs to specifically target different regions of a gene to create a null mutation. ● (10) Evaluate and select guide RNAs based on a given set of criteria to selectively target a single gene. ● (11) Design PCR experiment to detect successful gRNA guided deletions in *A. thaliana*. In-lab: • Introduce their finished CRISPR-Cas9 construct into *E. coli* using bacterial transformation. • Perform colony PCR to identify bacterial colonies successfully transformed with their construct.
4 and 5	Students will be able to:

	● Effectively communicate the project and results in a scientific report. ● Create and present slides that summarize the project and their results to their peers. In-lab: ● Analyze DNA sequences of CRISPR-Cas9 constructs to check for PCR induced mutations. ● Transform *A. tumefaciens* with their CRISPR-Cas9 construct. ● Transform *A. thaliana* plants with their CRISPR-Cas9 construct by floral dip.

Appendix 3. Suggested pacing for the remote learning course organized into 5 weeks of instruction.

Week	Activities	Materials
1	**Introduction to Research Project and Primer Design** -Watch the project introduction video. -Design primers for maize *Actin*. -Assign genes to students. -Students design primers to amplify across their assigned gene on their own. -Students record their primers in a notebook entry based on the provided example. -Students complete Quiz 1	-Project introduction video (Appendix 5) -Project report guidelines (Appendix 6) -Primer design - maize *Actin* (Appendix 7) -Primer design - *A. thaliana* (Appendix 8) - Example Notebook Post - Primer Design (Appendix10) -Quiz 1 (Appendix 11)
2	**Phylogeny and Homologs** -Watch the “Gene evolution and homology” video in class. -Discuss the “Phylogeny and Using Plaza” reading in class. -Students create a phylogenetic tree for their assigned gene and use it to identify homologs. -Students complete a notebook entry.	-Gene evolution and homology video (Appendix 5) -Phylogeny and Using Plaza activity (Appendix 13) -Example Notebook Post - phylogenetic tree (Appendix 14)
3	**Guide RNA Design** -Use the E-CRISP protocol to review CRISPR-Cas9 and design guide RNAs of a specific example in class. -Students design guide RNAs for their assigned gene. -Students record their guide RNAs in a notebook entry and in a class spreadsheet. -Students complete Quiz 2	-E-CRISP protocol (Appendix 15) -Example Notebook Post - guide RNAs (Appendix 16) -Example class spreadsheet (Appendix 17) -Quiz 2 (Appendix 18)
4	**Project Presentation and Future Directions** -Discuss the Project Presentation guidelines in class. -Students create a short project presentation. -Discuss future directions for the project using the “Future Directions” reading in class.	-Presentation guidelines (Appendix 20) -Future Directions (Appendix 24) -Example student report (Appendix 25)
5	**Student Presentations** -Students present their results to the class. -Students take a final exam for formal assessment.	-Example presentation score sheet (Appendix 22) -Example Student Presentation (Appendix 23) -Final exam (Appendix 27)

Appendix 4. Modified research project course showing additional in lab activities.

Week	Activities	Materials
1	-Introduce project to class. -Assign candidate genes to students. -Students use BLAST and phylogenetic comparison to identify close homologs for their gene. -Design guide RNAs targeting assigned gene and close homolog(s). -Instructor orders primers to PCR amplify guide RNAs the following week.	-Selection of Target Gene(s) (Appendix 30) -Guide RNA Design (Appendix 31) -Designing and Ordering primers (Appendix 32)
2	-PCR amplify guide RNA cassette. -Golden Gate clone guide RNAs into CRISPR-Cas9 vector and transform *E. coli* with Golden Gate reaction mix.	-PCR amplification of guide RNA cassette (Appendix 34) -Golden Gate Cloning Protocol and *E. coli* transformation (Appendix 35)
3	-Introduction to transformation in bacteria and plants -Colony PCR	-Colony PCR of *E. coli* transformants (Appendix 36)
4	-Grow positive colonies overnight for minipreps. -Students perform DNA minipreps and make glycerol stocks of overnight cultures. -Send colonies for sequencing. -Students check sequences for errors. -Transform vector into *A. tumefaciens.*	-DNA minipreps of CRISPR/Cas9 constructs (Appendix 37) -Transformation of *A. tumefaciens* (Appendix 38) -Sequence analysis (Appendix 39)
5	-Transform plants using floral dip. -(optional) Demonstrate selection of transgenic seeds.	Transformation of *A. thaliana* (Appendix 40)

Appendix 5: List of videos

Week 1: Project introduction: https://youtu.be/vBNoT--RL9U

Week 2: Gene evolution and homology: https://youtu.be/-FkjzC8VYXQ

Appendix 6: Project report guidelines

Report Guidelines

The report is based on your research project. In the essay you will discuss the purpose of the project, address the experimental strategy, and present your findings. The report will be evaluated on completeness and accuracy based on the following guidelines.

The Report Format:

• Must be submitted as a Word document (.doc or .docx).
• Use 12-point font.
• Double spaced.
• Default margins (1 inch top and bottom, 1.25 inches left and right).
• One extra return between titles and subtitles.
• The final draft should be 4-5 pages in length (excluding citations, figures, and tables).
• Place figures and tables at the end of document. Refer to the figures in the text: "As seen in Figure 1…", or "Three new alleles of *PIK3* were identified (Fig. 1) including a …".
• Paper must include at least three (3) literature references. Use PubMed or Google Scholar to find these. The course book, handouts, lectures, and videos can be used but DO NOT count as one of the 3 primary literature references. Refer to the citation section in these guidelines.

Due Dates and Points:

• (15 Points) Intro draft due Monday, May 18, 5:00 P.M. The draft must be complete and written in proper English. It is NOT an outline. Do not use idioms, colloquialisms, or texting abbreviations. Instructors reserve the right to determine "completeness." The draft will be submitted on iLearn.
• (15 Points) Methods & Data draft (or Experimental Design) due Monday, May 25 by 5:00 P.M. Same rules as Intro draft.
• (70 Points) Completed paper with all corrections and with all the sections, due Monday, June 8, 5:00 P.M. Submit this via iLearn. Note that in order to earn the maximum of 70 points all drafts must be completed on their respective due dates and receive at least 8 points in each

Contents of the Report Intro Draft:

• Title (make it descriptive to give a sense of what the project entails)
• Background on cell division in plants
    o Why is cell division a regulated process?
    o What is known about division plane orientation in plants?
    o Discuss what is known about TAN1 and the mutant *tan1*?
    o What is a model organism? How is *Arabidopsis thaliana* a good model?
    o What is CRISPR-Cas9? The length of this section, or draft, should be about 2 pages of text.

Methods & Results draft (Experimental Design):

Methods:
• Briefly describe the methods you used for the project in complete sentences.
Data:
• Describe what **each** of your figures shows and how that result relates to the project goal.
• What is your assigned gene (name and locus)?
    o Known/reported function
    o Gene structure (exon/intron)
    o Primers (location, sequences, product size)
• Phylogenetic tree (known homologs)
• Gene expression pattern
• Guide RNA sequence
    o Primers for gRNA construct (location, sequences)
    o Make sure it is a unique construct (different than your partner's)
• Make sure to include legends for your figures and refer to them in the text when appropriate. An example is shown below to discuss your assigned gene's structure. "MAP95-6 has 9 exons and 8 introns (Fig 1)."

Example figure and legend:

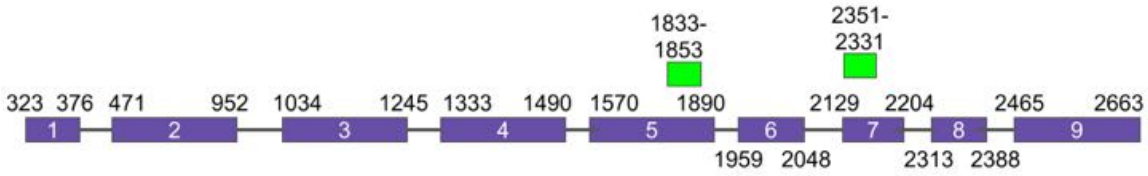


Figure 1. A gene diagram of MAP65-9 (AT5G62250). The purple rectangles represent the gene exons connected by lines representing introns. The numbers represent the nucleotide positions of each exon in the genomic sequence. The small green boxes represent the location of the primers to identify if the sample is cDNA or genomic DNA.

The length of this section, or draft, should be about 2 pages of text (not including figures and tables).

Final Draft (Conclusion & Future Directions):

• Combine the Intro and Methods & Data drafts. Please make corrections or edits previously suggested.
• Add a new section for Conclusions & Future Directions
    o How to make a successful plasmid construct that can knock out your intended target gene/homolog?
    o How is *Agrobacterium tumefaciens* used in the context of knocking out genes in the *Arabidopsis thaliana* genome?
    o How would a plant lacking a functional protein encoded by your gene help advance the study on cell division in plants?

The whole draft should be about 5 pages of text.

Styling Guidelines Genetic Nomenclature:

• Wild type gene: *TANGLED1*
• Wild type gene symbol: *TAN1*
• Mutant gene symbol: *tan1*
• Protein: TANGLED1
• Protein Symbol: TAN1

Typically, the first time the gene (or protein) is referred to spell out the name with the symbol in parentheses, e.g., “TANGLED1 (TAN1) is….”
Species nomenclature:
Scientific species names are given as Genus species: e.g., *Arabidopsis thaliana*. If the binomial is referred to subsequently use *A. thaliana*.

Plagiarism:

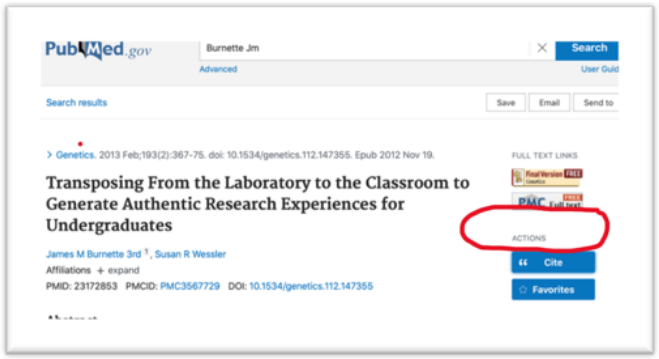
Plagiarism will not be tolerated. The drafts will be submitted to SafeAssign to check for plagiarism. Do not use direct quotes in the report. Re-phrase in your own words and cite the source of the information. Any instance of plagiarism will result in a zero for the assignment and referral to student conduct committee for review. Please refer to the Academic Dishonesty Policy for more information
http://www.cs.ucr.edu/cs14/cs14_04spr/handouts/ucr_aca_dis_policy.html.

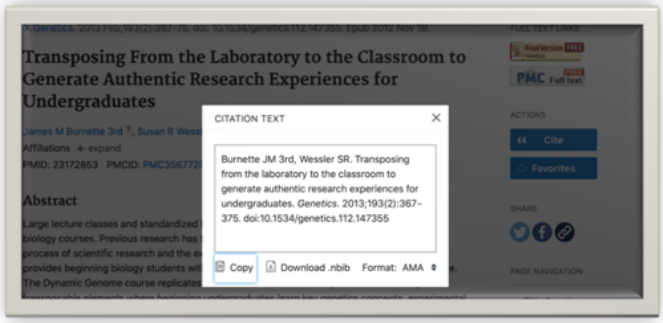
Citations:

In-text citations:
Information cited from other sources is credited by placing the last name of the author(s) and year in parentheses. Citations in the body of the report look as follows:

CCA1 may play a role in regulating the rhythmic transcriptome (Nagel et al., 2015). Gene promoters serve as species-specific switches to turn a gene on or off (Burnette and Wessler, 2016).

Works cited section:
At the end of the report include a “Literature Cited” section and list the full citations in alphabetical order. To get the proper citation use PubMed: 1. Click the blue “ Cite box and 2. Copy and paste the default citation.





Note the web-link for a literature source is not correct. Check with instructor for help finding the correct citation format. Acceptable formats include MLA and APA, just be consistent. The Purdue Online Writing Lab offers useful formatting guidelines for different sources of materials: https://owl.english.purdue.edu

Appendix 7: Primer design - maize *ACTIN*

Primer Design – Maize Actin

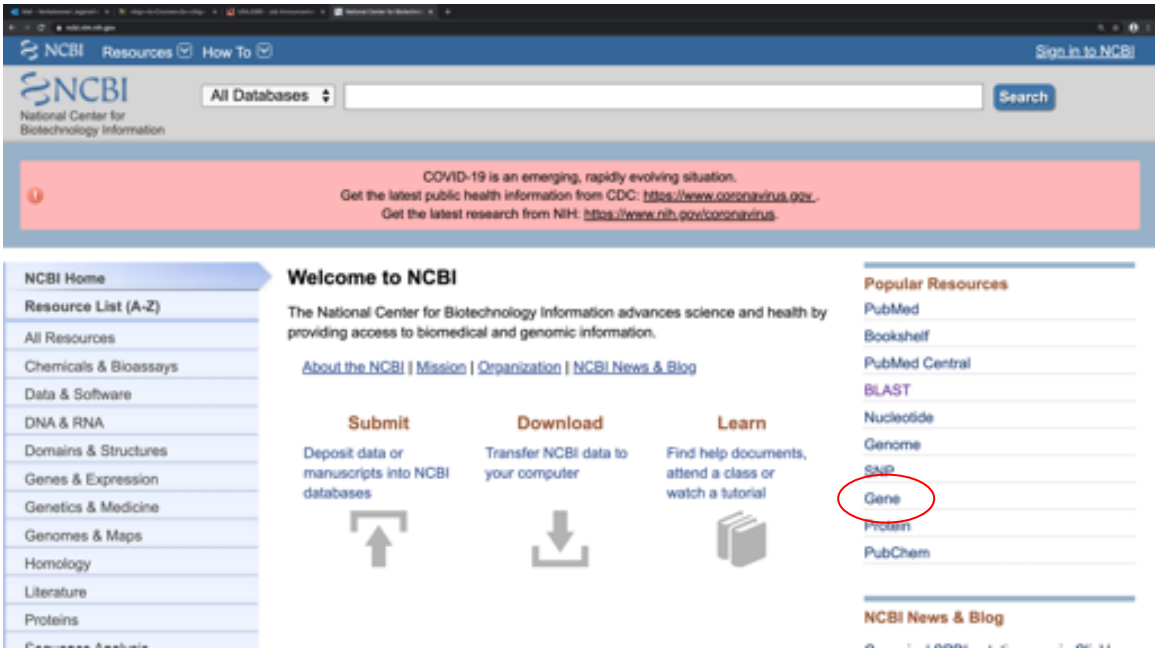
In previous experiments you learned about using PCR to amplify the *ACTIN* gene from genomic DNA and cDNA and to study transposable element variation in maize. In addition to this you learned about how to analyze the DNA sequencing results using NCBI BLAST and to compare gene sequences through MUSCLE. In today's activity you are going to learn about primer design.

Activity

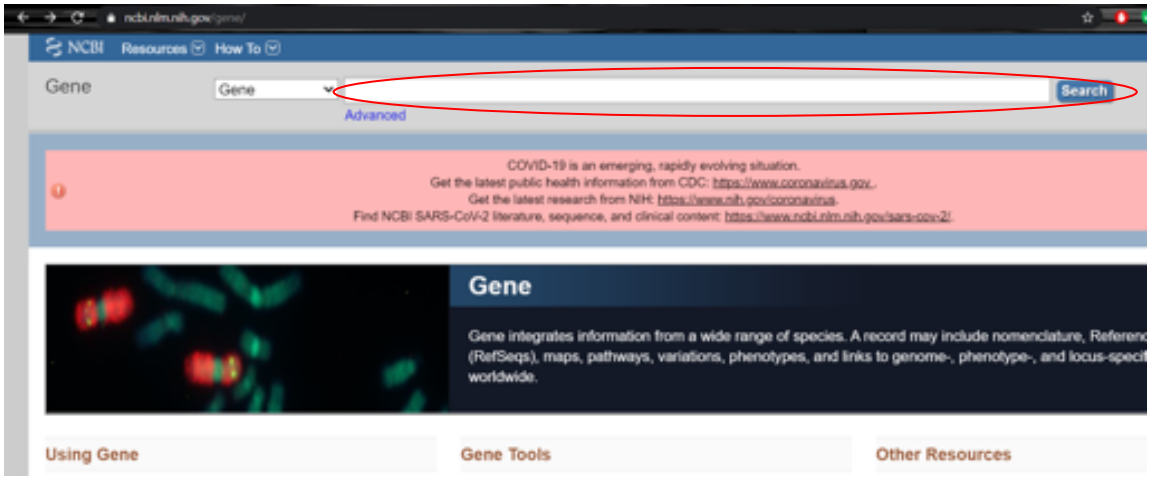
Use the following steps to design primers using NCBI's Primer Blast:

Part 1 – Choose a gene of interest using NCBI. Copy the gene sequence (both genomic and coding sequence (CDS)).
Part 2 – BLAST Align the two sequence against GenBank database.
Part 3 – Check the graphic summary. Analyze the intron and exon positions.
Part 4 – Primer design tool: Primer Blast.
Part 5 – Write down the primer positions and amplicon sizes of gDNA and CDS.

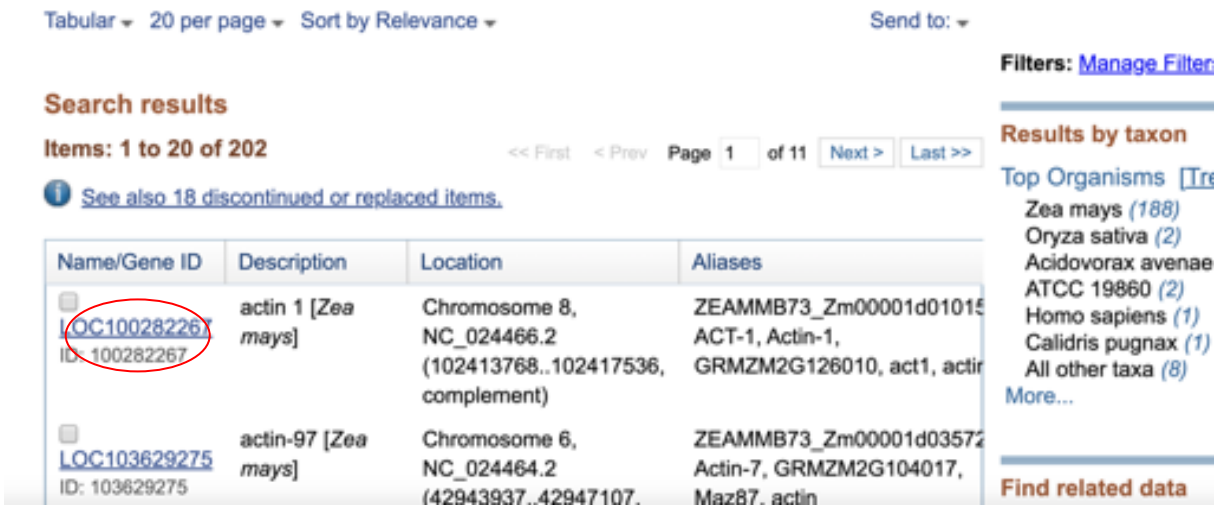
Part 1 – Choose a gene of interest using NCBI



Click on "Gene" – This will take you to NCBI's Gene Search



Next we'll look up the maize *ACTIN* gene. To do so type "maize actin" into the search bar and press "Search".

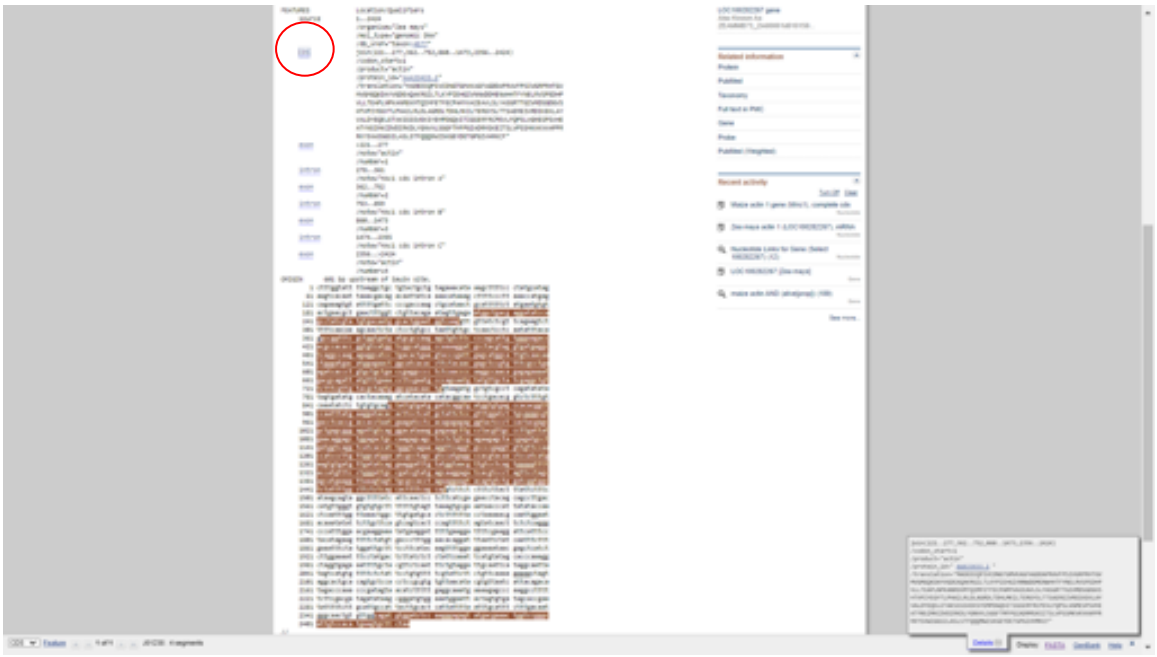


Click on "LOC100282267" under the Name/Gene ID.

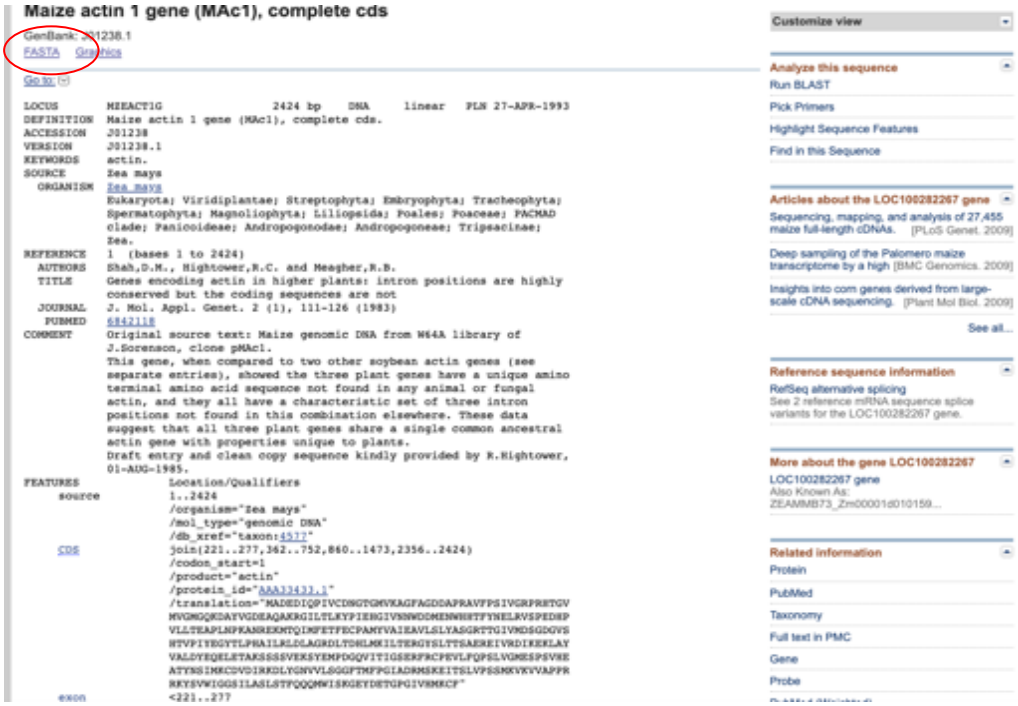
Click on “Nucleotide” under “Related information” to get the link to the nucleotide entry of maize *ACTIN-1*.

Scroll down and click on the “Maize actin 1 gene (MAc1), complete CDS.”

Under “features” you can see this gene contains 4 exons denoted by their position within the sequence. For example, exon 1 is from 221 to 277.



If you click on “CDS”, the **c**o**d**ing **s**equence will be highlighted in the genomic sequence below.



Maize actin 1 gene (MAc1), complete cds

GenBank: J01238.1
FASTA Graphics

Go to:

```
LOCUS       MIEACT1G                2424 bp    DNA     linear   PLN 27-APR-1993
DEFINITION  Maize actin 1 gene (MAc1), complete cds.
ACCESSION   J01238
VERSION     J01238.1
KEYWORDS    actin.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACMAD
            clade; Panicoideae; Andropogonodae; Andropogoneae; Tripsacinae;
            Zea.
REFERENCE   1  (bases 1 to 2424)
  AUTHORS   Shah,D.M., Hightower,R.C. and Meagher,R.B.
  TITLE     Genes encoding actin in higher plants: intron positions are highly
            conserved but the coding sequences are not
  JOURNAL   J. Mol. Appl. Genet. 2 (1), 111-126 (1983)
   PUBMED   6842118
COMMENT     Original source text: Maize genomic DNA from W64A library of
            J.Sorenson, clone pMAc1.
            This gene, when compared to two other soybean actin genes (see
            separate entries), showed the three plant genes have a unique amino
            terminal amino acid sequence not found in any animal or fungal
            actin, and they all have a characteristic set of three intron
            positions not found in this combination elsewhere. These data
            suggest that all three plant genes share a single common ancestral
            actin gene with properties unique to plants.
            Draft entry and clean copy sequence kindly provided by R.Hightower,
            01-AUG-1985.
FEATURES             Location/Qualifiers
     source          1..2424
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
     CDS             join(221..277,362..752,860..1473,2356..2424)
                     /codon_start=1
                     /product="actin"
                     /protein_id="AAA33433.1"
                     /translation="MADEDIQPIVCDNGTGMVKAGFAGDDAPRAVFPSIVGRPRETGV
                     MVGMGQKDAFVGDEAQAKRGILTLKYPIEHGIVSNWDDMENWHETFYNELRVSPEDHP
                     VLLTEAPLNPKANREKMTQIMFETFECPAMYVAIEAVLSLYASGRTTGIVMDSGDGVS
                     HTVPIYEGYTLPHAILRLDLAGRDLTDHLMKILTERGYSLTTSAEREIVRDIKEKLAY
                     VALDYEQELETAKSSSSVEKSYEMPDGQVITIGSERFRCPEVLFQPSLVGMESPSVHE
                     ATYNSIMKCDVDIRKDLYGNVVLSGGSTMFPGIADRMSKEITSLVPSSMKVKVVAPPR
                     RKYSVWIGGSILASLSTFQQMWISKGEYDETGPGIVHRKCF"
     exon            <221..277
```

Customize view

Analyze this sequence
Run BLAST
Pick Primers
Highlight Sequence Features
Find in this Sequence

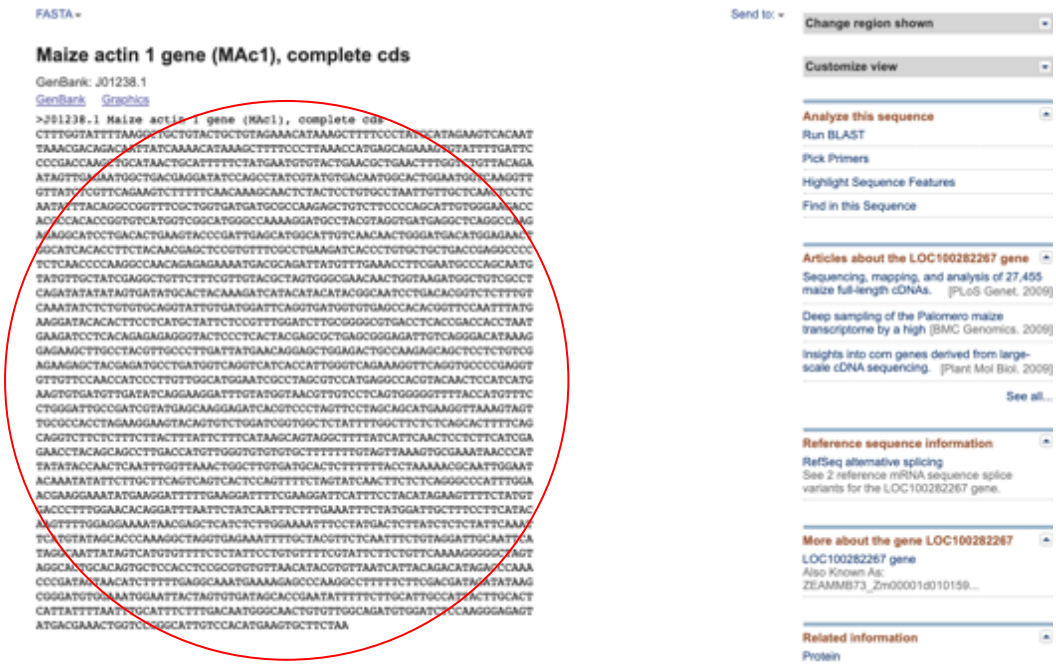
Articles about the LOC100282267 gene
Sequencing, mapping, and analysis of 27,455 maize full-length cDNAs. [PLoS Genet. 2009]
Deep sampling of the Palomero maize transcriptome by a high [BMC Genomics. 2009]
Insights into corn genes derived from large-scale cDNA sequencing. [Plant Mol Biol. 2009]
See all...

Reference sequence information
RefSeq alternative splicing
See 2 reference mRNA sequence splice variants for the LOC100282267 gene.

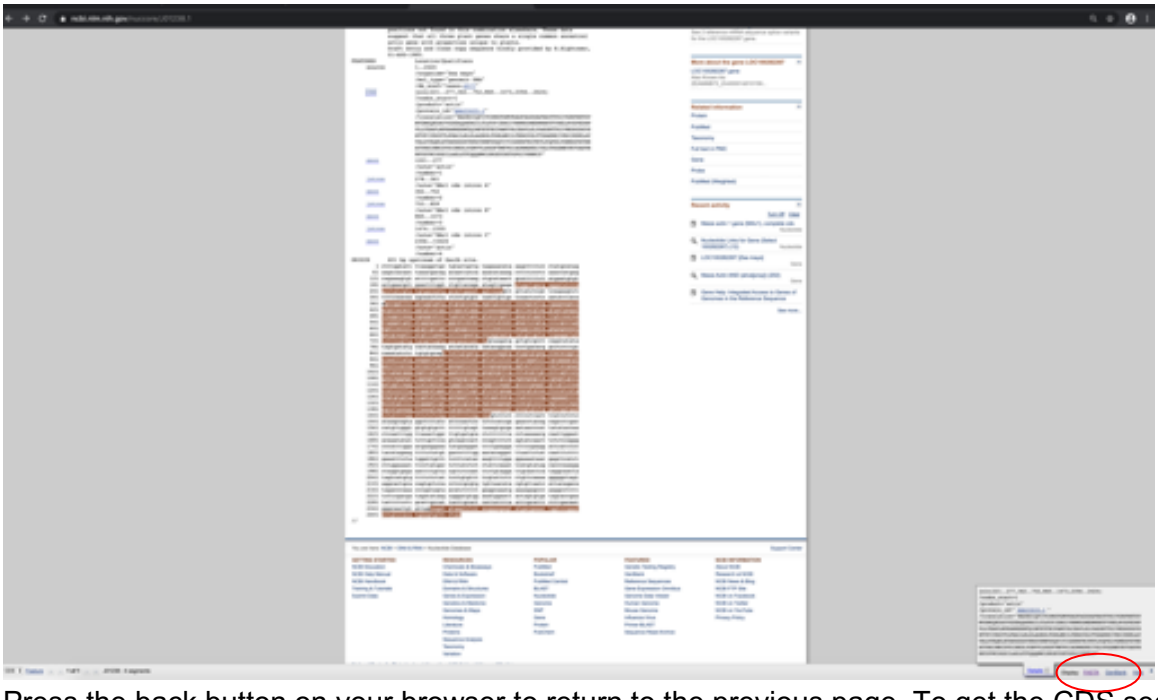
More about the gene LOC100282267
LOC100282267 gene
Also Known As:
ZEAMMB73_Zm00001d010159...

Related information
Protein
PubMed
Taxonomy
Full text in PMC
Gene
Probe

To get the genomic DNA sequence click the FASTA icon on the top left corner.



Create a new word document, copy the genomic DNA sequence and paste it. Give a name to the document and save it.



Press the back button on your browser to return to the previous page. To get the CDS sequence click on **“**CDS” to highlight the gene’s coding sequence and then under the box that pops up in the right lower corner, press “FASTA.”

FASTA

Maize actin 1 gene (MAc1), complete cds

GenBank: J01238.1

GenBank Graphics

```
>J01238.1:221-277,362-752,860-1473,2356-2424 Maize actin 1 gene (MAc1), complete
cds
ATGGCTGACGAGGATATCCAGCCTATCGTATGTGACAATGGCACTGGAATGGTCAAGGCCGGTTTCGCTG
GTGATGATGCGCCAAGAGCTGTCTTCCCCAGCATTGTGGGAAGACCACGCCACACCGGTGTCATGGTCGG
CATGGGCCAAAAGGATGCCTACGTAGGTGATGAGGCTCAGGCCAAGAGAGGCATCCTGACACTGAAGTAC
CCGATTGAGCATGGCATTGTCAACAACTGGGATGACATGGAGAACTGGCATCACACCTTCTACAACGAGC
TCCGTGTTTCGCCTGAAGATCACCCTGTGCTGCTGACCGAGGCCCCTCTCAACCCCAAGGCCAACAGAGA
GAAAATGACGCAGATTATGTTTGAAACCTTCGAATGCCCAGCAATGTATGTTGCTATCGAGGCTGTTCTT
TCGTTGTACGCTAGTGGGCGAACAACTGGTATTGTGATGGATTCAGGTGATGGTGTGAGCCACACGGTTC
CAATTTATGAAGGATACACACTTCCTCATGCTATTCTCCGTTTGGATCTTGCGGGGCGTGACCTCACCGA
CCACCTAATGAAGATCCTCACAGAGAGAGGGTACTCCCTCACTACGAGCGCTGAGCGGGAGATTGTCAGG
GACATAAAGGAGAAGCTTGCCTACGTTGCCCTTGATTATGAACAGGAGCTGGAGACTGCCAAGAGCAGCT
CCTCTGTCGAGAAGAGCTACGAGATGCCTGATGGTCAGGTCATCACCATTGGGTCAGAAAGGTTCAGGTG
CCCCGAGGTGTTGTTCCAACCATCCCTTGTTGGCATGGAATCGCCTAGCGTCCATGAGGCCACGTACAAC
TCCATCATGAAGTGTGATGTTGATATCAGGAAGGATTTGTATGGTAACGTTGTCCTCAGTGGGGGTTTTA
CCATGTTTCCTGGGATTGCCGATCGTATGAGCAAGGAGATCACGTCCCTAGTTCCTAGCAGCATGAAGGT
TAAAGTAGTTGCGCCACCTAGAAGGAAGTACAGTGTCTGGATCGGTGGCTCTATTTTGGCTTCTCTCAGC
ACTTTTCAGCAGCAGATGTGGATCTCCAAGGGAGAGTATGACGAAACTGGTCCGGGCATTGTCCACATGA
AGTGCTTCTAA
```

This should take you to a page with just the gene CDS. Copy the sequence and paste it on to the same document as your genomic sequence.

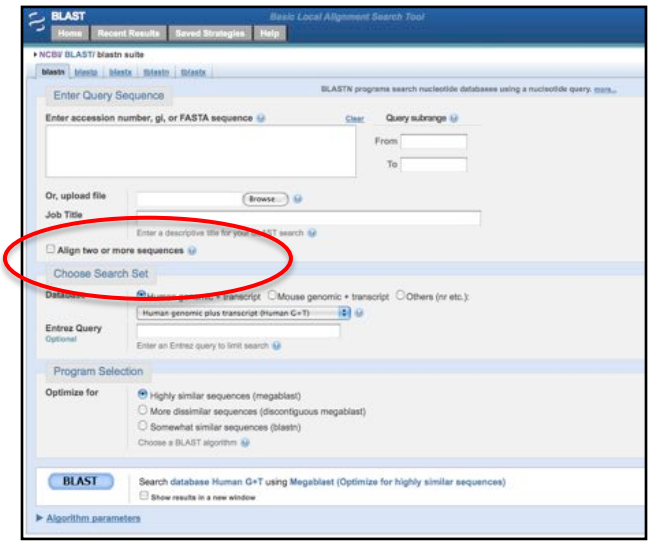
Next you will use BLAST to align the genomic sequence and CDS of maize actin.

Part 2 -- Compare the cDNA to the genomic DNA sequences.

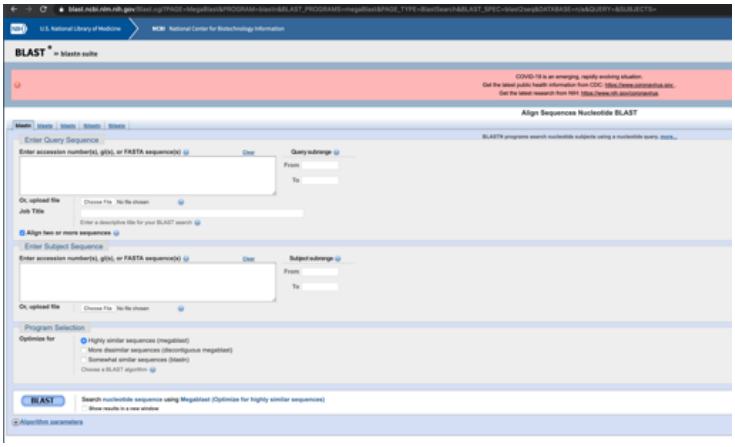
The CDS and genomic DNA sequences need to be aligned to determine where they are similar and where they differ. While this could be done by hand (also called manually) for short PCR sequences, it would be time consuming. Fortunately, a modified version of blastn called Blast2Sequences exists that performs alignments quickly and accurately. In this case the query (Genomic Actin DNA) will be compared to a single subject (CDS) instead of to a database.

1. Open a web browser, go to the Blast Website (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) and click on the 'nucleotide blast' link.

2. Check the 'Align two or more sequences' checkbox.



A second text box should appear:

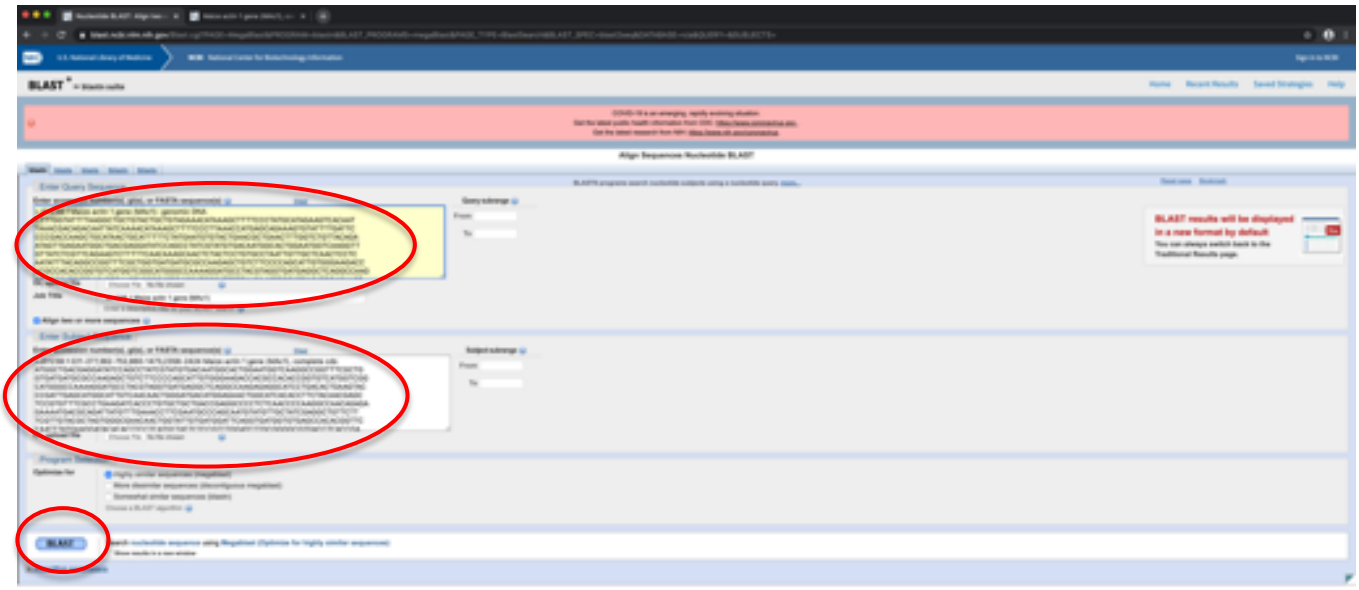


3. A) Enter the ‘Actin Genomic DNA sequence’ in the Query (top) textbox and B) the Subject (bottom) ‘Actin CDS sequence’ in the bottom text box. C) Click ‘Blast.’

A.

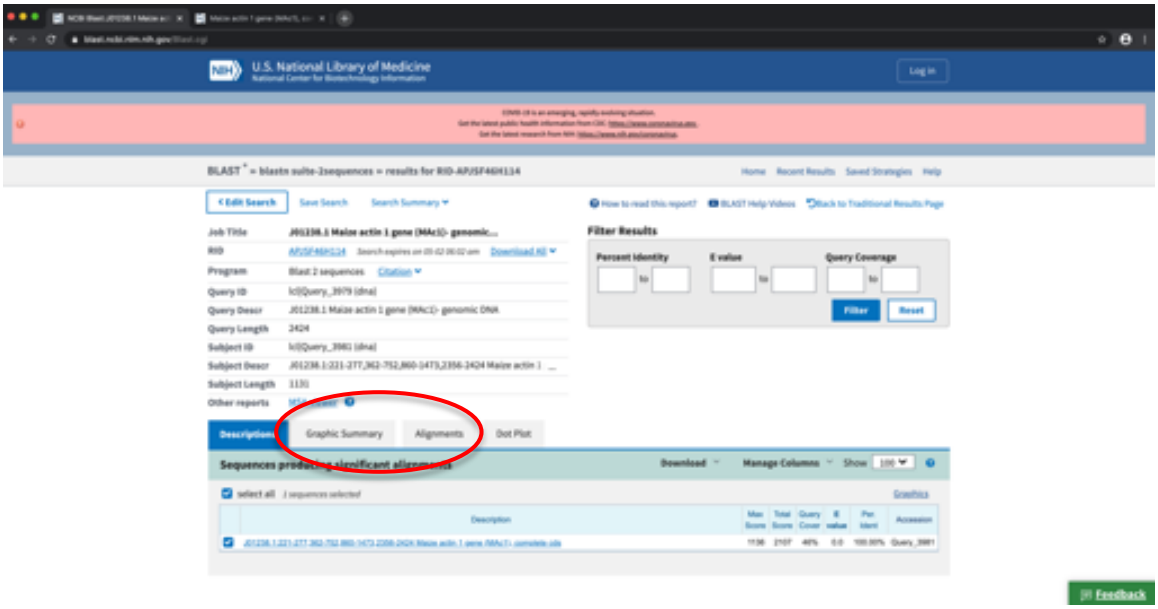
B.

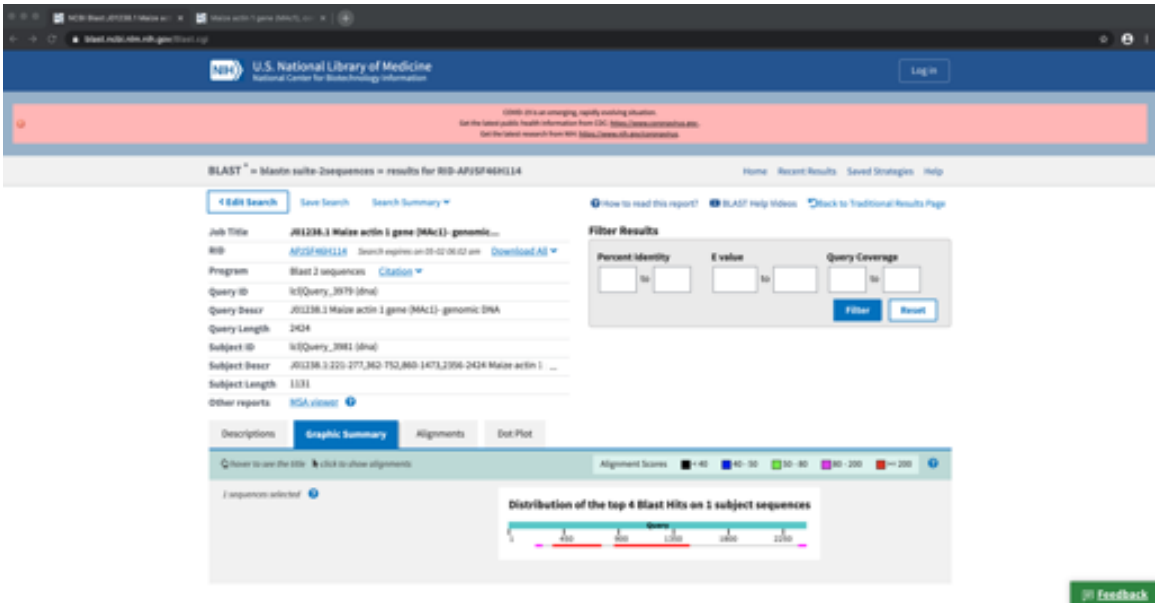
C.



Part 3 -- **Click on the graphic summary.**

The results are presented diagrammatically in the top half of the page with the query shown as a red, thick rectangle. Any similarity between the query and the subject is shown as thin rectangles below the query. The color of the rectangle indicates the hit score; the higher the score the better the hit.

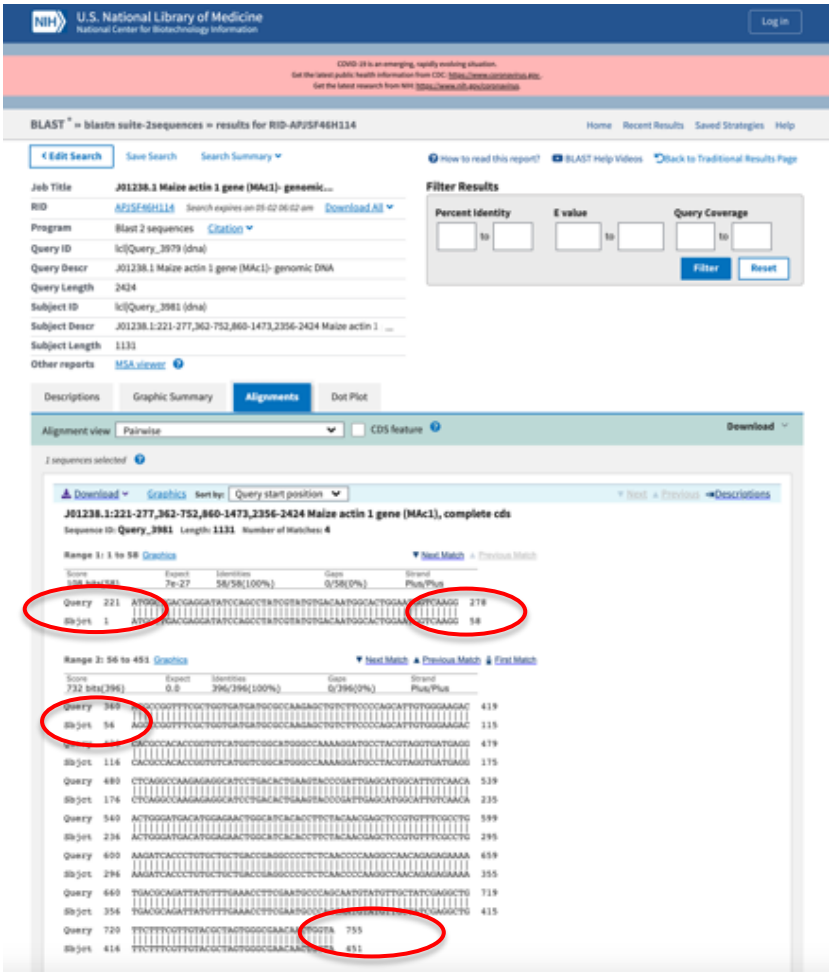




In this case there are four matching sequences between the query (genomic DNA) and the subject (CDS). Why are there four matching sequences and what does each gap represent?

At the bottom of the page are alignments of the four matching sequences nucleotide by nucleotide. When bases at the same position are the same a vertical line is placed between them.

The alignments can be put in order by query start position by going to the “Alignments” tab and then selecting “query start position” from the “Sort by” dropdown menu.

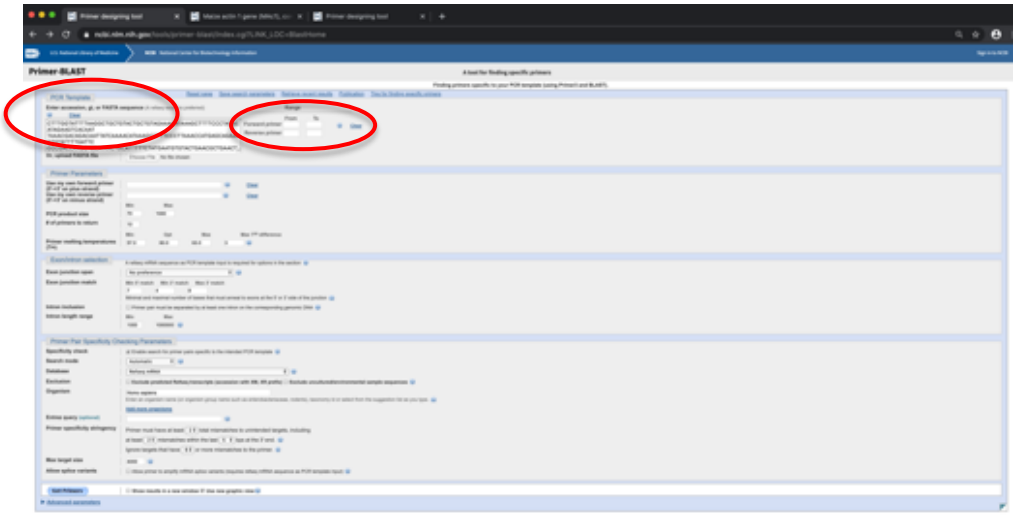

U.S. National Library of Medicine
National Center for Biotechnology Information
Log in
BLAST » blastn suite-2sequences » results for RID-AP1SF46H114
Home Recent Results Saved Strategies Help
Edit Search
Save Search
Search Summary
How to read this report?
BLAST Help Videos
Back to Traditional Results Page
Job Title J01238.1 Maize actin 1 gene (MAc1)- genomic...
RID AP1SF46H114
Program Blast 2 sequences
Query ID lcl|Query_3979 (dna)
Query Descr J01238.1 Maize actin 1 gene (MAc1)- genomic DNA
Query Length 2424
Subject ID lcl|Query_3981 (dna)
Subject Descr J01238.1:221-277,362-752,860-1473,2356-2424 Maize actin 1
Subject Length 1131
Other reports MSA viewer
Filter Results
Percent Identity
E value
Query Coverage
Filter
Reset
Descriptions
Graphic Summary
Alignments
Dot Plot
Alignment view Pairwise
CDS feature
Download
1 sequences selected
Download
Graphics
Sort by: Query start position
Descriptions
J01238.1:221-277,362-752,860-1473,2356-2424 Maize actin 1 gene (MAc1), complete cds
Sequence ID: Query_3981 Length: 1131 Number of Matches: 4
Range 1: 1 to 58 Graphics
Next Match
Previous Match
Score 7e-27 Identities 58/58(100%) Gaps 0/58(0%) Strand Plus/Plus
Range 3: 56 to 451 Graphics
Next Match
Previous Match
First Match
Score 732 bits(396) Expect 0.0 Identities 396/396(100%) Gaps 0/396(0%) Strand Plus/Plus
Range 3: 448 to 1062 Graphics
Next Match
Previous Match
First Match
Score 1136 bits(615) Expect 0.0 Identities 615/615(100%) Gaps 0/615(0%) Strand Plus/Plus
Range 4: 1062 to 1131 Graphics
Next Match
Previous Match
First Match
Score 130 bits(70) Expect 2e-33 Identities 70/70(100%) Gaps 0/70(0%) Strand Plus/Plus

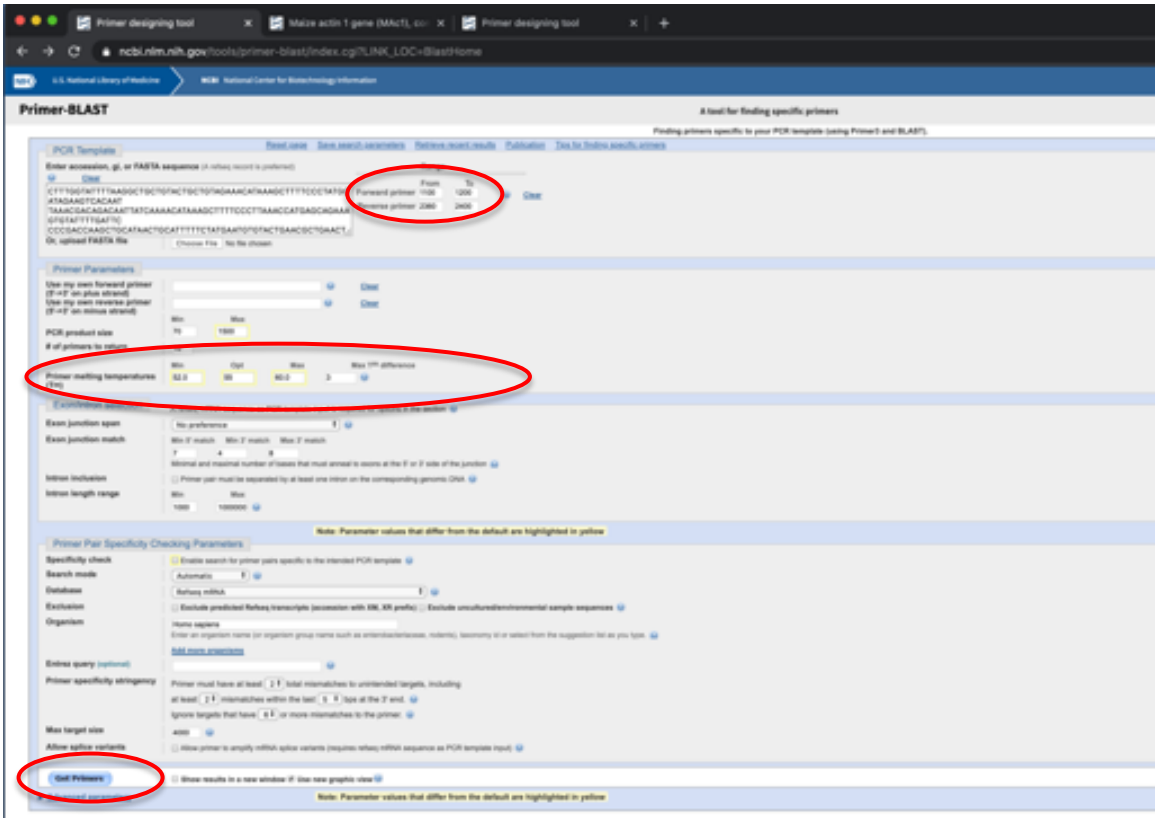
5. Draw the gene structure for the maize Actin sequence. Indicate the start and stop locations of the exons. You need this data for primer design in the following parts.

Part 4 – Primer design tool: Primer Blast

1. Open a web browser, go to the Blast Website (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) and click on the 'Primer- BLAST' link.

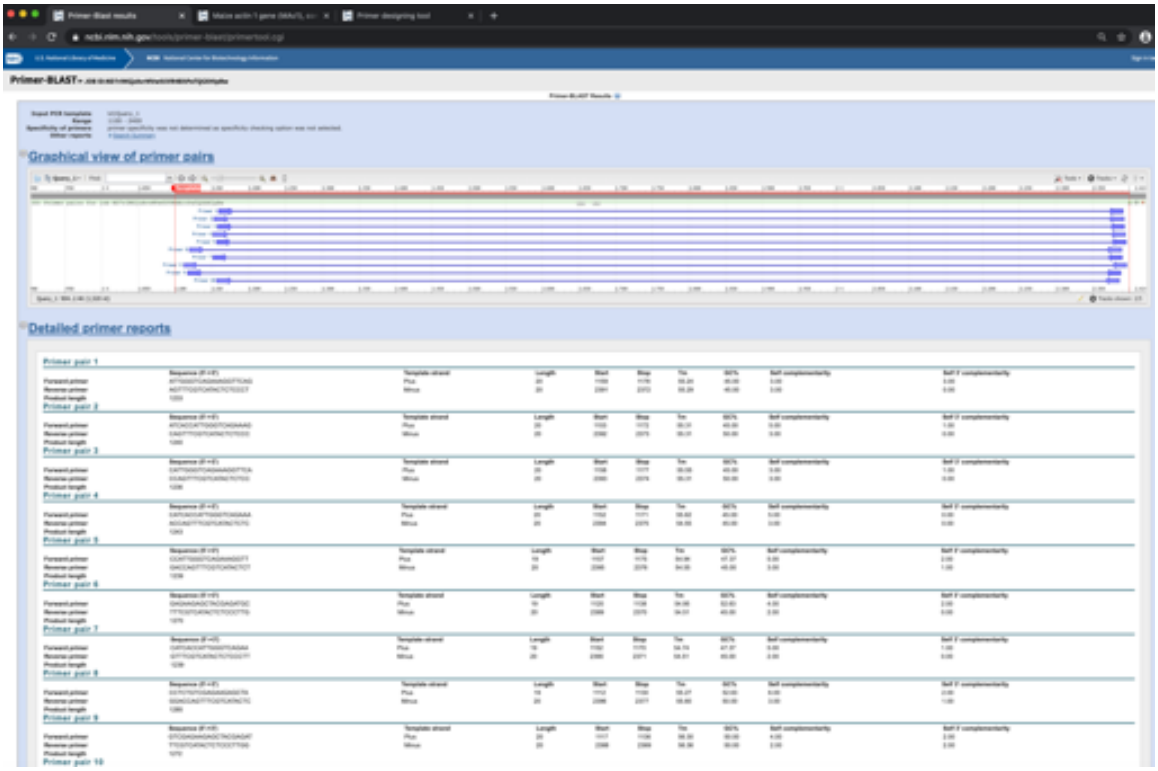
Copy the genomic DNA sequence and paste the sequence inside the PCR- Template box.
Put the nucleotide range for both forward and reverse primer based on the region you want to amplify that can distinguish between gDNA and CDS.
From the given example, the region to amplify is from exon 3 and 4 (nucleotide positions from 859 to 2324).



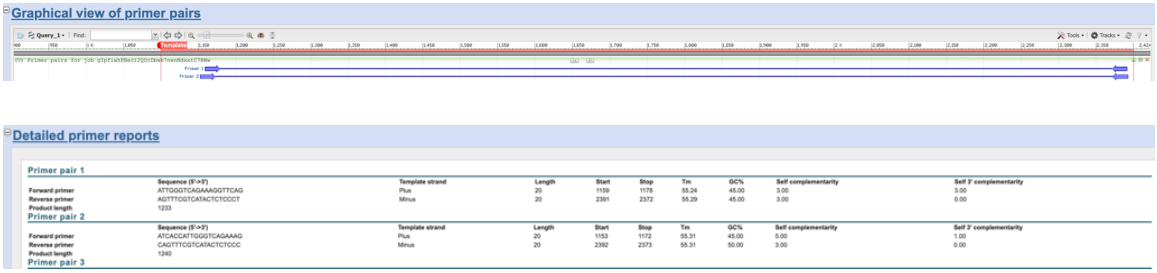


Click the get primer button.

When designing primers to amplify a large region, make sure to adjust the maximum “PCR Product size.”



Part 5 -- Primer position and amplicon sizes of gDNA and cDNA



>J01238.1 Maize actin 1 gene (MAc1), complete cds

CTTTGGTATTTTAAGGCTGCTGTACTGCTGTAGAAACATAAAGCTTTTCCCTATGCATAGAAGTCACAAT
TAAACGACAGACAATTATCAAAACATAAAGCTTTTCCCTTAAACCATGAGCAGAAAGTGTATTTTGATTC
CCCGACCAAGCTGCATAACTGCATTTTTCTATGAATGTGTACTGAACGCTGAACTTTGGTCTGTTACAGA
ATAGTTGAGAATGGCTGACGAGGATATCCAGCCTATCGTATGTGACAATGGCACTGGAATGGTCAAGGTT
GTTATCTCGTTCAGAAGTCTTTTTCAACAAAGCAACTCTACTCCTGTGCCTAATTGTTGCTCAACTCCTC
AATATTTACAGGCCGGTTTCGCTGGTGATGATGCGCCAAGAGCTGTCTTCCCCAGCATTGTGGGAAGACC
ACGCCACACCGGTGTCATGGTCGGCATGGGCCAAAAGGATGCCTACGTAGGTGATGAGGCTCAGGCCAAG
AGAGGCATCCTGACACTGAAGTACCCGATTGAGCATGGCATTGTCAACAACTGGGATGACATGGAGAACT
GGCATCACACCTTCTACAACGAGCTCCGTGTTTCGCCTGAAGATCACCCTGTGCTGCTGACCGAGGCCCC
TCTCAACCCCAAGGCCAACAGAGAGAAAATGACGCAGATTATGTTTGAAACCTTCGAATGCCCAGCAATG
TATGTTGCTATCGAGGCTGTTCTTTCGTTGTACGCTAGTGGGCGAACAACTGGTAAGATGGCTGTCGCCT
CAGATATATATAGTGATATGCACTACAAAGATCATACATACATACGGCAATCCTGACACGGTCTCTTTGT
CAAATATCTCTGTGTGCAGGTATTGTGATGGATTCAGGTGATGGTGTGAGCCACACGGTTCCAATTTATG
AAGGATACACACTTCCTCATGCTATTCTCCGTTTGGATCTTGCGGGGCGTGACCTCACCGACCACCTAAT
GAAGATCCTCACAGAGAGAGGGTACTCCCTCACTACGAGCGCTGAGCGGGAGATTGTCAGGGACATAAAG
GAGAAGCTTGCCTACGTTGCCCTTGATTATGAACAGGAGCTGGAGACTGCCAAGAGCAGCTCCTCTGTCG

Forward Primer (1159) ⟹

AGAAGAGCTACGAGATGCCTGATGGTCAGGTCATCACCATTGGGTCAGAAAGGTTCAGGTGCCCCGAGGT
GTTGTTCCAACCATCCCTTGTTGGCATGGAATCGCCTAGCGTCCATGAGGCCACGTACAACTCCATCATG
AAGTGTGATGTTGATATCAGGAAGGATTTGTATGGTAACGTTGTCCTCAGTGGGGGTTTTACCATGTTTC
CTGGGATTGCCGATCGTATGAGCAAGGAGATCACGTCCCTAGTTCCTAGCAGCATGAAGGTTAAAGTAGT
TGCGCCACCTAGAAGGAAGTACAGTGTCTGGATCGGTGGCTCTATTTTGGCTTCTCTCAGCACTTTTCAG
CAGGTCTTCTCTTTCTTACTTTATTCTTTCATAAGCAGTAGGCTTTTATCATTCAACTCCTCTTCATCGA
GAACCTACAGCAGCCTTGACCATGTTGGGTGTGTGTGCTTTTTTTGTAGTTAAAGTGCGAAATAACCCAT
TATATACCAACTCAATTTGGTTAAACTGGCTTGTGATGCACTCTTTTTTACCTAAAAACGCAATTGGAAT
ACAAATATATTCTTGCTTCAGTCAGTCACTCCAGTTTTCTAGTATCAACTTCTCTCAGGGCCCATTTGGA
ACGAAGGAAATATGAAGGATTTTTGAAGGATTTTCGAAGGATTCATTTCCTACATAGAAGTTTTCTATGT
GACCCTTTGGAACACAGGATTTAATTCTATCAATTTCTTTGAAATTTCTATGGATTGCTTTCCTTCATAC
AAGTTTTGGAGGAAAATAACGAGCTCATCTCTTGGAAAATTTCCTATGACTCTTATCTCTCTATTCAAAT
TCATGTATAGCACCCAAAGGCTAGGTGAGAAATTTTGCTACGTTCTCAATTTCTGTAGGATTGCAATTCA
TAGGCAATTATAGTCATGTGTTTTCTCTATTCCTGTGTTTTCGTATTCTTCTGTTCAAAAGGGGGCTAGT
AGGCACTGCACAGTGCTCCACCTCCGCGTGTGTTAACATACGTGTTAATCATTACAGACATAGACCCAAA
CCCGATAGTAACATCTTTTTGAGGCAAATGAAAAGAGCCCAAGGCCTTTTTCTTCGACGATAGATATAAG
CGGGATGTGGAAATGGAATTACTAGTGTGATAGCACCGAATATTTTTCTTGCATTGCCATTACTTGCACT

CATTATTTTAATTTGCATTTCTTTGACAATGGGCAACTGTGTTGGCAGATGTGGATCTCCAAGGGAGAGT

⟸ **Reverse primer (2391)**

ATGACGAAACTGGTCCGGGCATTGTCCACATGAAGTGCTTCTAA

Primer pair 1 are highlighted with the start nucleotide. The expected amplicon size is 1233 bp and 353 bp for genomic DNA and cDNA respectively.

Appendix 8: Primer design – *A. thaliana*

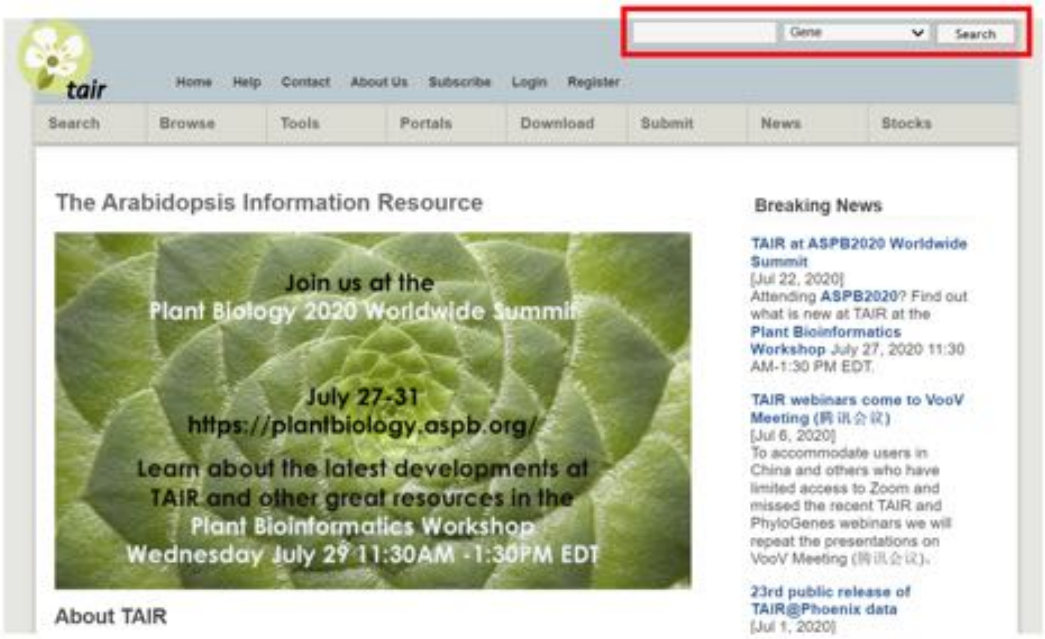
In previous experiments you learned about PCR to amplify the *ACTIN*-1 gene from genomic DNA and cDNA and also to study genome variation in maize. In addition to this you learned about how to analyze the DNA sequencing results using NCBI BLAST and to compare gene sequences through MUSCLE. In today's activity you are going to learn about primer design and design primers for an *A. thaliana* gene.

Activity

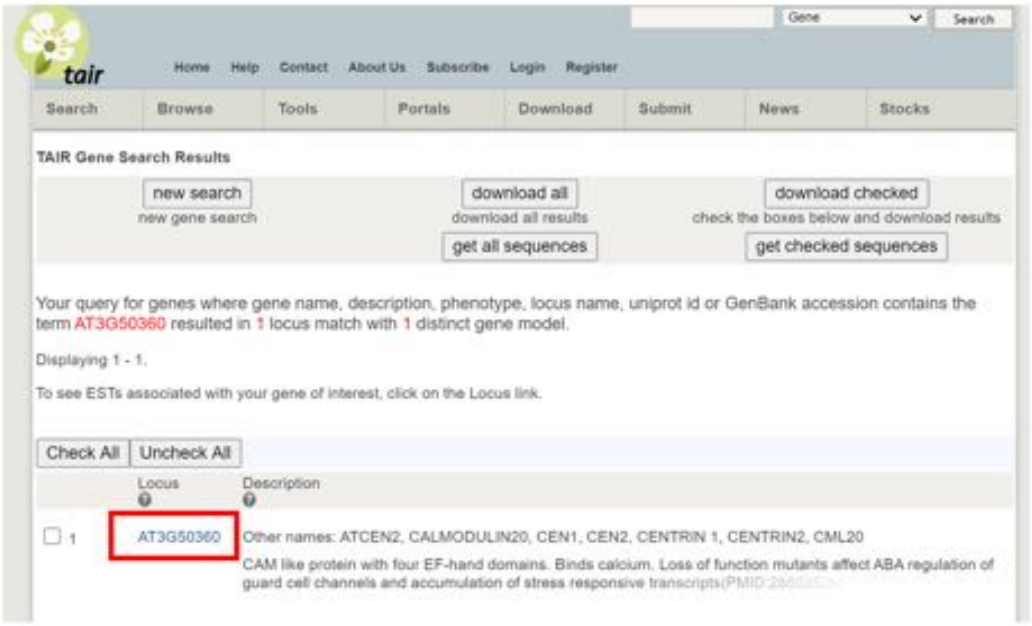
Use the following steps to design primers using NCBI's Primer Blast:

Part 1 – Look up your assigned gene using TAIR. Copy the gene sequence (both genomic and CDS).
Part 2 – BLAST Align the two sequence against GenBank database.
Part 3 – Check the graphic summary. Analyze the intron and exon positions.
Part 4 – Primer design tool: Primer Blast.
Part 5 – Write down the primer positions and amplicon sizes of gDNA and CDS.

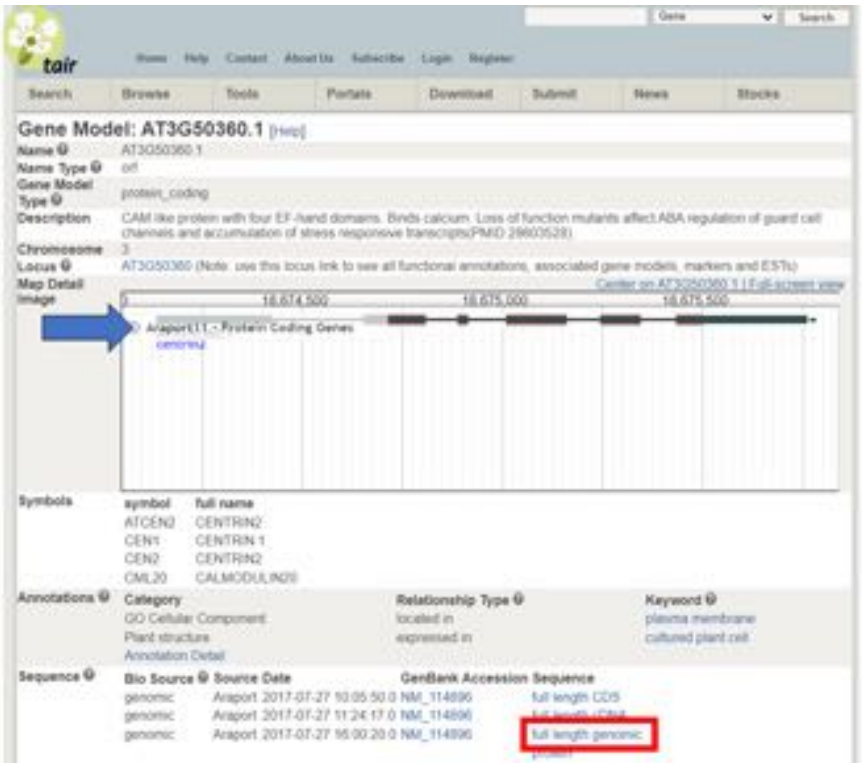
Part 1 – Looking up your gene



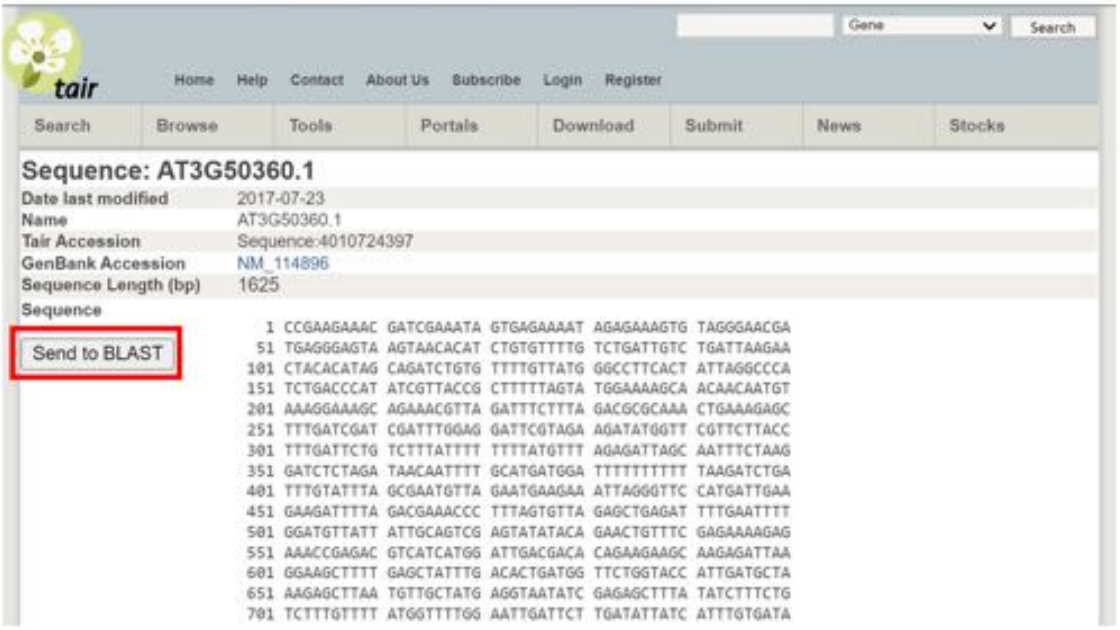
Go to arabidopsis.org. Enter your gene identifier into the search bar. Your gene identifier should resemble a series of letters and numbers starting with "AT". For example the gene identifier for CEN2 is AT3G50360. Press "Search"



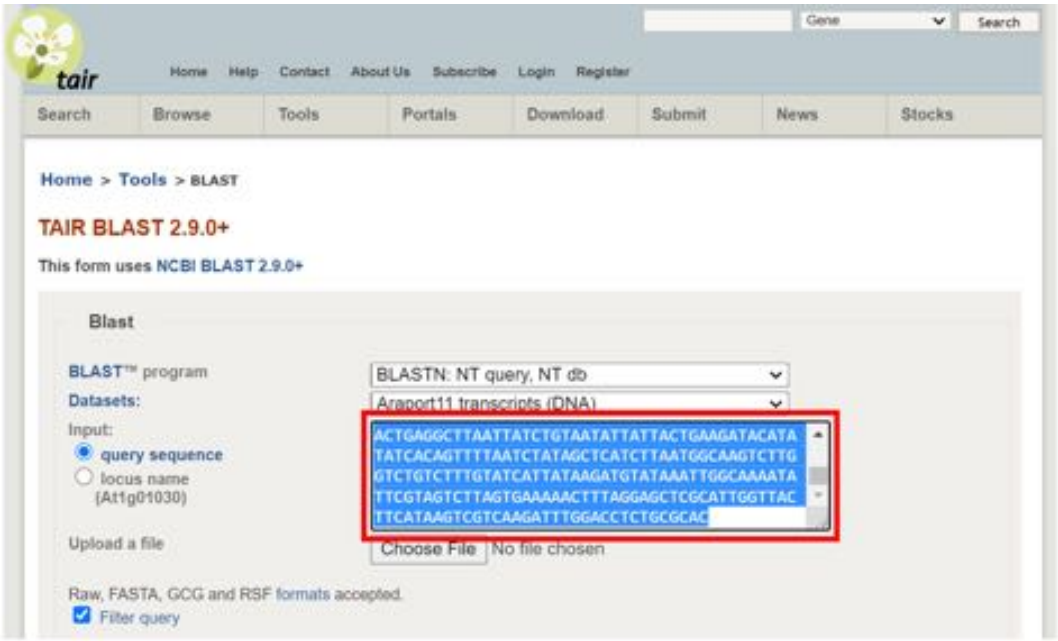
Click on your gene identifier under “Locus”. Make sure the numbers match because you may see multiple results show up.



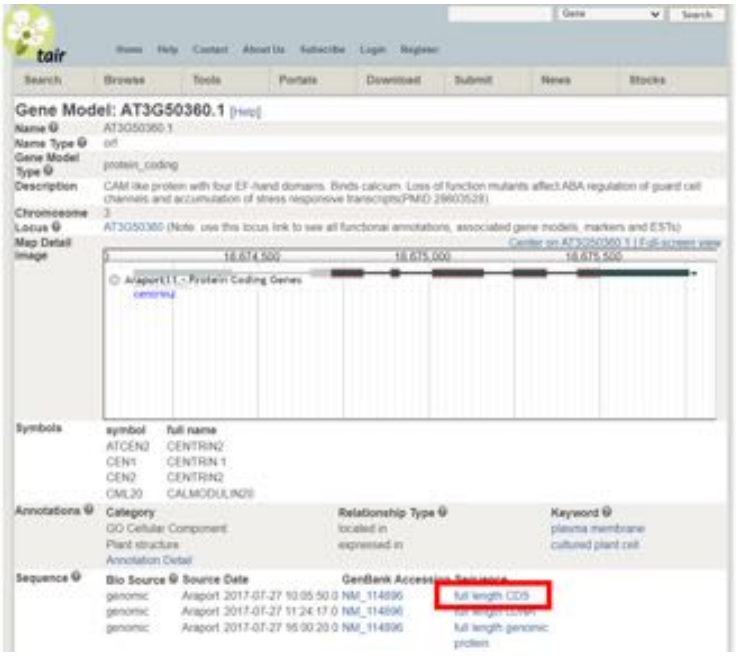
On this page you can find information on your gene, including a gene map showing its exons and introns (indicated by the blue arrow). For now, scroll down and click on “full length genomic” under the “Sequence” section.



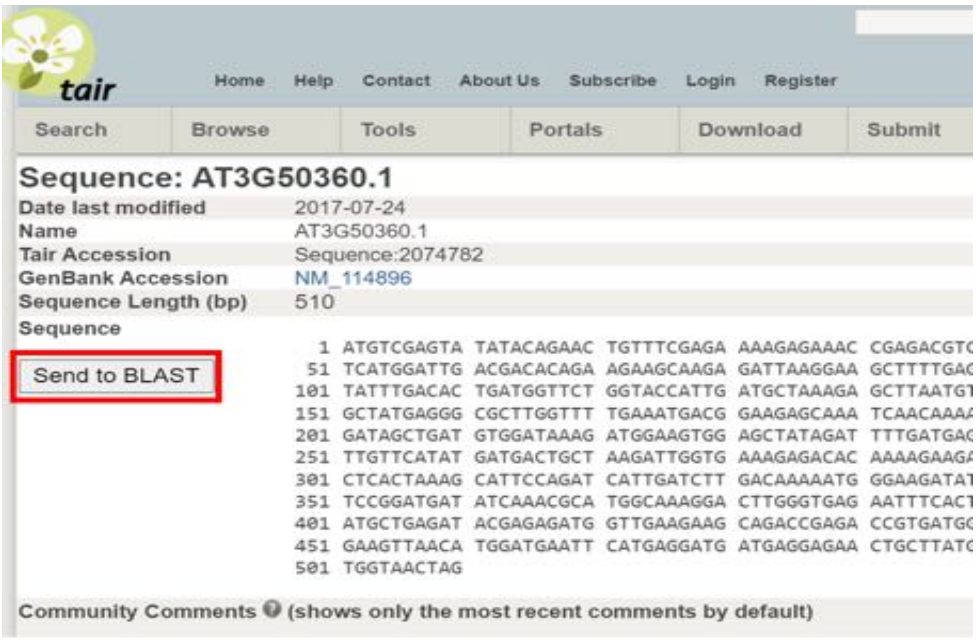
Here you'll find the genomic sequence of your gene. Press "Send to BLAST".



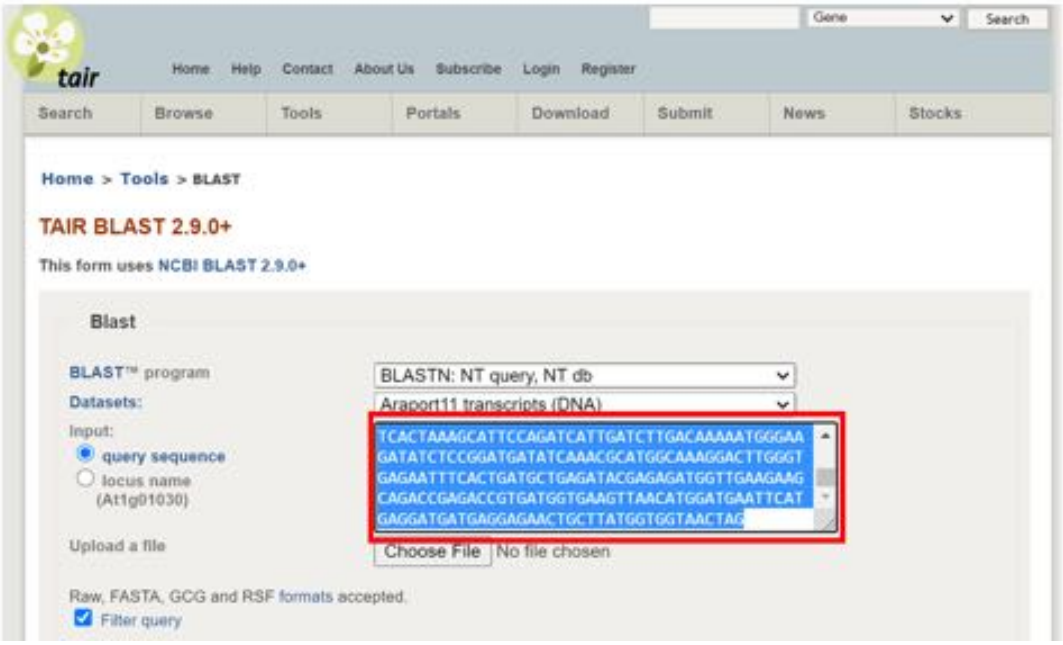
Create a new word document or google sheet, copy the genomic DNA sequence and paste it there. Make sure you scroll down to highlight the entire sequence. Give a name to the document and save it.



Hit the back button on your web browser to return to the gene summary page. Once there click on “full length CDS” in the “Sequence” section.



Click on “Send to BLAST”



Copy the sequence and paste it on to the same document you put your genomic sequence in.

Now you are ready to proceed to Part 2 of the primer design activity of “Primer Design – Maize Actin” where you will use BLAST to align your 2 sequences using the genomic DNA as the query and CDS as the subject. Then in Part 4 you’ll proceed to design primers for your *A. thaliana* gene.

Appendix 9: Lab Notebook Rubric

Categories	Points	Justification
Date	1	Actual start date of the experiment. 0.5 if date is incorrect. Date format does not matter. *Example: Jan 11, 2020 or 1/20/20.*
Purpose	4	Purpose should address a biological question related to the project in 1-3 sentences. *Example: "The purpose today is to design guide RNAs to target Cas9 to the A. thaliana gene AT1G20010 (TUB5)." Partial credit is allowed i.e. "to design guide RNAs for the gene TUB5" 2 pts; "to design guide RNAs" 1 pt.*
Materials	2	Materials and reagents can be found in the protocols provided to students. They should be included as a bulleted or numbered list. For remote learning, students should list the web resources used in the materials. For in lab learning, students should list any reagents, equipment, and web resources used.
Methods	3	Methods can be found in the protocols provided to students. Students should include the detailed methods in their lab notebook and include notes about any changes that deviated from the provided protocols.
Data and Results	6	This can consist of tables or images (gels or pictures) related to the specific experiment. If there are images, they must be annotated or provide a caption that explains what is in the image. *Example: If the student includes an agarose gel of their PCR results, they should label the lane with DNA ladder and what was loaded into each subsequent lane. Approximate band sizes should be indicated.*
Analysis and Conclusion	4	Some synthesis/interpretation of the experiment that addresses the purpose/biological question in the context of the project in a short paragraph. *Example: When choosing guide RNAs, students should justify why their guide RNAs are a good choice by explaining the specificity, location of the guide RNAs within the gene, etc.*
Total	**20**	

Appendix 10: Example notebook post - primer design

Date: 5/6/20

Purpose: This week we learned the basics of primer design and designed primers to amplify across the *A. thaliana* gene *MAP65-9* (AT5G62250). By doing so we can identify the coding sequence and genomic versions of our gene using PCR. These primers can also be used later to detect deletions within MAP65-9 created using CRISPR-Cas9.

Materials:

- NCBI - primer BLAST
- TAIR

Methods:

Part 1: Obtain the cDNA and gDNA for my gene

1. Go to arabidopsis.org. Enter your gene identifier into the search bar. Your gene identifier should resemble a series of letters and numbers starting with “AT”. Press “Search”
2. Click on your gene identifier under “Locus”. Make sure the numbers match because you may see multiple results show up.
3. On this page you can find all sorts of information on your gene, including a gene map showing its exons and introns (indicated by the blue arrow). For now, scroll down and click on “full length genomic” under the “Sequence” section.
4. Here you’ll find the genomic sequence of your gene. Press “Send to BLAST”.
5. Create a new word document or google sheet, copy the genomic DNA sequence and paste it. Make sure you scroll down to highlight the entire sequence. Give a name to the document and save it.
6. Hit the back button on your web browser to return to the gene summary page. Once there click on “full length CDS” in the “Sequence” section.
7. Click on “Send to BLAST”
8. Copy the sequence and paste it on to the same document you put your genomic sequence in.

Part 2: Make a gene diagram of my gene

1. Open a web browser, go to the Blast Website (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) and click on the ‘nucleotide blast’ link.
2. Check the ‘Align two or more sequences’ checkbox.
3. Enter your gene’s genomic DNA in the Query (top) textbox and B) the Subject (bottom) CDS sequence in the bottom text box. C) Click ‘Blast.’
4. Click on the graphic summary. The results are presented diagrammatically in the top half of the page with the query shown as a red, thick rectangle. Any similarity between the query and the subject is shown as thin rectangles below the query. The color of the rectangle indicates the hit score; the higher the score the better the hit.

5. At the bottom of the page are alignments of the four hits nucleotide by nucleotide. When bases at the same position are the same a vertical line is placed between them. The alignments can be put in order by query start position.
6. Draw the gene structure for your gene sequence (Do NOT copy and paste the “graphic summary” from BLAST. You should make your own using google draw, google slides, hand drawing, etc.). Indicate the start and stop locations of the exons. You need this data for primer design in the following parts.

Part 3: Design primer for my gene

1. Open a web browser, go to the Blast Website (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) and click on the ‘Primer- BLAST’ link.
2. Copy the genomic DNA sequence and paste the sequence inside the PCR- Template box.
3. Put the nucleotide range for both forward and reverse primer based on the region you want to amplify that can distinguish between gDNA and cDNA.
4. Click the get primer button.

Results:

>AT5G62250 CDS

ATGTCCAAATCTCAAATCGAATCAACTTGGTCATCTCTTCTACAAGAACTTGAGATTATATG
GAAAGAGGTTGGAGAAACTGAAACAGAAAGGGAGAAGATTTTGATTGAGATCGAAGAAGAA
TGCAGAGAAGTCTACAATCGAAAAATCGAAAAGGTTAAAGAAGAAAAGATTCGGATAAAAC
AAGAGATTGCTGATTCTGAAGCAAGAGTTATTGATATATGTTCTGTAATGGAAGAGCCACCA
ATTCTTGGAAGACACCATCAATCTGATCAACAAAGTGGTAATGGGAGAAGCTTGAAAGATG
AGTTAGTAAAGATTCTTCAGAAACTTGAAGAAATGGAAAAGAGAAAATCAGAGAGGAAGATT
CAGTTTATTCAAGTAATTGATGATATAAGATGTGTAAGAGAAGAGATTAATGGAGAATCTGA
TGATGAGACTTGTTCATCTGATTTTTCTGCTGATGAATCTGATTTATCTCTTAGAAAGCTTGA
AGAGTTACATAGAGAGCTTTACACACTTCAAGAACAAAAGAGGAACCGGGTGAAACAGATT
CAAGATAATATAAGAACTCTTGAATCGCTTTGTTCGGTTCTTGGTTTGAATTTTCGAGAAACT
GTTACCAAGATTCACCCAAGTTTAGTAGATACTGAAGGGTCAAGAAGTATAAGTAATGAAAC
ACTTGACAAGTTAGCTTCATCAGTACAACAATGGCATGAGACAAAGATTCAAAGAATGCAA
GAACTTCAAGATCTCGTGACAACGATGCTTGAGTTTTGGAATTTAATGGATACACCAGCAG
AAGAACAACAAAAGTTCATGGATGTATCATGTAATATAGCTGCTACTGTTTCTGAAATAACC
AAACCCAATAGTCTTTCTATAGATTTGCTAGAAGAGGTTAAAGCTGAGCTATGTCGGTTGGA
GGAGTTGAAGTGGAGCAAAATGAAAGAACTTGTTCTAAAGAAAAGGTCAGAGCTTGAAGAG
ATATGTAGAAGAACACACATTGTTCTTGAAGAAGAAGATATCGCGGTAGAGAATGTAATCAA
AGCCATTGAATCAGGAGATGTGAACCCTGAAAATATACTAGAACAGATCGAGTATCGAGCT
GGGAAAGTGAAAGAGGAAGCTCTAAGCAGAAAAGAGATTCTTGAGAAAGCTGATAAATGGT
TGAATGCTTGTGAGGAAGAGAATTGGCTTGAAGAGTATAATCAGGATGAAAACCGATACAA
CGCTGGAAAAGGATCTCATTTAATCCTCAAACGCGCAGAGAAAGCTCGTGCACTTGTTAAT
AAACTTCCAGCTATGGTTGAAGCATTAGCTTCCAAGATTACAATATGGGAATCAGAGAAAGA
ATATGAGTTTCTCTTTGATGGTAATCGCCTACTTTCAATGCTCGAAGAGTATACGGAACTCA

GAGAAGAGAAAGAACAAGAACGCCGCAGAAAGAGGGATCTGAAGAAACATCAAGGTCAAG
TGACATCAGAGCAAGACAAAGGAAGTGTTACGAAGCCCCAAAGCGCGAAAAAGGGTCTTA
AAGTGTCAACTAACAAGAGATTTGTCTCATCACCTCATACTCCTCAAACTGATTCGCCTCAC
TCAGCAAAATCGAATCAATCGTTTAGTACTCCTCTATCACGCCATGGCTGA

>AT5G62250 Genomic sequence

TTTTTTTTTGACTAGATTATTTTTATTTGATTCCTAAATTTTATCCGAATTTTTTTCAAGGAAAA
ACCAACGGCTGTGTCCTTCAATCCAGGTAAAGAATCTCGTATTTCTTCCAGGAAAATTCAAA
AGAAAAAAGTTTCAGCAATTAAAATGATTATCATTTAGAAAGAAAAAAATTCTTGTGTTTTCT
AAAAAGATTTTGTTATCGTTTGAGAAATTTTATGAAGTAACCTATTGAGGAAAAAAAACATCC
TCATAGGATTGAATAAGATAACGAATGTGGAAGAATTTTGAATTTATTTGCAGAGTTTGGAG
ATAAATAATGTCCAAATCTCAAATCGAATCAACTTGGTCATCTCTTCTACAAGAACTTGAGG
TTTGTGACGATTGCTTCAACTCCAAGAAAACGCTTTTTTTTTTCAATTACAAAAAGGCATTTC
CATTTTGATGTGGATTCTGTTTTCTGATGCAGATTATATGGAAAGAGGTTGGAGAAACTGAA
ACAGAAAGGGAGAAGATTTTGATTGAGATCGAAGAAGAATGCAGAGAAGTCTACAATCGAA
AAATCGAAAAGGTTAAAGAAGAAAAGATTCGGATAAAACAAGAGATTGCTGATTCTGAAGC
AAGAGTTATTGATATATGTTCTGTAATGGAAGAGCCACCAATTCTTGGAAGACACCATCAAT
CTGATCAACAAAGTGGTAATGGGAGAAGCTTGAAAGATGAGTTAGTAAAGATTCTTCAGAA
ACTTGAAGAAATGGAAAAGAGAAAATCAGAGAGGAAGATTCAGTTTATTCAAGTAATTGATG
ATATAAGATGTGTAAGAGAAGAGATTAATGGAGAATCTGATGATGAGACTTGTTCATCTGAT
TTTTCTGCTGATGAATCTGATTTATCTCTTAGAAAGCTTGAAGAGTTACATAGAGAGCTTTAC
ACACTTCAAGAACAAAAGGTTTTTTCTAAACTTTACCTTAAATTTGATTTCATTTCTTGATCAA
AAAGTTAACTAATACTTTACTGTTTATTTTGGCTCAGAGGAACCGGGTGAAACAGATTCAAG
ATAATATAAGAACTCTTGAATCGCTTTGTTCGGTTCTTGGTTTGAATTTTCGAGAAACTGTTA
CCAAGATTCACCCAAGTTTAGTAGATACTGAAGGGTCAAGAAGTATAAGTAATGAAACACTT
GACAAGTTAGCTTCATCAGTACAACAATGGCATGAGACAAAGATTCAAAGAATGCAAGAAG
TGAGTTAGAATGTTTTTAAACATTTTACCATCTTAAAAACTCTGCTCTTTGTTTCTTATATATA
TATGCTATCTTATGGATTACAGCTTCAAGATCTCGTGACAACGATGCTTGAGTTTTGGAATT
TAATGGATACACCAGCAGAAGAACAACAAAAGTTCATGGATGTATCATGTAATATAGCTGCT
ACTGTTTCTGAAATAACCAAACCCAATAGTCTTTCTATAGATTTGCTAGAAGAGGTAACTTC
ACAACAACTTTACATCAAATCTTTTTGTAAACTTACTGAGACTATTTTCGCCGATGGATTGCT
TAATGTTTTAAATAGGTTAAAGCTGAGCTATGTCGGTTGGAGGAGTTGAAGTGGAGCAAAA
TGAAAGAACTTGTTCTAAAGAAAAGGTCAGAGCTTGAAGAGATATGTAGAAGAACACACATT
GTTCTTGAAGAAGAAGATATCGCGGTAGAGAATGTAATCAAAGCCATTGAATCAGGAGATG
TGAACCCTGAAAATATACTAGAACAGATCGAGTATCGAGCTGGGAAAGTGAAAGAGGAAGC
TCTAAGCAGAAAAGAGATTCTTGAGAAAGCTGATAAATGGTTGAATGCTTGTGAGGAAGAG
AATTGGCTTGAAGAGTATAATCAGGTTTGTCTTTTTTTTATTCATTAAAGATTCAGTTCTGAC
ATTTTGCCGTTACTTAATGTTTTTTTGTATAGGATGAAAACCGATACAACGCTGGAAAAGGA
TCTCATTTAATCCTCAAACGCGCAGAGAAAGCTCGTGCACTTGTTAATAAACTTCCAGGTTT
GATATAAATGATCCCAACGTACACATTGTCATATTGATGAATGCTTATCTTACTAAAAATCTT
TTTCACTCTTCTTAGCTATGGTTGAAGCATTAGCTTCCAAGATTACAATATGGGAATCAGAG
AAAGAATATGAGTTTCTCTTTGATGGTGTATGATCTTATCTTCTTCTCTCAAAAAAAAACCCC
TCTTCTTCTCTTTTCTCCTTTTTCTCTTTTATATACTTCATTTGATCTCTTAATCTGATTTTCTT
CATTCAGAATCGCCTACTTTCAATGCTCGAAGAGTATACGGAACTCAGAGAAGAGAAAGAA

CAAGAACGCCGCAGAAAGAGGGTATTAGTTTTACATGCAAATCATATTTTTAATCAAATTTT
CTATTGTTCCAAGAATCACCAAGAACAATGTTTGTAGGATCTGAAGAAACATCAAGGTCAAG
TGACATCAGAGCAAGACAAAGGAAGTGTTACGAAGCCCCAAAGCGCGAAAAAGGGTCTTA
AAGTGTCAACTAACAAGAGATTTGTCTCATCACCTCATACTCCTCAAACTGATTCGCCTCAC
TCAGCAAAATCGAATCAATCGTTTAGTACTCCTCTATCACGCCATGGCTGAAAAAGATAGAG
CTAAGATTTCAAAAACAGTCTTTTTATAGATCAACTAGTAATAGTTTCTTTCTTATATACACAC
TGTCACACATGTTATGTACAAGTATAGTTCTAATGCAACCATAGAAACAAGACATGCTAATT
CAAATAGAGGGTCTCGAAACTGATGATTCTTGATCATTTGTAATCTGTCAAACTCAGGATCC
AAGTGTGTACTATACACTGATTGTTTTGCGGAAAGATTAGTTTGACTTTGACCAACCAATAT
ACTT

Map65-9 Gene Diagram with primer positions:

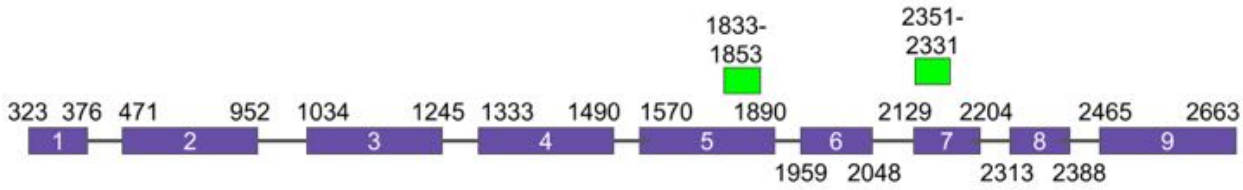


Figure 1. Gene diagram of Map65-9 showing its exons and introns in purple. The position of the primers I designed are in green.

Primer Design Conditions:

I want to amplify across introns 5 and 6 since my introns are short. The conditions I used for my primer search are as follows:

Forward primer: 1809 to 1880

Reverse primer: 2318 to 2380

Temperatures:

Min: 52°C

Optimal: 55°C

Max: 60°C

Here are the primers I picked:

Forward primer:

Sequence	Length	Start	Stop	Tm °C	GC %	self complementarity	3' complementarity
GCTGATAAATGGTTGAATGCT	21	1833	1853	54.87	38.1	3	0

Reverse primer sequence:

Sequence	Length	Start	Stop	Tm °C	GC %	self complementarity	3' complementarity
GAGTTCCGTATACTCTTCGAG	21	2351	2331	55.13	47.62	6	3

Genomic product length: 519 bp

cDNA product length: genomic product length 519 - introns in region 150 = 369 bp

Conclusion:

In this experiment I designed primers to distinguish the MAP65-9 cDNA from genomic DNA. I designed my primers to amplify across introns 5 and 6, so there will be 150bp difference between the cDNA (369 bp) and genomic DNA (519bp). These PCR products should be big enough to easily see on a gel.

Appendix 11: Quiz 1

Quiz 1

Instructions

All answers must be typed. Use the Word Formula editor for calculations. Please put answers in BLACK. Upload the document to the assignment for grading.

1. (5 points) Use BLAST to determine the following information.

```
> Sequence 1
AGCAAGATGGATGTAGGCAGCAAAGAAGTCCTGATGGAGAGCCCGCCGGTGAGTGTGCATGTGGAAGGGA
GTGTGTGTGTGTGTTTGTGTGTCTGTGTGTGCAGGTGAGTGGTGGGTCCTGCCTCCTCCAACCTCTCAGC
CTGTTTCCTCATCCAAATTCATTCGCTCAACTAACCTAGGACTGTGCTAAATTGGGAGAGGGACCCTAGC
CTAATGGTAGGTTGGTCATAGTAGACTGGAGAGATCGGTGTGGCAGCCACTAATTCCCATACGAGGGGCT
GTGAGCCACCACGGAGGTCAGTGGAAAGGGCTGCTAAGAGCATTGGATAAGGAGCCACTACTTAGGAGGT
CGCAGAAAGCTTCACAGAGAAAATAAGCTGGCCCCAGAAGACCAGGTCAGAGCTCACTAGGTGGAGAAGG
AGGAAGGCATTATGGGAAGAAGGGATCATATGCATAAAGGCATGGGGGTGTGATACAGCATGGGTTTGGG
GGGGACAGAGAGGGACTGGGTATGGCTGGGATGCCAGGAAATAGCAGGAGATGAGACTGGTGAAACAGGC
TCCCTTGGGTCTCTCCAAGCTAGACTTCTCTAAGCATCAATGGGTTTGTTGAAATCCAGGCCCCTTCCTC
CAGAAGCCTTCTCTGATCTCCTCAGCCCCCTGCTGCATGCCCCTAGCATTGGTTGTTCAGCCCTGGGAGA
GTGTGGAGGGTTTCCAGATCGTATAAGAAGAAGACGGAGCATCTTCACAACCTCACACCTGTCTCCCTGC
CTGCCTCACCCCCTTCAGGACTACTCAGCAGTCCCAGGGGGCCGGCTCCGCATCCCCTGCTGTCCTGTGA
ACCTCAAACGCCTTCTTGTCGTGGTCGTGGTGGTGGTTCTTGTCGTCGTGGTGATTGTAGGGGCCCTGCT
CATGGGTCTTCACATGAGCCAGAAACATACTGAGATGGTGAGCAGGCCTGGGATGGGGTGGACCGGACAG
GTAGCATTTTGCCCGAGGGGAAGGAGGGAAAGTGAGGGGCACCTAGTGGAAGGACAGGTGGGAGCAGGGT
GGTAATGCTGGCCTGAGCTTTTTTACCAGGCCCCTGTCCCAGGGTGGGATAGCCTGTTCCCTCCAGCACC
TGGTTCCCCTCGGCCTCTCTAAGCACCCAAAGGAAGAAAGAATTGCATTTGAGTAAAGAGGACTGAGGGT
GGGCATGAGGCAAAGGAATGAGTGGAAGAGAGGAATGCCAGGAAGCACTACGCCCTCCCCCAGGTCCTAG
AGATGAGCCTCGCAGGGCCAGAAGCCCAGCAACGCTTGGCCCTGAGTGAGCATGTGGGAACCACTGCCAC
CTTCTCCATTGGCTCTAGTGGCAATGTGGTCTATGACTACCAGCGGGTGAGTATGACTCCTGACCCTAGG
GCTGGGGGACCAATAACAAGACAATGCTAGTCCTGCCCAGCTATGCTGCTTCTTACATGTCATCCATCTC
TCCCCACCCCCAGCCCTTTTACCTCCCCTGGGGGGGTGCCTGAGCCCCAGATTCCAATATGGCTCCCCCC
AGGGCCCATCCTAGAGGGAGGTGGCCAAGGCCCAGGGTTAGAGAGTAAACTAGGCAGTGCCTTCCTTAAC
CCACTCCATCGTTCCCGTGCACAGCTCCTGATTGCCTACAAGCCAGCCCCGGGAACCTGCTGCTATGTCA
TGAAGATGTCTCCGCAGAGTATGCCGAGTCTTGAGGCTCTCACCAAAAAATTCCAGAACTTCCAGGTGAG
TGTGCATGTAAGGAGTGGGTTATCTCCCTGCCAGGGCTGCTGGGAGGAGCATTTGCAATGACGACTCCTT
GTCACCTGTAAGGCCTCATGGACTGCCAGGTGAACGCCCTGCCTAGACCCATGCACCAAACCTTTCCTAC
CTGTTCCTCATCCCTACCTACAGCCATGCCACTGGCCTCAGCTGAGCCCCACACCCCCTGGGGGCTTCTG
ACTCCAGCACAGCCCCTTCTTTACACAGAGAGATTACCAGGCACACCTGAAGTCCCATAATAAGGGCCGC
AGGTGAGACAGAAACTCAGTTCCCAGGCTCTCAAATCAGATGCTCTTTCCAGGCCAAGCCCTCGACGCCT
ACCTCTAAGCTGGGCCAGGAGGAGGGCTGTGTCGCTGGCTCAGCACCCTCCGGGGACCTGGCCTTCCTGG
GCAGCACCATGAGCACCCTGTGTGGCGAAGTGCCCCTCTTGTACATCTAGGAAACATCAGGTGAGCAGGT
GTGACCTCCAGGGCCCTGGAGCAAGCAAGTGACAGTGAGTGACCACTGGTCCCTCCCCCTCCTCTTTCTC
AGTGTGAAAGGAGCCTGAATTCAGACTGAGGCACAAAGGCTTTTTGGGACACTAGTCTGCCATCTTCTGT
CTGCTACCTTTCCATTAAAGTTGTTATTCTTTGCCCCAACAACTCATCTCCGAATTTATTGGCCTGTGGC
```

AGGGAGAGCAGAGTGAACTTGGGTTCAGTAACACCCCCAGTCCCTCTTCTTCACATAAACTGTGTTCCTT
TGCTTCACAGGGCCTCAGGAAGCCCCAAGAGGACAGCAAAGATCCAGGAGCAAAGAGTCTTGTGCAGACT
CACAGGAAGCCGCTTCTGGGACACCACGGGGACTGGCCCTGGAGAAATGGGAGCTGTGGGGAGAGGTGGGCAGAGGA
GAAGCAGCTGTTAGGGGCCCGGGGGCTTCTACCACCAAAGAAT

A. What is the name of the gene?
B. What is the next best BLAST hit?
C. Is the next best BLAST hit an Ortholog or a Paralog?
D. What is the length for the gene?
E. Is the sequence Genomic DNA or mRNA?

2. (2 points) Align the previous sequence with sequence below and answer the following questions.

> Sequence 2
ATGGATGTAGGCAGCAAAGAAGTCCTGATGGAGAGCCCGCCGGACTACTCAGCAGTCCCAGGGGGCCGGC
TCCGCATCCCCTGCTGTCCTGTGAACCTCAAACGCCTTCTTGTCGTGGTCGTGGTGGTGGTTCTTGTCGT
CGTGGTGATTGTAGGGGCCCTGCTCATGGGTCTTCACATGAGCCAGAAACATACTGAGATGGTCCTAGAG
ATGAGCCTCGCAGGGCCAGAAGCCCAGCAACGCTTGGCCCTGAGTGAGCATGTGGGAACCACTGCCACCT
TCTCCATTGGCTCTAGTGGCAATGTGGTCTATGACTACCAGCGGCTCCTGATTGCCTACAAGCCAGCCCC
GGGAACCTGCTGCTATGTCATGAAGATGTCTCCGCAGAGTATGCCGAGTCTTGAGGCTCTCACCAAAAAA
TTCCAGAACTTCCAGGCCAAGCCCTCGACGCCTACCTCTAAGCTGGGCCAGGAGGAGGGCTGTGTCGCTG
GCTCAGCACCCTCCGGGGACCTGGCCTTCCTGGGCAGCACCATGAGCACCCTGTGTGGCGAAGTGCCCCT
CTTGTACATCTAG

A. How many exons?
B. How many introns?

3. (4 points) Use Primer Blast to design primers which will work in both sequences.

A. Paste screenshot of Primer blast output for sequence 1.
B. How big is the amplicon for sequence 1?
C. Paste screenshot of Primer blast output for sequence 2.
D. How big is the amplicon for sequence 2?

4. (5 points) What is the difference between *tan1* mutants in maize and *tan1* mutants in Arabidopsis? Why must a *tan1, air9* double mutants be used in Arabidopsis? Be sure to define synthetic mutant.

5. (3 points) List three advantages of using *Arabidopsis thaliana* as a model organism.

6. (5 points) Define Preprophase Band (PPB) and its role in cell division in plants. How does TAN1 play into that role?

Appendix 12: Quiz 1 Answer Key

1. Use BLAST to determine the following information.
   A. Sus scrofa SP-C gene
   B. Bos taurus surfactant protein C
   C. Ortholog
   D. 2710 base pairs
   E. Genomic DNA

2. Align the previous sequence with sequence below and answer the following questions.

   A. 5 exons
   B. 4 introns

3. Use primer BLAST to design primers which will work in both sequences.
   Answers vary, below is an example of an acceptable answer:
   A.

   B. 1268 base pairs
   C.

   D. 367 base pairs

4. What is the difference between *tan1* mutants in maize and *tan1* mutants in Arabidopsis? Why must *tan1 air9* double mutants be used in Arabidopsis? Be sure to define synthetic mutant.

tan1 mutants in maize have a strong mutant phenotype with cell division plane orientation defects and short stature. The *tan1* mutant in Arabidopsis has no mutant phenotype and resembles wild type plants. *tan1 air9* Arabidopsis double mutants have a strong mutant

phenotype with division plane orientation defects and short stature, so we can study division plane orientation in the double mutant. They have a synthetic mutant phenotype where the double mutant phenotype is more severe than the single mutant phenotypes combined. The *air9* single mutant in Arabidopsis resembles wild type plants as well.

5. List three advantages of using *Arabidopsis thaliana* as a model organism.

1. They are small plants that are easy to grow and care for compared to larger plants that require a field or greenhouse.
2. They have a short generation time, so it doesn't take them too long to grow and produce seeds. This makes it easier to study genetics and make transgenic Arabidopsis.
3. *A. thaliana* is a well-studied plant and there are many databases and tools available to work on them.

6. Define Preprophase Band (PPB) and its role in cell division in plants. How does TAN1 play into that role?

The PPB is a ring of microtubules and proteins that forms around the premitotic nucleus in plants cells, and it marks the future location of the new cell wall after the cell divides. TAN1 associates with the PPB and remains at the division site after the PPB disassembles when the cell enters mitosis. TAN1 plays a role in the molecular memory of the PPB and helps guide the phragmoplast to construct the new cell wall at the site originally marked by the PPB.

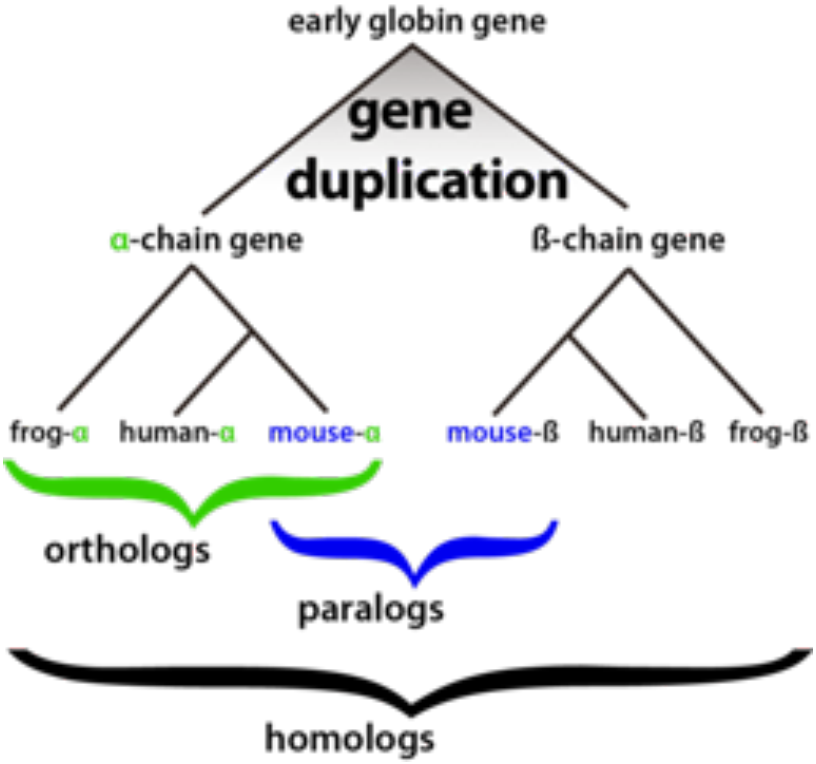
Appendix 13: Phylogeny and using PLAZA

Background:

Just like how organisms evolve over time, gene do too! As a result, we see similar but distinct genes in different species (orthologs) and sometimes gene duplications within a species (paralogs). Orthologs and paralogs are both types of homologs. Phylogenetic trees are useful for examining the evolutionary history of a target and its homologous genes.

It is common for genes to undergo duplication. Sometimes these duplication events are local (small portions of a chromosome), and sometimes this is due to whole genome duplication. The pair of genes created from a gene duplication event are called paralogs. If these do not acquire inactivating mutations (there is minimal selection to keep two copies of the same thing active), the **paralogs** may eventually evolve a subset of functions of the original gene or entirely new functions. For example, two paralogs of a gene that is expressed in shoots and roots may evolve into one that is expressed in shoots only and the other that is expressed in roots only.

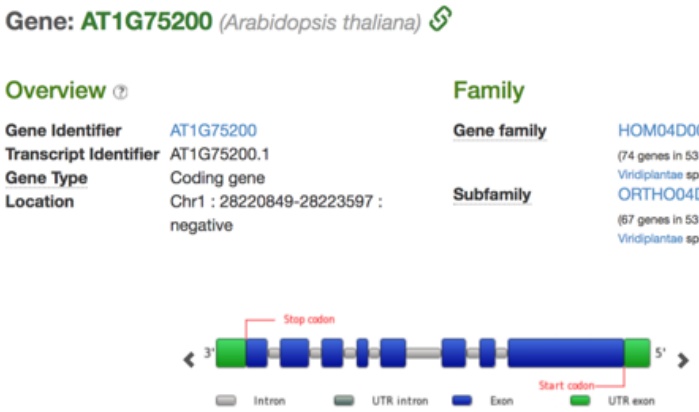
Orthologs are another form of homologous genes. These are genes that are separated by speciation rather than duplication. That is, equivalent genes may be found in different species that share a common ancestor because the ancestor carried this gene. Orthologs may often retain similar functions in different species.

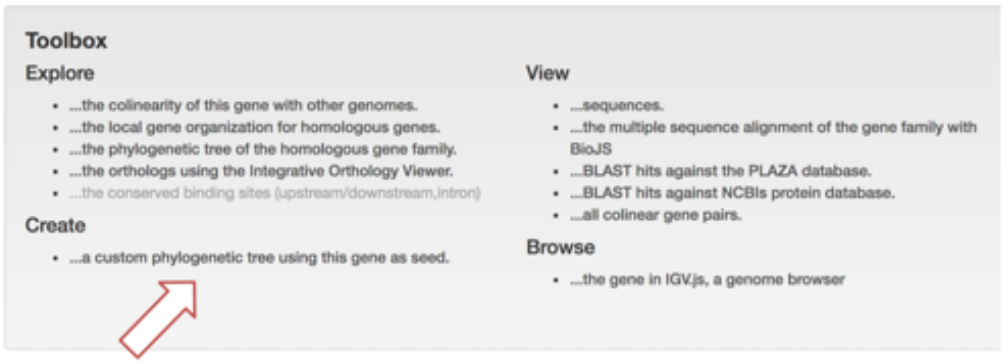


How to Make a Phylogenetic Tree with PLAZA

1. Go to PLAZA https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_dicots/
At the top of the page enter your gene identifier.

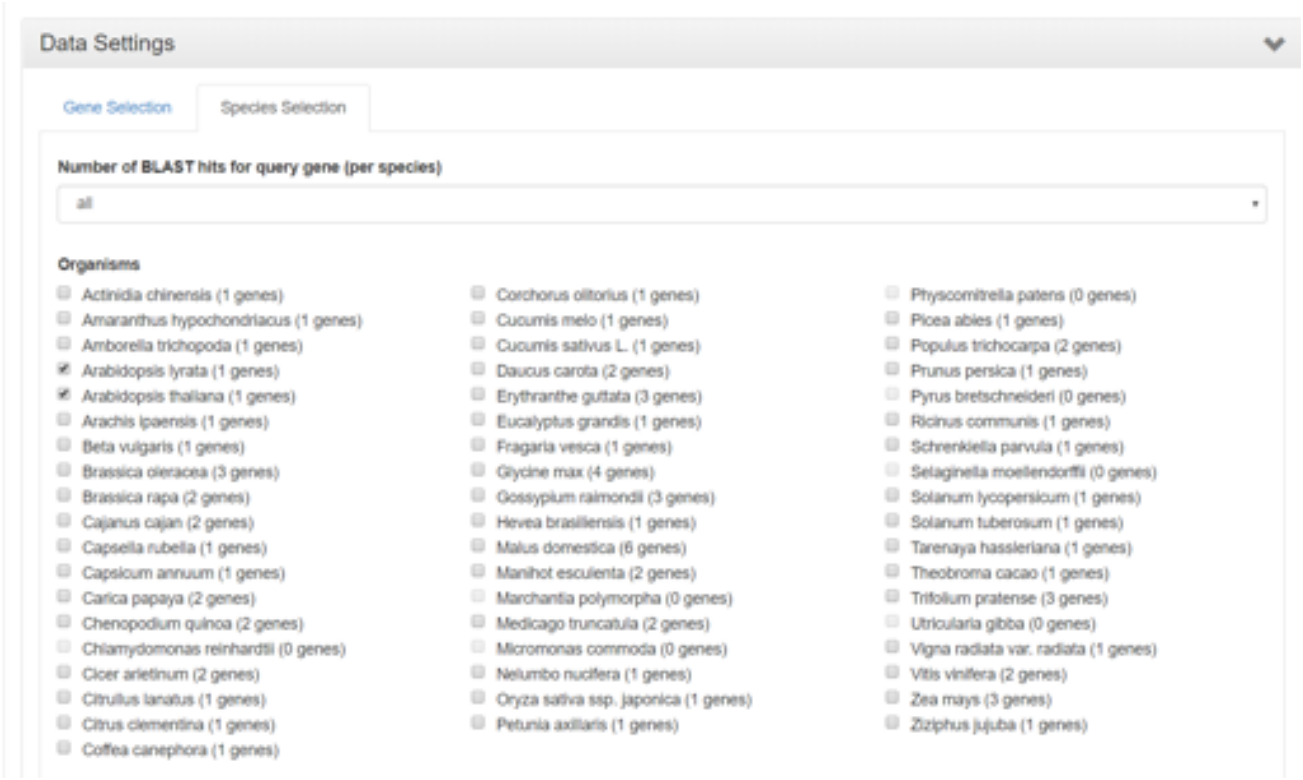
2. This will take you to a page with a diagram of your gene. Scroll down on this page to a gray box that says “Toolbox”. Under “Create” click “...a custom phylogenetic tree using this gene as a seed”.





3. On the next page, scroll down to the data settings and click on the Species Selection tab. There you can pick different plant species to add to your tree. *Arabidopsis thaliana* and *Arabidopsis lyrata* are good picks. If you'd like you can try adding some other species like *Zea mays* (corn), *Oryza sativa* (rice), and *Solanum lycopersicum* (tomato).

NOTE: If you get an error message when it tries to generate your tree, you'll have to come back to the species selection and remove some species.



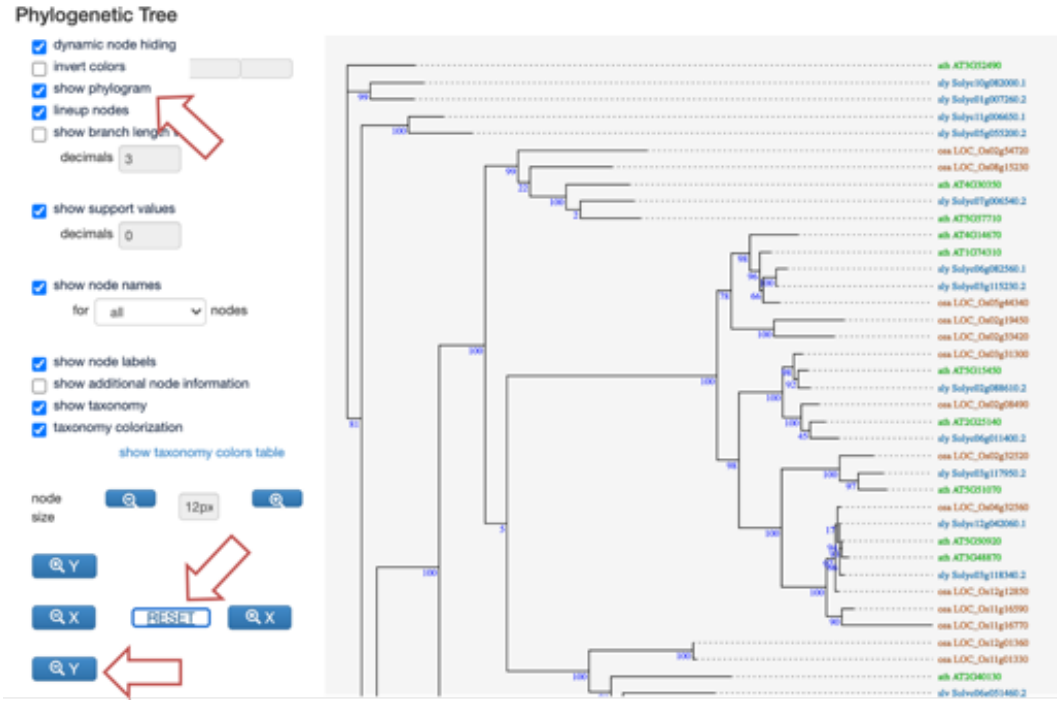
4. After you checked which species you want to try, scroll down and press the blue "Create MSA/Tree" button.

5. There are several useful buttons to be familiar with.

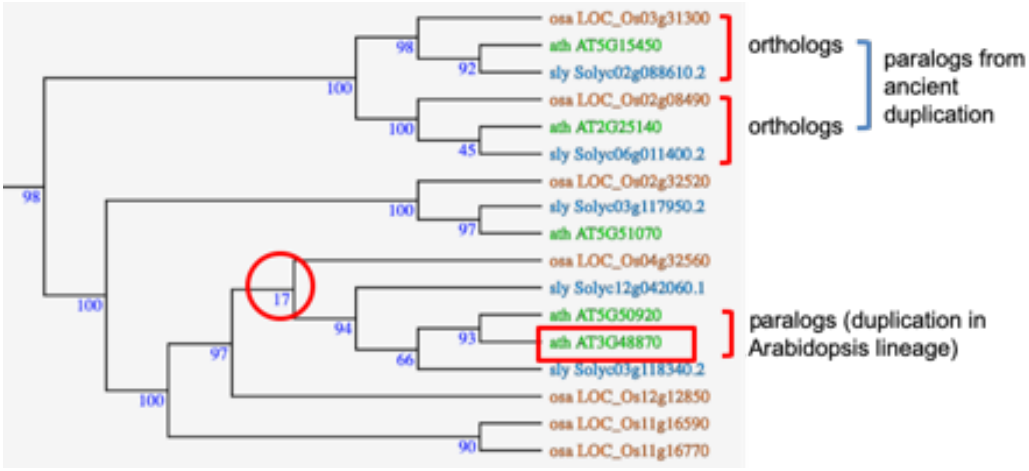
- First it's helpful to **uncheck "show phylogram"**. This present the tree with the same structure, but ignores the branch lengths (which are indicators of sequence divergence).
- To zoom in on your gene you can use the scroll wheel on your mouse. The **"+ Y"** button is useful for adding space between genes if your tree is really dense.
- The tree can be finicky sometimes as you try to zoom in and move around. If you get lost you can always press **RESET** to reset the tree view.

5. If you need help locating your gene on the tree you can put your gene identifier in the **search box** and it will highlight your gene in yellow on the tree.

Note: If you do this and you still don't see your gene on the tree, use the "+Y" button or scroll wheel on your mouse to zoom in.



Below is an example where I zoomed in on my gene (highlighted in yellow). Make sure to save a good close-up view for your lab notebook and report. In the example below, the blue bracket is pointing to an example of a **paralog** (you can tell because it is in the same species as my gene ath - *Arabidopsis thaliana*). The red brackets point to examples of **orthologs** of similar genes in different species.



Appendix 14: Example notebook post – phylogenetic tree

Date: 5/13/20

Purpose: Today I'll be creating a phylogenetic tree for my gene. I'll then use this tree to identify paralogs and orthologs for my gene.

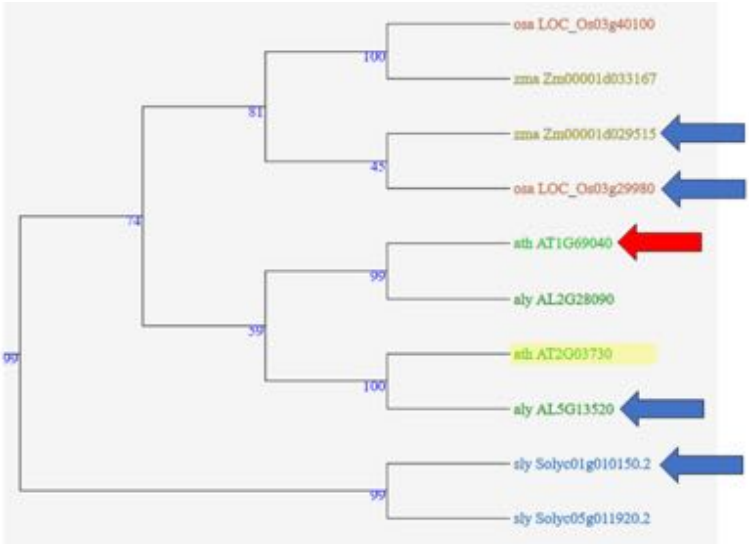
Materials:

- PLAZA 4.0
- TAIR

Methods:

1. 1. Go to PLAZA https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_dicots/. At the top of the page enter your gene identifier.
2. This will take you to a page with a diagram of your gene. Scroll down on this page to a gray box that says "Toolbox". Under "Create" click "...a custom phylogenetic tree using this gene as a seed".
3. On the next page, scroll down to the data settings and click on the Species Selection tab. There you can pick different plant species to add to your tree. *Arabidopsis thaliana* and *Arabidopsis lyrata* are good picks. If you'd like you can try adding some other species like *Zea mays* (corn), *Oryza sativa* (rice), and *Solanum lycopersicum* (tomato).
4. After you checked which species you want to try, scroll down and press the blue "Create MSA/Tree" button.

Results:



Phylogenetic tree for AT2G03730 (ACR5) highlighted in yellow compared to *A. lyrata, S. lycopersicum, Z. mays,* and *O. sativa*. Blue arrows indicate orthologs. The red arrow points to a paralog.

Conclusion:

My gene has orthologs in A*. lyrata, S. lycopersicum, Z. mays,* and *O. sativa*. This suggests that this gene is conserved across different plant species. It has one close paralog. I looked up this gene ID on TAIR. This makes sense because these genes appear to have similar amino acid lengths.

Appendix 15: E CRISP protocol

Designing guide RNAs for CRISPR Cas9

Purpose:

Your goal is to design guide RNAs to disrupt your *A. thaliana* gene of interest. These genes were chosen because the protein encoded by the genes likely interact with TAN1 or AIR9 or there is evidence they are involved with some aspect of division plane orientation in Arabidopsis.

Today you'll be designing two sets of guide RNAs to be used in the CRISPR-Cas9 system to mutate your gene of interest. These guide RNAs will be used to remove (also sometimes referred to as "knocking out") your gene in the *air9* single mutant. If your gene of interest acts in the same pathway as TAN1, we expect to see a severe phenotype similar to the short stature and misplaced cell walls seen in the *tan1 air9* double mutant. If your gene is not in the pathway, or is unimportant for division plane orientation, we expect to see no change in the *air9* single mutant phenotype.

The CRISPR-Cas9 system:

Cas9 is a double-stranded nuclease that cleaves DNA at specific locations based upon the guide RNA it is carrying. The guide RNA can be 17-20 nt long and can be any sequence as long as it is followed by the correct protospacer adjacent motif (PAM) in the genome being targeted. The PAM for Cas9 is NGG, where N is any of the four nucleotides. (See figure below for an example of how Cas9 binds to and cleaves its target.) Other enzymes with similar activity to Cas9 have been found that have different PAM requirements

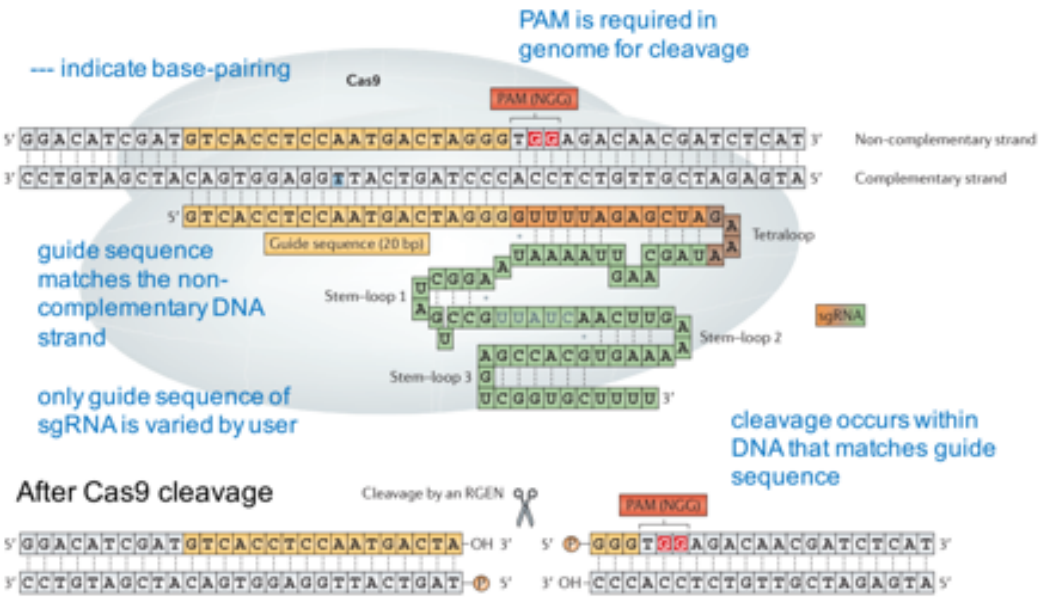


Figure adapted from Kim & Kim, 2014, *Nat Rev Gen*

Guide RNA design considerations:

If G makes up 25% of the bases in a genome and is equally distributed, on average a guide RNA for Cas9 could be designed every 16 bp. However, this is typically not the case, and further considerations must be made when selecting a gRNA. First, if the guide sequence (+PAM) of a guide RNA is present at more than one location in the genome, Cas9 will cleave at more than one place.

A primary concern with CRISPR-Cas9 is “off-targeting”. This is when Cas9 cleaves DNA that is similar to the guide sequence in another part of the genome, rather than only cleaving at the desired location. Researchers typically design guide RNAs to avoid “off-targeting”, or use Cas9 variants that have higher fidelity. Note that even with some off-targeting, genome editing with Cas9 induces far, far fewer mutations than traditional methods of mutagenesis, which makes it ideal for developing genetically engineered organisms.

Second, some guide sequences enable more effective target cleavage than others based on a number of factors that are not always easy to predict. This is why it can be particularly useful to design and try several different guide sequences.

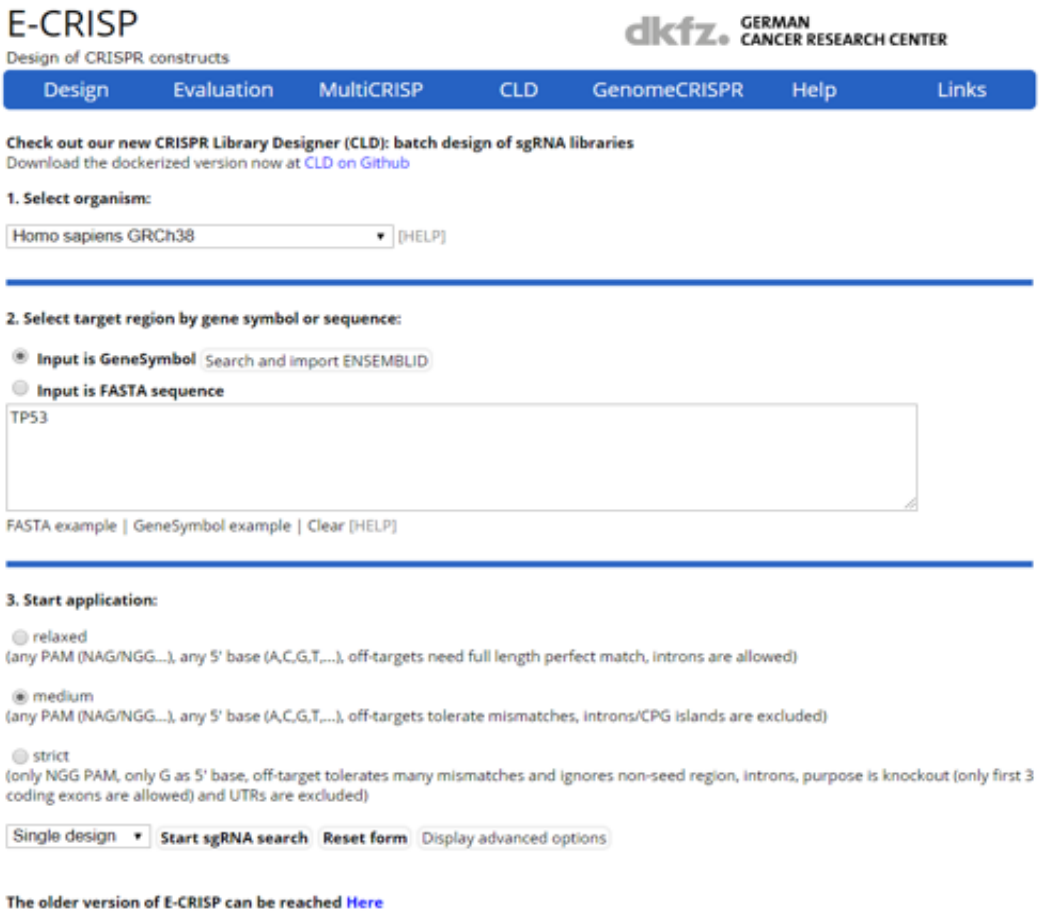
Third, the purposes of the experiment must be taken into consideration. In many cases, the goal is simply to introduce a loss-of-function mutation into a gene. This requires targeting an exon, as the small insertion/deletion (indel) mutations caused by Cas9 are unlikely to have an effect if they occur in introns. Mutations in non-coding regulatory regions adjacent to the gene may have an effect on gene expression but this is usually unpredictable. In contrast, an indel in an exon will cause a frameshift mutation approximately 2/3 of the time. Frameshift mutations cause the

ribosome to incorporate the wrong amino acids as it translates an mRNA, effectively stopping translation of the correct protein sequence at the point of the mutation. Loss-of-function mutations will generally work well if placed closer to the beginning of the gene than the end, but caution must be exercised with mutations very near the beginning of the gene as translation could potentially proceed from an alternative downstream start codon.
We are going to design two guide RNAs to target each gene. This has the potential to create large deletions between the two target sites, which are effective at knocking out a gene's function and are also very easy to detect.

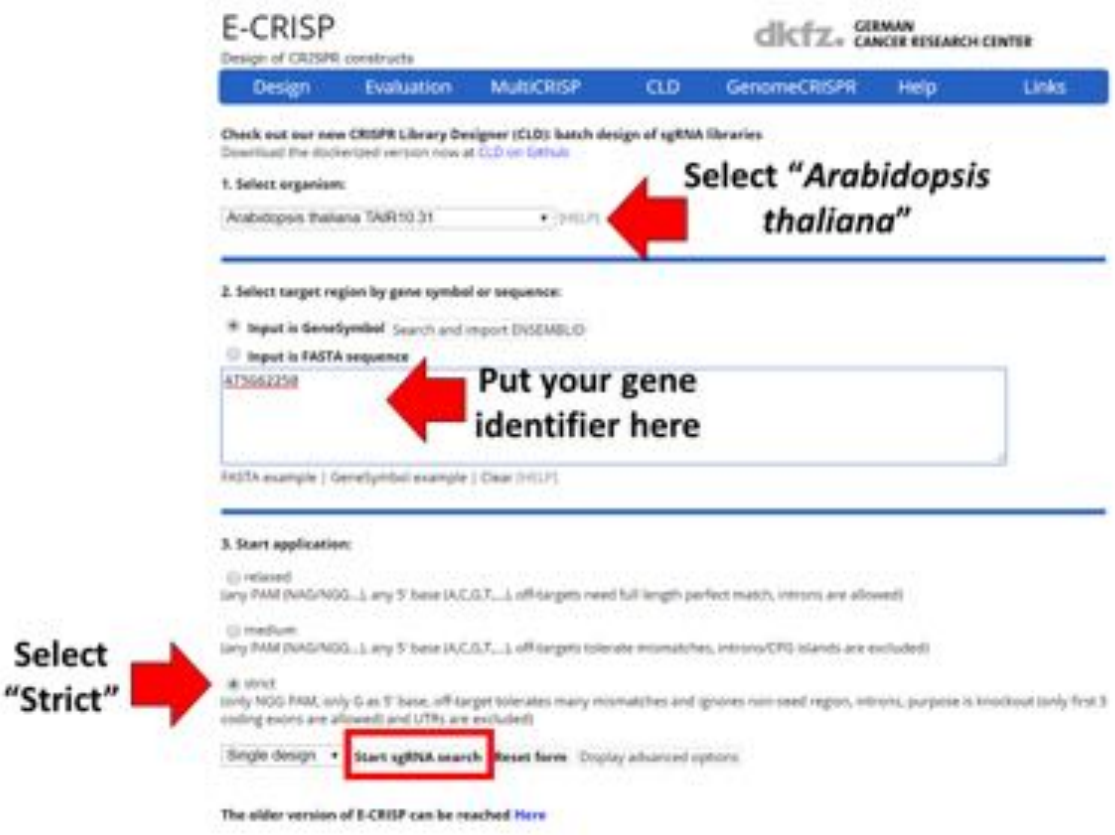
E-CRISP is a tool that we can use to easily design guide RNAs in plants. The CRISPR-Cas9 vector we will introduce into plants can express two gRNAs. So you'll use E-CRISP to design multiple pairs of guide RNAs to remove a part of your gene and your whole gene.

Selecting Guide RNAs Using E-CRISP

1. Go to http://www.e-crisp.org/E-CRISP/ You should see this screen below.



2. Now you'll need to make the following changes to the setting:
    - From the "Select organism:" drop down menu. Select "*Arabidopsis thaliana* TAIR10.31"
    - Put your gene identifier into the box in the "Select target region by gene symbol or sequence"
    - Under "Start application" select "strict"
3. Press "start sgRNA search"!



4. Your results should show up in a new tab in your browser! You'll see something like the image below. If you have fewer than 5 guides refer to "Tips and Troubleshooting" at the end of this document.

E-CRISP

Design of CRISPR constructs

dkfz. GERMAN CANCER RESEARCH CENTER

Design | Evaluation | MultiCRISP | GenomeCRISPR | Help | Links

Download a tabular report for all query sequences together

Download a Excel formated tabular report for all query sequences together

Download a GFF-File for all query sequences together

Query name: AT5G62250 Query length: 3341 Query location: 5::25005756::25008096

Total number of possible designs = 150
Number of successful designs = 15
Number of designs that hit a specific target = 15
Number of designs excluded because they were not located in a coding sequence = 3
Number of designs excluded because they did not hit any gene = 5
Number of designs excluded because their nucleotide composition was not within the given ranges = 10
Number of designs excluded because their nucleotide composition contained TTTT = 117

S: Specificity score A: Annotation score E: Efficiency score
for more information please see the Help pages

Name	Nucleotide sequence	SAE-Score	Target	Matchstring	Number of Hits
AT5G62250_3_477	TGGTGTCTTCCAAGA ATTGG NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_4_477	TGATGGTGTCTTCCA AGAAT NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_1_477	TATTGATATATGTTCT GTAA NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_2_477	ATGGAAGAGCCACCA ATTCT NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_0_954	TAGCTTCATCAGTACA ACAA NGG		AT5G62250::MAP65-9	Matchstring Info	1

5. As you scroll down, you'll first see all of your potential guide RNAs.

S: Specificity score A: Annotation score E: Efficiency score
for more information please see the Help pages

Name	Nucleotide sequence	SAE-Score	Target	Matchstring	Number of Hits
AT5G62250_3_477	TGGTGTCTTCCAAGA ATTGG NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_4_477	TGATGGTGTCTTCCA AGAAT NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_1_477	TATTGATATATGTTCT GTAA NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_2_477	ATGGAAGAGCCACCA ATTCT NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_0_954	TAGCTTCATCAGTACA ACAA NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_2_2385	AGTACTCCTCTATCAC GCCA NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_0_1908	GAACAGATCGAGTAT CGAGC NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_2_1431	GTCGGTTGGAGGAG TTGAAG NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_3_1431	TCTTGAAGAAGAAGA TATCG NGG		AT5G62250::MAP65-9	Matchstring Info	1
AT5G62250_1_1431	AGCTGAGCTATGTCG		AT5G62250::MAP65-9	Matchstring Info	1

Each column has useful information about the individual guides:

Column	Meaning
Name	The ID of the guide. This is of the form: ID of the input sequence_randomNumber_randomNumber.
Nucleotide Sequence	The guide RNA target sequence
SAE Score	S: Specificity score A: Annotation score E: Efficiency score. See the table below for more information.
Target	The gene that is targeted by this gRNA.
Match String	A colored match screen, which indicates at a glance how good the alignment is. You can click on “Matchstring info” and it’ll bring up a window that shows how well the guide matches the target sequence, where a green "M" for a match. A "X" for a mismatch, an "I" for an insertion in the guide RNA.
Number of Hits	The number of locations this CRISPR design targets, or, the number of times this CRISPR appears in the output table (one for each target). If your gene has multiple hits, it’s good to check the “Matchstring info” for the off target region. If there are a number of mismatches “X”s for the second gene, the guide is still useful! See the “Tips and Troubleshooting” at the end of the document for details.

Each guide also has an SAE score that will help you determine whether or not it is a really good guide RNA or just an okay one. All of the scores are out of 100 and the higher each one is the better the guide RNA is predicted to be.

- S = Specificity score - this score indicates how specific this guide RNA is for the target. The more off targets (other places in the genome other than our target gene that the guide can bind), the lower this score is. The default excludes any potential guide RNAs

with more than 5 off targets, so the “S” score should be pretty good for most of your guide options.

- A = Annotation score - this score indicates whether the guide is in a good region of the gene. Guides that land in protein coding exons tend to score higher.
- E = Efficiency score - this score indicates how likely it is the guide will result in a cut at the site. This score is based on a number of factors regarding the sequence of the guide RNA.

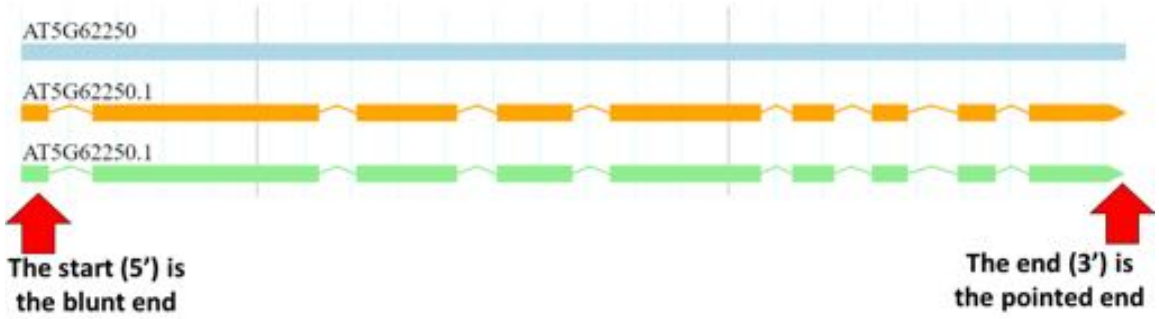
If you’d like to read more details about the SAE score (optional) you can check out this about page: http://www.e-crisp.org/E-CRISP/aboutpage.html

6. Scroll down further and you’ll find a diagram of your gene that includes the locations of the guide RNAs. There are several important features to note here!

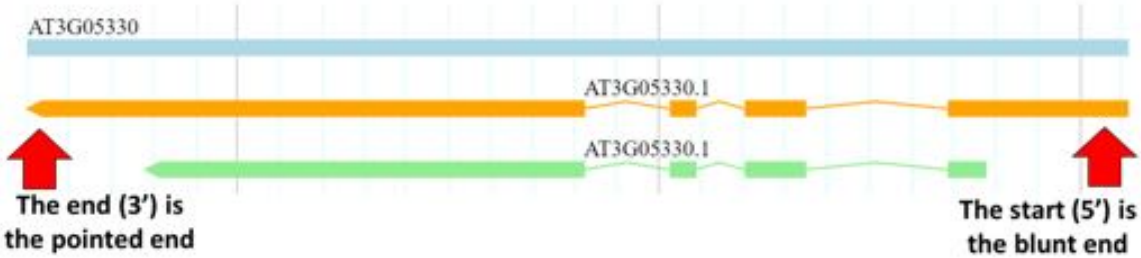
Orientation of Your Gene:

The orientation of your gene will vary sometimes depending on which strand of DNA your gene is encoded on. You can tell which is the start/front of your gene (also called the 5’ end) because it’ll be the blunt end. The end of your gene will be the part of the gene diagram (also known as the 3’ end) that is pointed. It’s good to know what orientation your gene is so you know which exons you are targeting (with exon 1 always being at the starting 5’ end of the gene). Below are examples:

Here you can see the start for this gene is on the left.

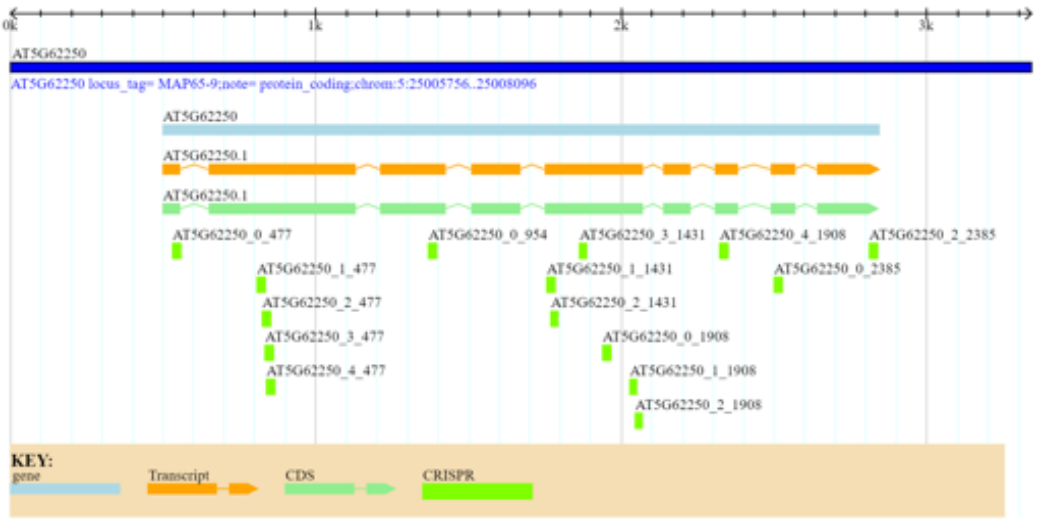


Pay attention though! Other genes start on the right side.



What the Different Colored Bars Represent:

You'll notice there appear to be multiple versions of your gene diagram in different colors. The key in the orange highlighted area notes what each diagram represents.

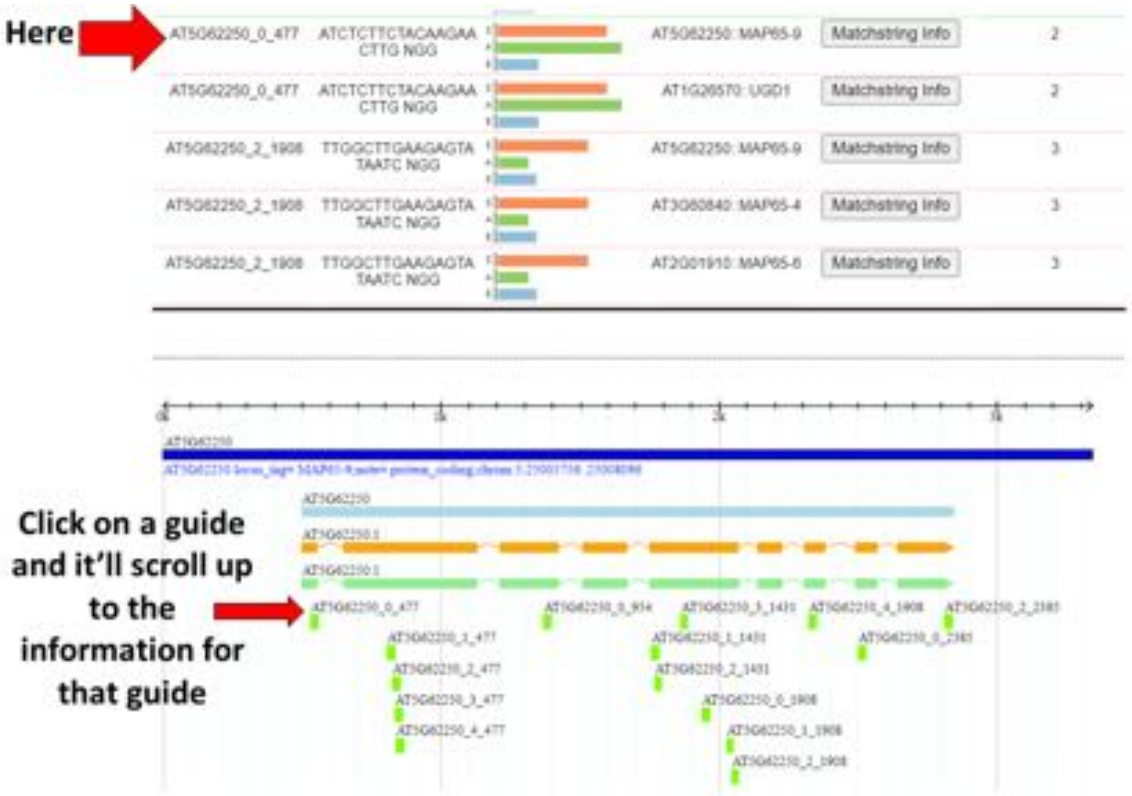


The **light blue** represents your entire gene, and won't be very useful for you. Exons and introns are not indicated. This represents your gene relative to its location within the chromosome (also called its locus). The chromosome is represented as the **dark blue** bar above it. Above that is a black scale bar with 1k (1000bp) tic marks. Each set of smaller tick marks represents 100bp. This can be useful for determining approximately how far apart your guide pairs are.

The **orange diagram** is the mRNA transcript of your gene. Here you can see the introns (represented as thin lines) and exons (represented as thick rectangles) for your gene. Sometimes it may look longer than the light green CDS diagram below it, because mRNA transcripts often include UTRs (untranslated regions). These are regulatory regions that are not translated into the final protein.

The **light green diagram** is the CDS (CoDing Sequence) for your gene. This will be the most useful for identifying guide RNAs because it represents the part of your gene that is translated into protein by the ribosome. We want to eliminate the protein so it makes sense to target the protein coding regions. This diagram also has introns and exons represented as thin lines and rectangles respectively.

The **bright green segments** below that are the locations of your potential guide RNAs within your gene. You can click on the names of the guide RNAs in the diagram to take you to the information related to that guide RNA in the list above. This is useful for comparing information on where the guide is located vs how good the guide is. Make sure to double check the names match for the guide on the diagram and the guide information. If you are zoomed in on your browser window, you may not see the correct guide RNA in the list when you click on the name in the diagram.



7. Now you are ready to make your guide RNAs!

Specific Guide RNAs You'll Design:

You should have a minimum of four guide RNA pairs by the end of the exercise (although you're welcome to create more!):

- Two pairs targeting a small section of one end of the gene (either the front or back)
    - These guides should be designed to remove a small section of the gene (approximately 50-300 nucleotides).
    - If there's only one good pair near the start or back, try designing a pair that would remove a region in the middle of the gene. It's good practice!
    - You'll put both pairs in your notebook. Either take a screenshot of the guides details or copy and paste the sequences of the guides. Then in your lab notebook conclusion discuss which guide RNA pair you would use out of the two (discuss why based on the SAE score and position of the guides). Feel free to confer with your classmates, TA, and instructor to see if they agree.

- Two pairs targeting the entire gene

- These guides should be designed to remove the entire gene (or as much as possible)
- If there is only one really good pair for removing your entire gene, try designing a pair to remove a large section of your gene (over 500bp). Even though you'll choose the whole gene pair as your favorite, it's still good practice.
- You'll put both pairs in your notebook. Either take a screenshot of the guides details or copy and paste the sequences of the guides. then in your lab notebook conclusion discuss which guide RNA pair you would use out of the two (discuss why based on the SAE score and position of the guides). Feel free to confer with your classmates, TA, and instructor to see if they agree.

Enter Your Best Guide RNAs Into the Class Datasheet

Finally, you'll put your favorite guide pairs you discussed in your conclusion (the best pair for targeting a small section of the gene and the best pair for targeting the whole gene) into the spreadsheet below. **Make sure you add them to your class section.**

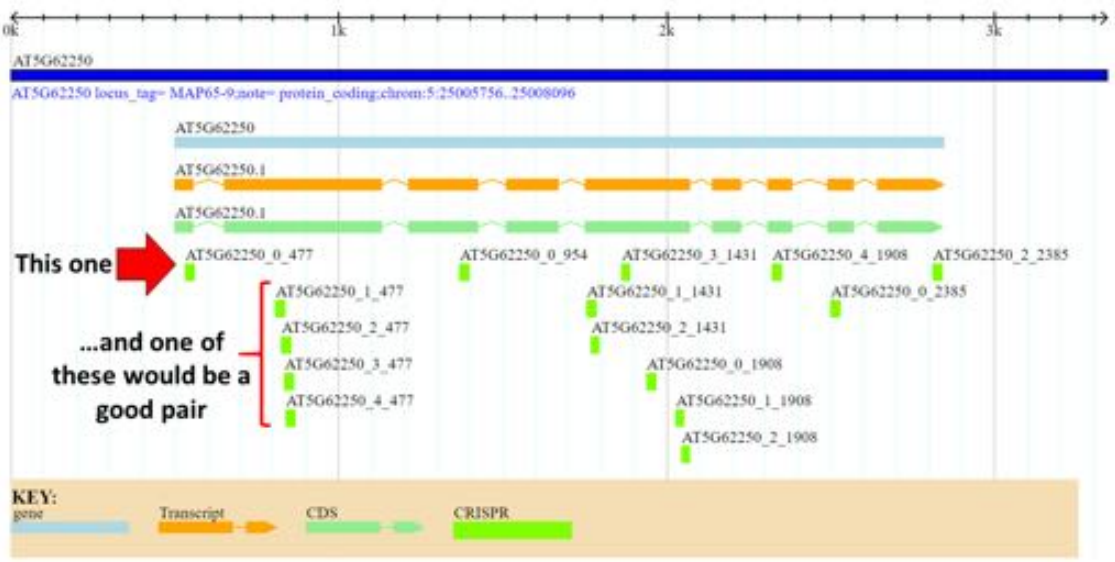
See Appendix 17 for example.

Tips and Troubleshooting

Example of how to target a part of a gene:

To remove part of my gene I need to find two guide RNAs within my region of interest. In this case, the start of my gene. You can think of it as deciding which two places you are going to cut in your gene to remove a piece (kind of like marking where you are going to cut a ribbon to remove a specific sized piece).

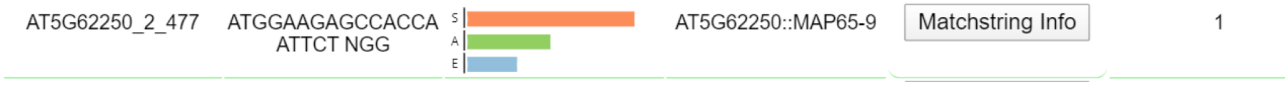
There is a guide in exon 1 (indicated by the red arrow). So this is a good place to have my first cut.
For my second cut I want to pick another site nearby. Here you can see I have a bunch of overlapping options in exon 2 (indicated by the red bracket) about 300bp away from my first guide. Any one of them would hypothetically work, so I could design multiple pairs using AT5G62250_0_477 and any one of bracketed options. (For example I could do AT5G62250_0_477 and AT5G62250_1_477, AT5G62250_0_477 and AT5G62250_2_477, etc.).



If you look at the details for each option, you can see they are all about equally good, which makes sense since they are overlapping within the same region.

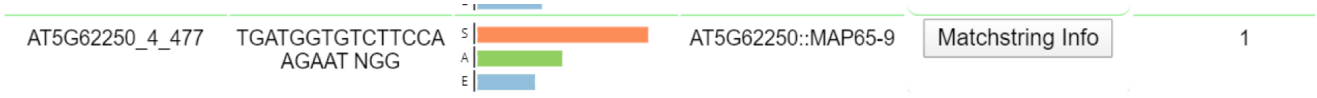
AT5G62250_1_477

AT5G62250_2_477



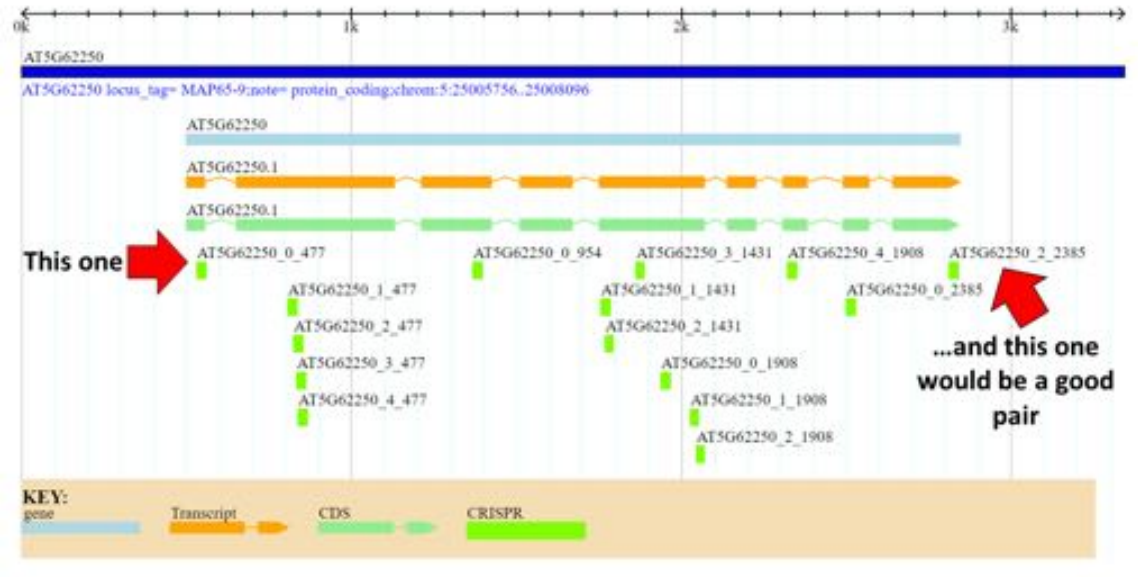
AT5G62250_3_477

AT5G62250_4_477



Example of how to target the entire gene:

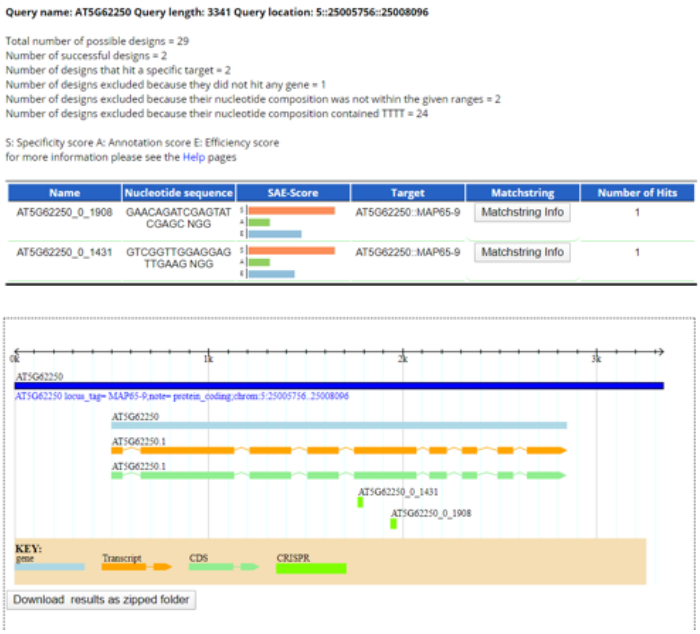
The principles are the same accept now I want to pick one guide near the start of my gene, and one guide near the end of the gene in order to cut out the entire gene.



So the best guides to remove the entire gene in this case would be AT5G62250_0_477 and AT5G62250_2_2385.

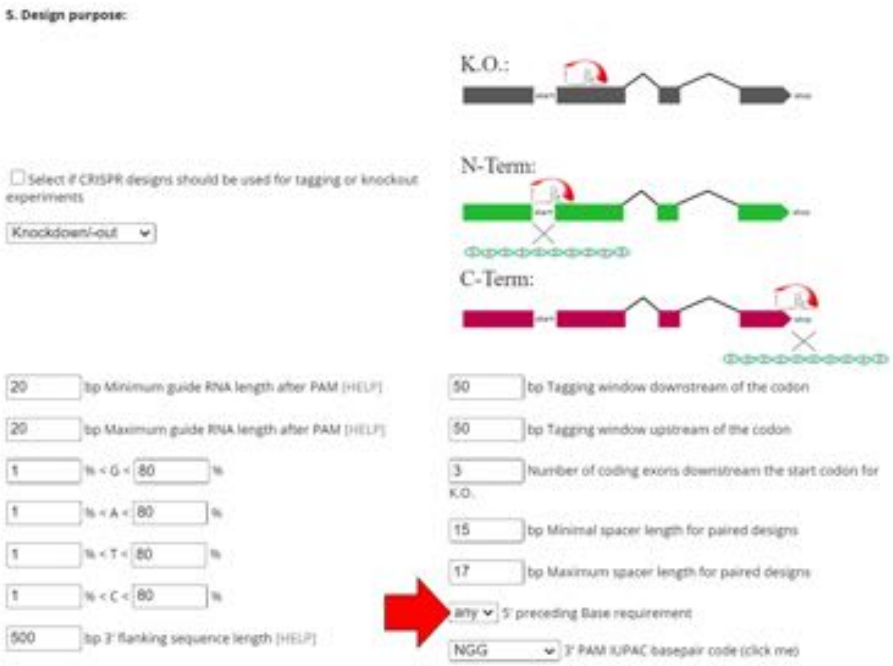
What if I only get a few guide RNAs or no good options near the ends of my gene?

If you get results that look like this with less than 5 guide RNAs and/or no guide RNA options at the ends of the gene (like the example below) you can modify the parameters for your search in the advanced options.

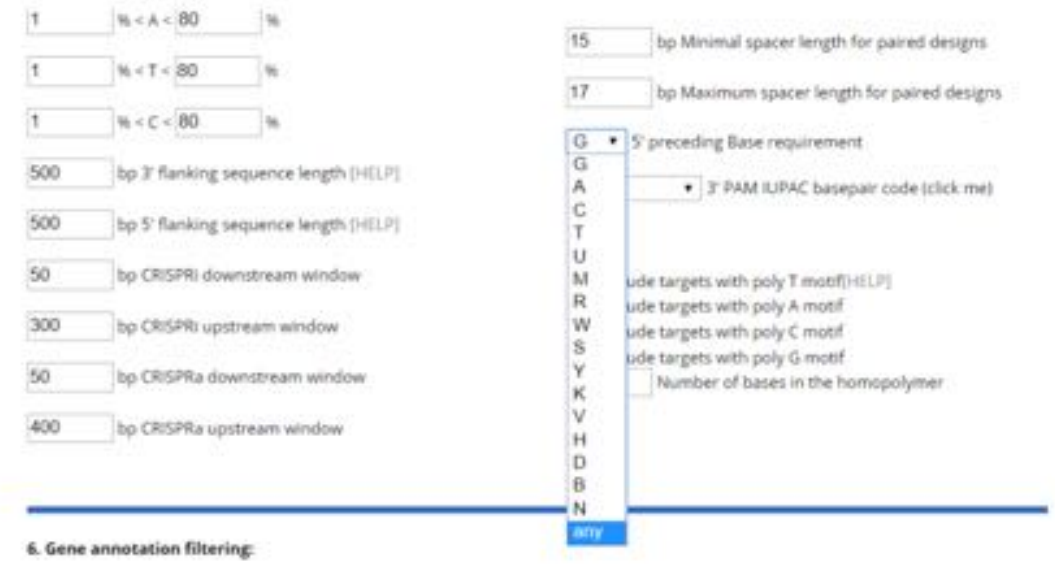


Go back to your first E-CRISPR page and press “display advanced options”. This will open up a section with a lot of options directly below the button.

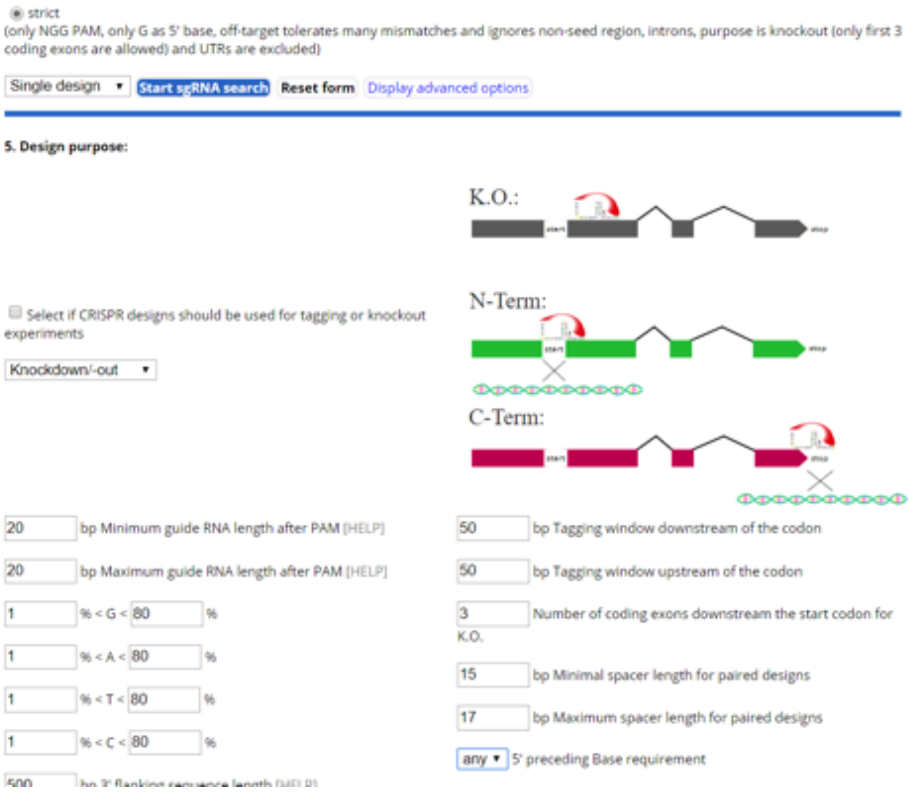
Find the “5’ preceding Base requirement” in the advanced options.



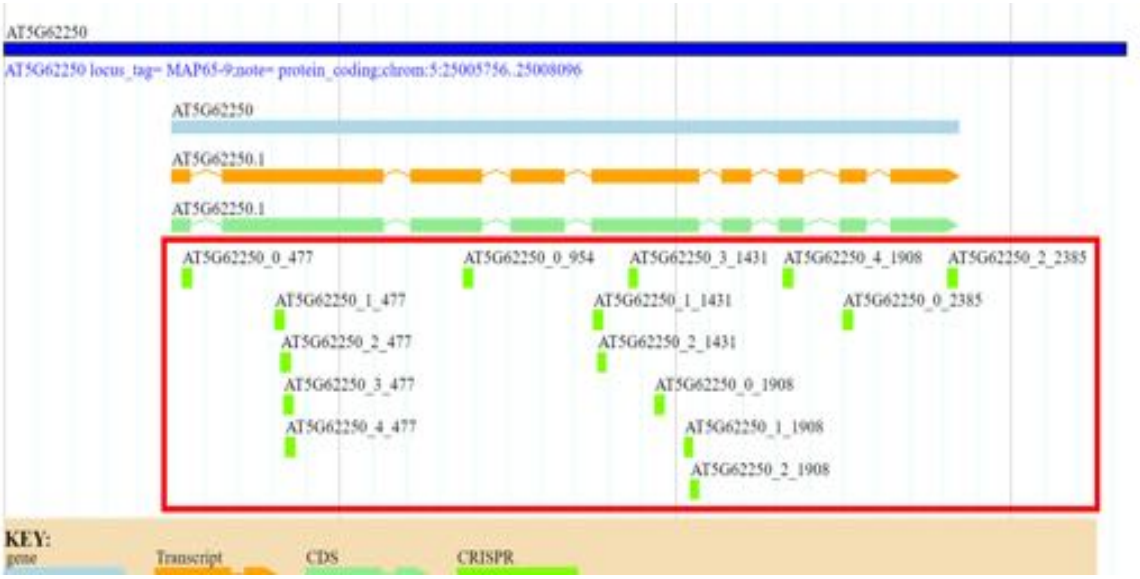
Open up the drop-down menu for the “5’ preceding Base requirement” and select “any”



Now press “start sgRNA search”.



You should now have many more options to choose from for your guide RNAs (see how many more there are in the example below)!

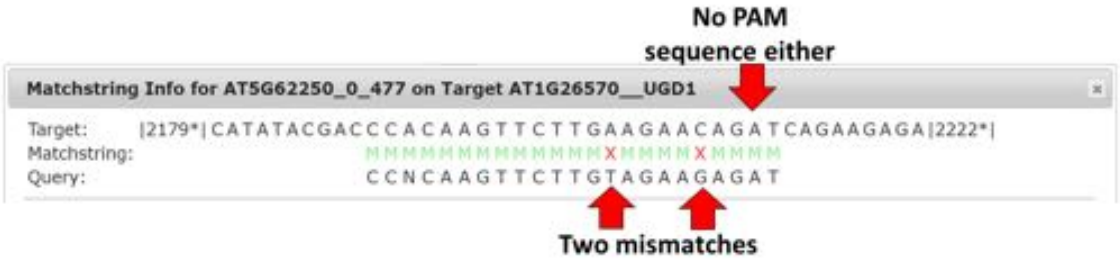


How do I make sure my guide is still usable if it has multiple hits?

Below is an example of a guide with 2 “hits” indicated by the red box. Ideally you want a guide with only 1 “hit” but sometimes a guide with multiple hits might be in a very useful position.

To check if a guide with multiple hits is still usable. Look through the guide list and find the second “hit”, which will have the same name (indicated by the red arrow) but a different gene identifier and name than your target gene (shown in the orange box). Click the “matchstring info” box (shown in the blue box) to bring up a summary of how well this guide matches the off-target gene.

So, we can see that this guide is still useful for targeting our gene of interest because the off-target site has multiple mismatches (indicated by red Xs) and it lacks a PAM sequence. This makes it extremely unlikely that this other site will be cleaved. Remember that in order for Cas9 to cut DNA the guide needs to be a very good match (no mismatches) and it needs a PAM sequence.



Appendix 16: Example notebook post – guide RNAs

Date: 5/18/20

Purpose: Today I will design pairs of guide RNAs to target my gene MAP65-9 (AT5G62250). I will design two pairs targeting a small section at the front of my gene and two pairs to remove the whole gene. Eventually, these guides RNAs will be used to mutate MAP65-9 in the *air9* single mutant. By doing so we'll determine if MAP65-9 enhances the *air9* mutant phenotype by examining the phenotype of the *map65-9 air9* double mutant.

Materials:

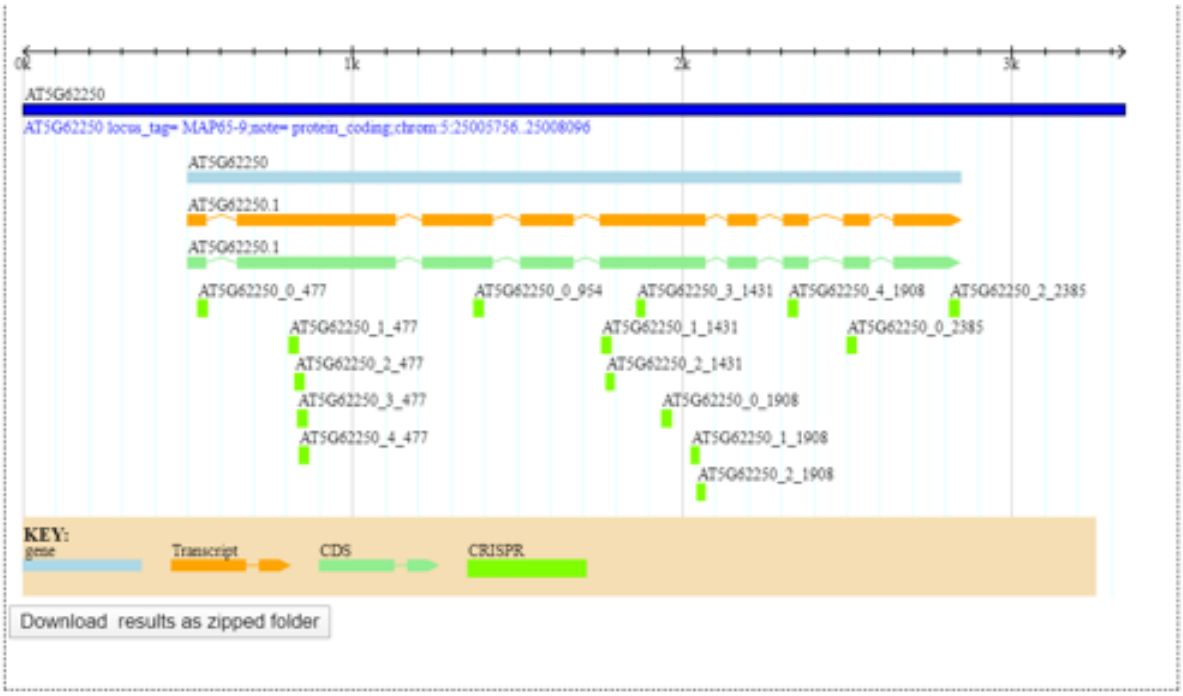
- http://www.e-crisp.org/E-CRISP/

Methods

1. Go to http://www.e-crisp.org/E-CRISP/ You should see this screen below
2. Now you'll need to make the following changes to the setting:
    - From the "Select organism:" drop down menu. Select "*Arabidopsis thaliana* TAIR10.31"
    - Put your gene identifier into the box in the "Select target region by gene symbol or sequence"
    - Under "Start application" select "strict"
3. Press "start sgRNA search"!
4. My initial search had too few guide options so I changed the settings to
5. Two pairs targeting a small section of one end of the gene (either the front or back)
    a. These guides should be designed to remove a small section of the gene (approximately 50-300 nucleotides).
    b. If there's only one good pair near the start or back, try designing a pair that would remove a region in the middle of the gene. It's good practice!
6. two pairs targeting the entire gene
    a. These guides should be designed to remove the entire gene (or as much as possible)
    b. If there is only one really good pair for removing your entire gene, try designing a pair to remove a large section of your gene (over 500bp). Even though you'll choose the whole gene pair as your favorite, it's still good practice.
7. Put your favorite guide pairs you discussed in your conclusion (the best pair for targeting a small section of the gene and the best pair for targeting the whole gene) into the class spreadsheet.

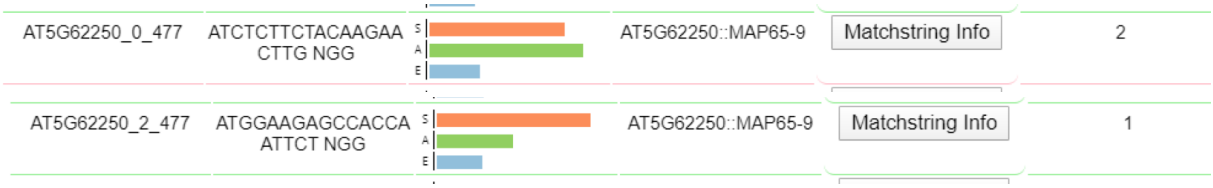
Results:

Below are my guide RNA options for MAP65-9



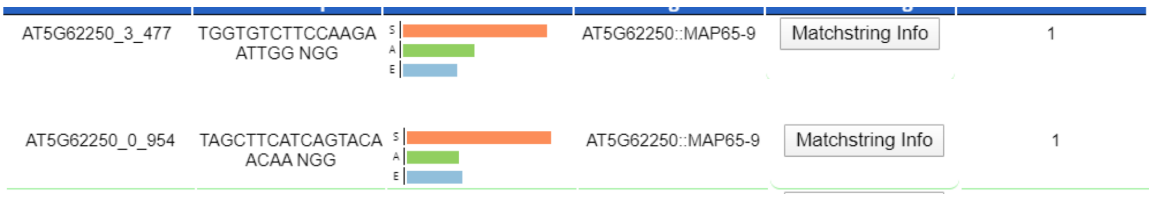
To remove a small section of the 5' part of my gene:

Pair 1: Removes part of exon 1 and exon 2
Notes: I noticed that AT5G62250_0_477 targets AT1G26570 too but that it contains 2 mismatches for this other gene.



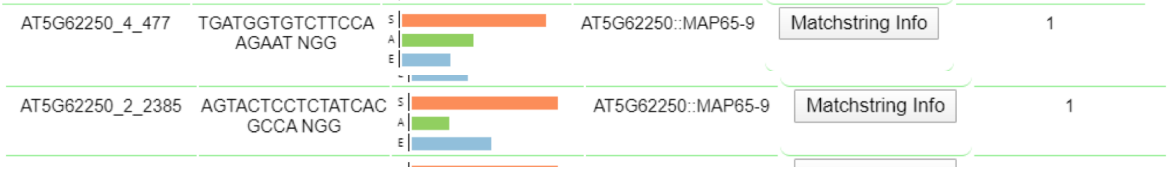
Here is the other gene the first guide targets:

Pair 2: Removes part of exon 2 and exon 3

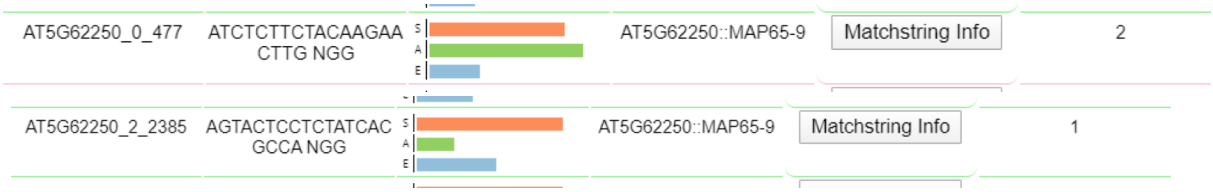


To remove the whole gene:

Pair 1: Removes exon 2 to exon 9 (the last exon)



Pair 2: Removes exon 1 to exon 9 (the last exon)
Note: Again AT5G62250_0_477 also may target one other gene (with 2 mismatches)



Conclusion:

To remove part of the gene I would choose pair 1. Even though one of the guides has one off target, it has two mismatches and no PAM sequence so it's unlikely it'll effectively cut the off-target site. This will remove exon 1 and part of exon 2, which is likely to disrupt the whole gene.

To target the whole gene, I would use my second pair. It uses the same exon 1 primer as my partial gene pair, which has one off target. However, it is still usable for the reasons I previously described. This is the best pair because one cuts in exon 1 and the other cuts in exon 9 (the last exon of my gene). This means it would remove almost my entire gene.

Appendix 17: Example class guide RNA spreadsheet

Name (Last, First)	Gene Name	Gene Identifier	Region (which exons or whole gene)	Guide 1	Guide 2
Mills, Alison	MAP65-9	AT5G62250	exon 1 and exon 2	ATCTCTTCTACAAGAACTTG NGG	ATGGAAGAGCCACCAATTCT NGG
			whole gene	ATCTCTTCTACAAGAACTTG NGG	AGTACTCCTCTATCACGCCA NGG
Doe, Jane	CEN2	AT3G50360	Exon 1 and Exon 2	TGGAGGATTCGTAGAAGAT A NGG	AAGAGCTTAATGTTGCTATG NGG
			Whole Gene	TGGAGGATTCGTAGAAGAT A NGG	ATGAGGAGAACTGCTTATGG NGG

Appendix 18: Quiz 2

Quiz 2

Instructions

All answers must be typed. Use the Word Formula editor for calculations. Please use BLACK font to answer the questions. Upload the document to the assignment for grading.

1. (8 points) Use TAIR to find out the following information. Perform the gene search for AT3G05330.

A. How many number of locus matches are there?
B. What is the name of the gene?
C. Describe its biological function.
D. Give the gene expression atlas.

2. (6 points) Use PLAZA to determine the following:

A. Gene structure of AT3G05330.
B. Build a homology tree for this gene with genes from Oryza sativa, Zea mays, and, Solanum lycopersicum species.
C. Find and label the orthologs and paralogs.

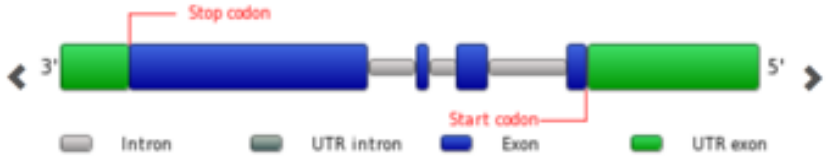
3. (6 points) Use E-CRISP to design sgRNA for AT3G05330.

A. Graphic summary of the result.
B. A pair of sgRNA sequences, location, and how many hits each one has.

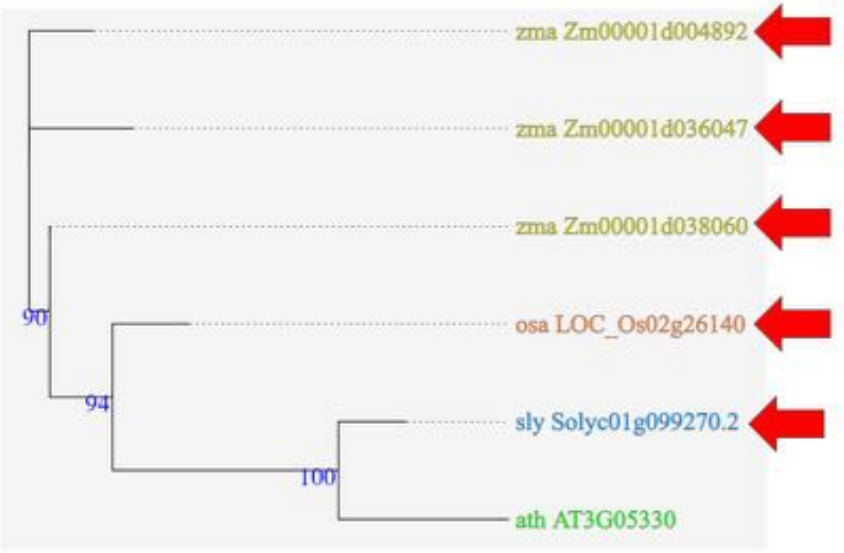
Appendix 19: Quiz 2 Answer Key

1. Use TAIR to find out the following information.
   A. 1
   B. TAN1
   C. Encodes a protein with moderate sequence similarity to the maize microtubule-binding protein TANGLED1. Involved in the identification of the division plane during mitosis and cytokinesis.
   D. It's highly expressed in the axis of inflorescence, young leaf, and pedicel.

2. Use PLAZA to determine the following:
   A.



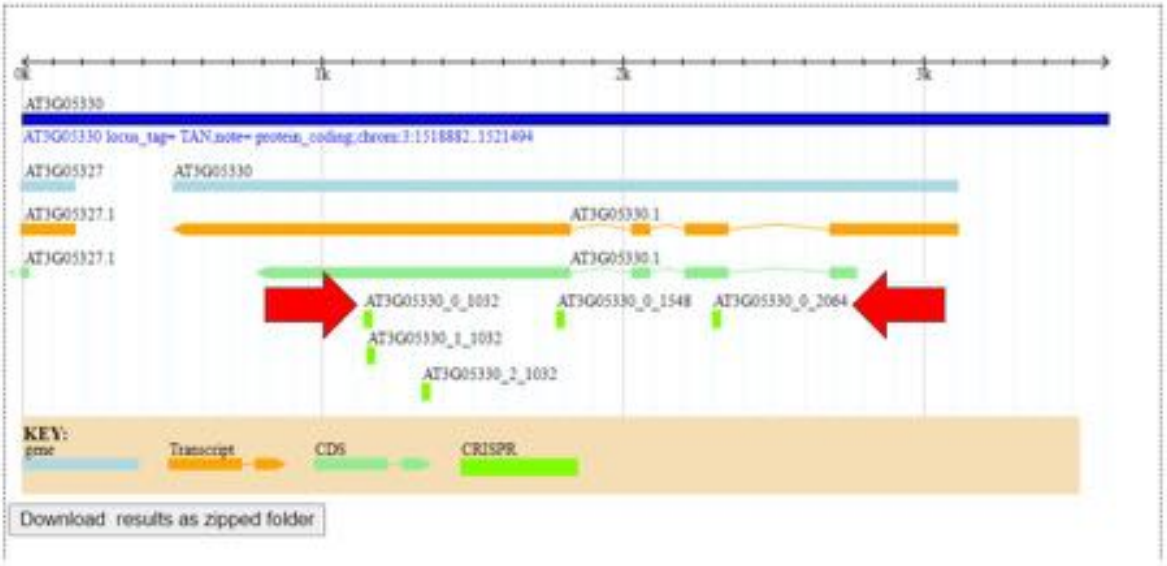
   B.



   C. Red arrows indicate orthologs. There are no paralogs in Arabidopsis for TAN1.

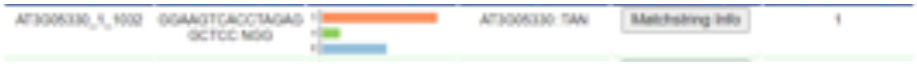
3. Use E-CRISP to design sgRNAs for AT3G05330.

A.



B.

This guide has 2 hits



The other guide has 1 hit

Appendix 20: Presentation Guidelines

Project Presentation

Please read these instructions carefully. The presentation is a quick overview of the background, experimental design for your project on *Arabidopsis thaliana*.

The presentation format

- Must be prepared using **Google Slide (preferred) or Powerpoint**.
- Share with instructor before 10 p.m. the night before your class
- 1 Title Slide with your name
- 5 additional content slides
    1. Background on project
        a. What is your assigned gene (name and locus)?
        b. Known/reported function
        c. Gene structure (exon/intron)
        d. PCR Primers (location, sequences, product size)

    2. Phylogenetic tree (known homologs)

    3. Gene expression pattern

    4. Guide RNA sequence
        a. gRNA primer pair – part of the gene (location, sequences)
        b. gRNA primer pair – whole gene (location, sequences)

    5. Summary
        a. How would you verify that gene was successfully knocked out?
        b. What might be the result of having a non-functional target protein for your assigned gene?

- No more than **5 minutes** total.
- Embed references in the bottom of each slide **not** as a separate “works cited” page.

Due dates and points

- (100 Points) Presentations will be during class on **Wednesday, June 3rd** (MW sections) or **Thursday, June 4th** (TR sections).

Appendix 21: Presentation Rubric

Categories	Points	Justification
Background	30	-Background information related to the project is briefly explained. -The purpose and relevance of the project is explained. -Name and known function of gene is explained (if information is available).
Data	29	-All data acquired over the course of the project is presented. This can include, but is not limited to, a diagram of the gene structure, any primers designed, phylogenetic tree showing closely related genes, expression pattern of the gene (if known), and any guide RNAs designed. -Rationale for primer and guide RNA designs should be explained.
Conclusions	12	-Hypothesis about gene's involvement in the project's proposed genetic pathway. -Future steps for the project are briefly described. -Potential results and outcomes are addressed.
Slides/Style	29	-Slides are legible, complete, and contain references where needed. -Information in the slides is accurate and relevant to the project. -Slides are presented clearly.
Total	**100**	

Sample scoring:

100 points for an outstanding presentation.

85 points for a good presentation that might need to improve on one or two of the categories.

70 points for an okay presentation that is missing major categories or requires significant improvements.

50 points for a weak presentation that is missing major elements and needs a complete overhaul.

Appendix 22: Example presentation score sheet

Background	Overview of cell division in plants	6
	Importance of division plane orientation in plants.	6
	Importance of *TAN1* and *AIR9* genes	6
	Assigned gene	6
	Known or Reported Function of assigned gene	6
Data	Gene Structure	6
	PCR primers	6
	Phylogenetic tree	5
	Gene expression pattern	6
	Guide RNAs	6
Conclusions	Hypothesis. (Is the gene part of TAN1 pathway?)	6
	Future directions; broader impacts	6
Slides/Style	Legible	6
	Accurate	6
	Complete	6
	References	6
	Delivery	5
Total Points		100

Appendix 23: Example student presentation

This student work was used with expressed written consent of the student.

Analyzing the Effect of Katanin P80 Subunit 1 of Cell Division in Arabidopsis Thaliana

Dynamic Genome Section 001

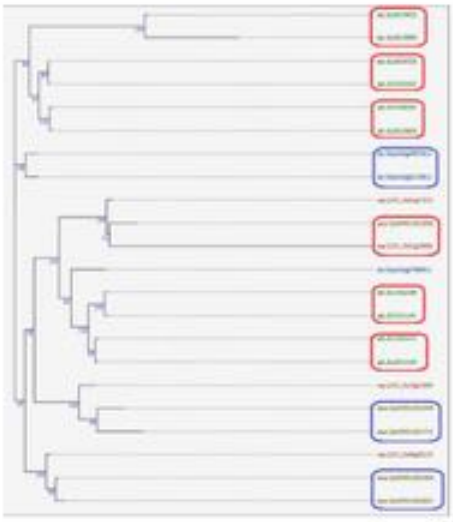
Background

Assigned Gene: KATANIN P80 Subunit 1
Locus: AT1G11160
Function: to target the katanin complex to branch points for severing microtubules
Structure: 

PCR Primers: FP → TCTGCTGATCCGGCAACTTT ; from 3800 bp to 3819 bp ; 20 bp long
RP → TTGTTGTACTGGCTCCTTCAGT ; from 4482 bp to 4461 bp ; 22 bp long
Genomic Product Length: 683 bp
cDNA product length: 265 bp

TAIR: (arabidopsis.org; blast.ncbi.nlm.nih.gov) BLAST: (blast.ncbi.nlm.nih.gov/primer-blast/ primertool) PLAZA: (https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_dicots/)

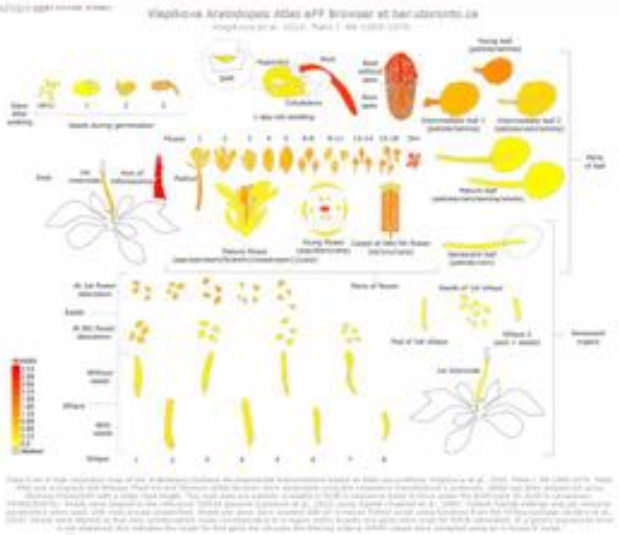
Phylogenetic Tree



Red: Orthologs
Blue: Paralogs

PLAZA: (https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_dicots/)

Gene Expression Pattern



Areas Most Expressed:

- Root
- Young Leaf
- Intermediate Leaf
- Axis of the Inflorescence
- Carpel of the 6th/7th Flower

TAIR: (arabidopsis.org; blast.ncbi.nlm.nih.gov)

CRISPR: (http://www.e-crisp.org/E-CRISP/;)

Guide RNA Sequence

To remove a small section at the beginning of my gene:

Removes part of exons 4 and 6

1. Sequence: GATCTAGTGACACCAATCTA NGG
   Location: 1856 bp to 1899 bp
2. Sequence: GATTGGAACTACTAGACCAG NGG
   Location: 2382 bp to 2425 bp

To remove the whole gene:

Removes part of exon 4 to part of exon 19 (last exon)

1. Sequence: GATCTAGTGACACCAATCTA NGG
   Location: 1856 bp to 1899 bp
2. Sequence: GATCTAGTGACACCAATCTA NGG
   Location: 6693 bp to 6736 bp

CRISPR: (http://www.e-crisp.org/E-CRISP/;)

Summary

In order to verify if the gene was successfully knocked out, I would have to observe the phenotype of the newly grown plant. I know for sure that the gene was knocked out if there is no phenotype, because then I know that there was no mutation. However, if my plant did show a phenotype in the division plane orientation, I would have to determine whether my gene Katanin P80 Subunit 1 was responsible, or if TAN1 was unsuccessfully knocked out. I could use PCR to confirm that my gene was responsible for the mutation.

The result of having a non-functional target protein for my assigned gene is clarification on its role in cell division in plants. If my gene turns out non-functional, then I know that the Katanin P80 Subunit 1 gene is not important for division plane orientation, or that it has a closely related gene that accounts for its loss. We know that proteins used for division plane establishment differ from those used for division plane maintenance. Thus, if the maintenance pathway shows that the gene does not resemble the phenotype of the tan1air9 double mutant, we know our gene is not important for division plane orientation.

Rasmussen, Carolyn G, and Marschal Bellinger. "An Overview of Plant Division-Plane Orientation." *The New Phytologist*, U.S. National Library of Medicine, 27 Apr. 2018, www.ncbi.nlm.nih.gov/pubmed/29701870.

Future Directions

In order to determine whether my gene is functional or not, I will use agrobacterium. First, I will clone Cas9 and my gRNAs into a special plasmid, and insert it into the agrobacterium. I will let the bacteria infect the plant and insert Katanin P80 Subunit 1 into the plant genome. Then, using I will dip the flowers into a mixture of the agrobacterium, some sugar water and a little bit of detergent, to express the plasmid in the single cell egg. I will allow the plants to grow and will then collect the seeds. To find the seeds that were transformed, I will include a selectable marker with my gene. This marker could be an antibiotic or an herbicide. I will grow the collected seeds on a plate containing the selection. Whichever seeds grow and survive will be the ones transformed because it contains the resistance gene. Finally, to get a closer look at the phenotype, I will transfer the plants to soil and analyze their seeds. I will look out to see if any of the seedlings have a stunted growth phenotype.

Wang, Z., Xing, H., Dong, L. *et al.* Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biol* 16, 144 (2015). https://doi.org/10.1186/s13059-015-0715-0
https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0715-0

Appendix 24: Future directions

Future Directions

How to make a successful plasmid construct that can knock out your intended target gene/homolog?

We're going to clone the designed guide RNAs into a specific plasmid called pHEE104E. A special feature of this plasmid is that it uses a conditional promoter that drives expression of the Cas9 protein (which cuts the double stranded DNA) and the guide RNAs we made (which tells the Cas9 where to cut) only in the single cell egg.

Why do we want our Cas9 and guide RNAs only expressed in the single cell egg?

Creating mutations with CRISPR-Cas9 is not an exact process. The Cas9 protein will cut where we want it to based on our guides, but how the organism repairs that cut is error prone. Usually Non-Homologous End Joining repairs the double strand break caused by Cas9 but it creates small insertions and deletions in the process, resulting in many slightly different mutations using one pair of guide RNAs. As a result, when you have an active promoter that expresses your Cas9 protein and guides in many different cells you can end up with plants that are mosaic (have different mutations in different patches of tissues). This can make it problematic to characterize the phenotype of the mutant, because your plant is a mix of different types of mutations (some of which may be stronger than others).

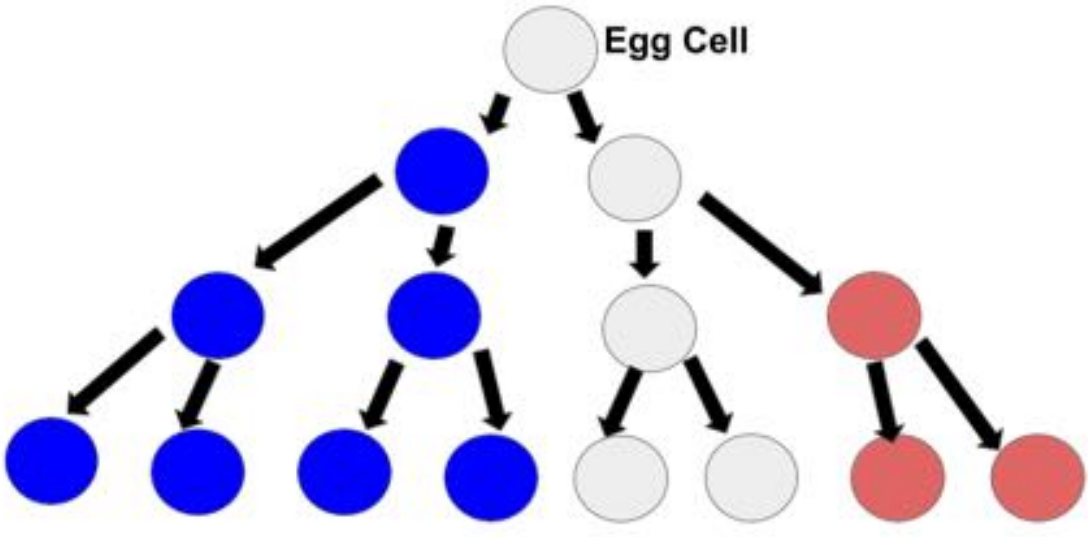


Figure 1. What can happen if you use a nonspecific promoter to drive expression of CRISPR-Cas9. Since the Cas9 is expressed all the time, this can lead to the creation of different variations of your mutated gene at different times in the plant's development. Here different variations of the mutation are indicated by blue and red. This leads to the creation of a

mosaic plant that has patches of different mutations and even non mutated patches (indicated in white). This makes it difficult to analyze the resulting plant.

Ideally we want the mutation to happen once, in the very first cell (the egg) so that all of the cells thereafter have the same uniform mutation. By using the egg cell specific promoter, our Cas9 protein and guide RNAs will only be active in the egg cell, giving us our desired uniform mutant.

This is the paper describing the development of the egg cell specific promoter plasmid. Feel free to take a look at it and/or cite it:

Wang, Z., Xing, H., Dong, L. *et al.* Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biol* 16, 144 (2015). https://doi.org/10.1186/s13059-015-0715-0
https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0715-0

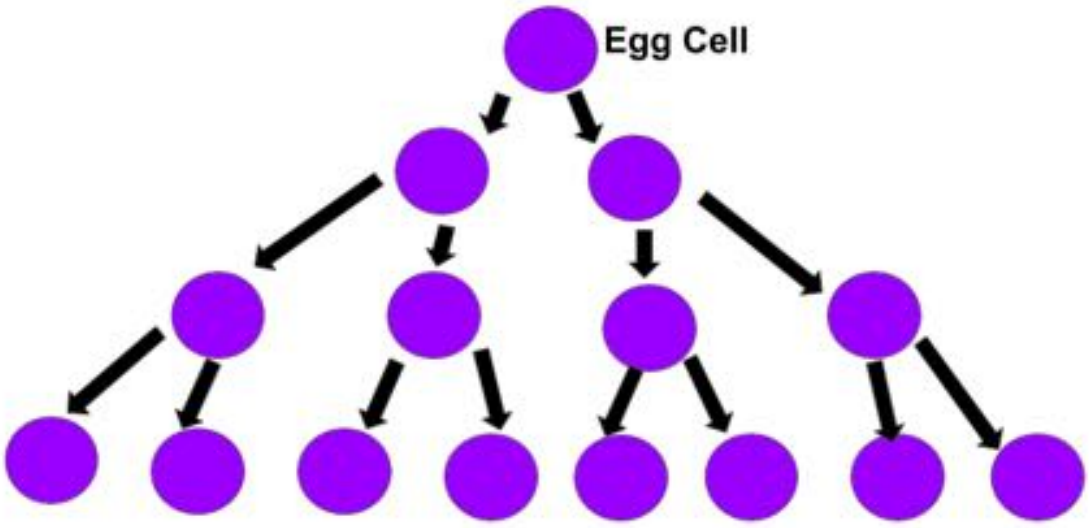


Figure 2. What happens if you use an egg cell specific promoter to drive expression of CRISPR-Cas9. In this case the Cas9 is only expressed in the egg cell. As a result, you only have a single mutation in the egg, which is then passed along to all the other cells of the plant as it develops. Here the mutation is indicated by the color purple. This is a much better plant to analyze because it is uniform (it has one mutation in all of its cells).

How is *Agrobacterium tumefaciens* used in the context of knocking out genes in the *Arabidopsis thaliana* genome?

Plant scientists have adapted a certain type of bacteria, called *Agrobacteria tumefaciens*, to introduce genes they want to study into plant genomes. *A. tumefaciens* infects plants and transfers a piece of its bacterial genome into the plants genome. In wild *A. tumefaciens*, the

genes they transfer make the plant synthesize nutrients required for growth. Scientist remove these nutrient producing genes, and swap in the genes they want to be introduced into the plant (in our case Cas9 and guide RNAs). The modified Agrobacteria then infect the plant and transfer the genes the scientist want into the genome of the plant.

It is easy to introduce new genes into *A. thaliana* because we can transform it using *A. tumefaciens*. After we introduce our gene of interest into the plasmid and into *A. tumefaciens*, we mix the *A. tumefaciens* with some sugar water and a little bit of detergent (to remove some of the protective wax on the plant). Then we dip the flower buds of the plant into this solution. In our case we'll dip the flowers of *air9* single mutants. We want to dip the flowers because this structure contains the egg cells of the plant. As described above, we want our CRISPR-Cas9 construct to be introduced into egg cells, so the majority of cells that generate the offspring plant contain the same mutation. So after we dip the plants, we allow them to grow and then we collect the seeds. *A. thaliana* produces thousands of seeds while the transformation process is rare. In order to find which of those few seeds were transformed with our gene of interest, we include a selectable marker with our gene of interest. This marker is usually an antibiotic or herbicide resistance gene. Then (similar to how we select for transformed bacteria) the collected seeds are grown on a plate containing the selection (either antibiotic or herbicide). The transformed seedlings that survive and grow (because they have our gene of interest and the accompanying resistance gene) are then transferred to soil so they can grow and produce seeds. We can then analyze the seeds produced from this plant to see if any of the seedlings have a stunted growth phenotype.

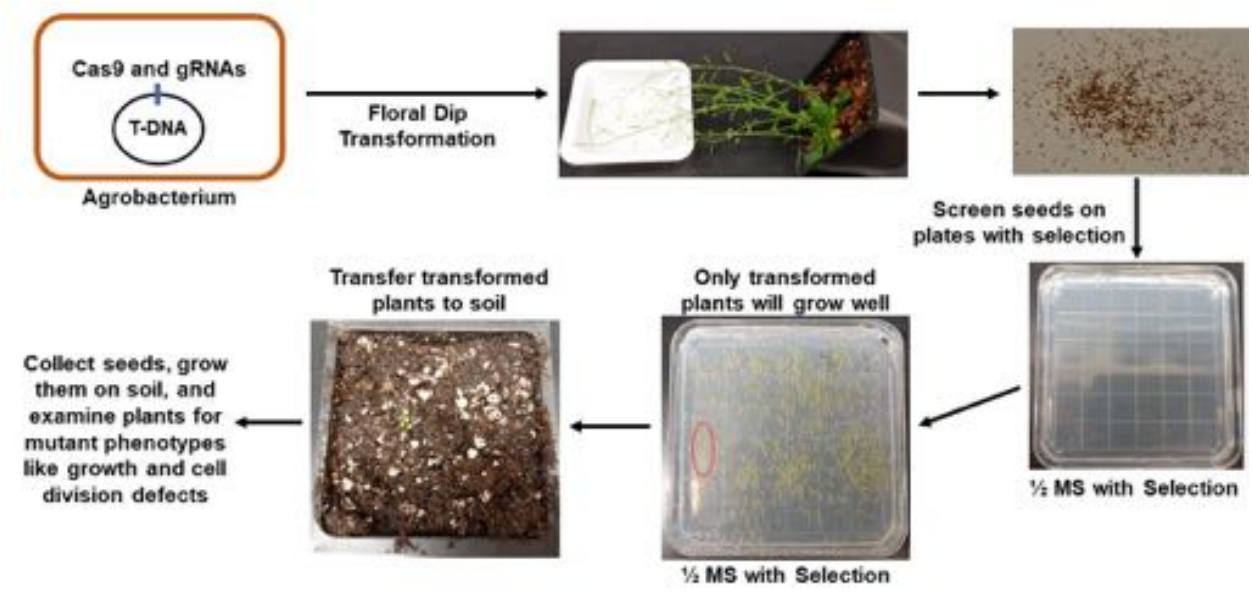


Figure 3. Diagram of *A. tumefaciens* mediated transformation by floral dip. We clone our gene of interest (in our case Cas9 with our guide RNAs) into a special plasmid (indicated as the T-DNA plasmid above). We then put this plasmid into *Agrobacterium tumefaciens*. Then we dip the plants in media containing *A. tumefaciens*, which inserts the gene of interest into the plant genome.

How would having a plant with a non-functional target protein for your gene help advance the study on cell division in plants?

We want to examine what happens to the growth of the plant when we mutate our assigned gene in the *air9* single mutant. As a review, the *air9* single mutant grows normally and the *tan1 air9* double mutant has a stunted growth phenotype due to division plane orientation defects. So, we're essentially creating our own new double mutants. For example, if I was assigned the gene *CEN2*, the guide RNAs I made are going to be used to eventually create a *cen2 air9* double mutant. If the *cen2 air9* double mutant resembles the *tan1 air9* double mutant with a stunted growth phenotype and misplaced cell walls, that would suggest that CEN2 is required for division plane orientation in parallel with AIR9. If the *cen2 air9* double mutant grows normally, we'll know that CEN2 either isn't important for division plane orientation or maybe it has a closely related gene that compensates for its loss (this is part of why it's helpful to check for paralogs).

Appendix 25: Example student report

This student work was used with expressed written consent of the student.

Investigating Division Plane Orientation in *A. thaliana* using Katanin P80 Subunit 1 Mutant

Cell division is critical to the life and functionality of multicellular organisms like plants and animals. It is a regulated process that controls the number of cells created and their geometry. The cell cycle is controlled by external and internal factors. Uncontrolled cell division is identified by a disease called cancer, creating tumors that interfere with biological processes and pathways (Sherr, 1996). Cancer is caused by the damage of genes that control cell regulation (Sherr, 1996). Unlike animal cells, plant cells have cell walls that are stiff and lock the cell in place, making geometry a basic element of how they choose their division plane (Guérin et al., 2016). In addition, symmetrical divisions are essential for growth for the majority of plant cells, as asymmetrical divisions are initiated only through specific signaling (Rasmussen, et al 1). Division plane orientation is the manner and organization in which new cell walls are created in plants (Rasmussen and Bellinger, 2018).

In this investigation, the division plane orientation of *Arabidopsis thaliana* in the *TAN1* pathway with the *tan1* gene mutant will be examined. The wild type division plane orientation of *Arabidopsis thaliana* has an organized and symmetrical character. *TAN1* is involved with the preprophase band (PPB). The PPB determines the new location of the cell wall and it dissociates before mitosis. Because of this, *TAN1* interacts with the microtubules and will keep the information from the PPB to hand over to phragmoplasts to develop their own microtubules (Rasmussen and Bellinger, 2018). When *tan1* is mutated in *Arabidopsis thaliana*, there is no significant change in the division plane orientation (Mir et al., 2018). The *air9* gene mutant is another pathway that shows a wildtype phenotype as well (Mir et al., 2018). However, when *tan1* and *air9* are combined to make a synthetic mutant, is created, the division plane orientation is changed drastically and is asymmetrical unlike the wildtype (Mir et al., 2018). According to "An overview of plant division-plane orientation," around half of the divisions in the *tan1* and *air9*

double mutant were completed in a different location than the PPB (Rasmussen and Bellinger, 2018)

Arabidopsis thaliana will be used in this research because it is a model organism. A model organism is a species that has been widely studied that has many resources provided for the scientific community. One of the advantages of using *Arabidopsis thaliana* as a model organism is that it is a small plant that makes it easier to manage (NSF, 2002). It also has a short generation time, or life cycle, which is ideal for several experiments or research (NSF, 2002). In addition, it has a nicely sequenced genome and many bioinformatics tools have resources available for research such as TAIR, The Arabidopsis Information Resource site. It can easily be manipulated for genetics research than any other plant (NSF, 2002).

CRISPR Cas9 is a genome editing program that allows you to modify genes and their functions of in organism's processes. It allows you to add, remove, and alter parts of the genome (GHR, 2017). Currently, it is popular for further research in medical applications like gene therapy and treating diseases (GHR, 2017). In this investigation, we will use it to identify genes that are involved in division plane orientation in the TAN1 pathway by creating mutants from specifically targeted genes. To achieve this, guide RNAs and primers will be designed to target specific areas that will be input into CRISPR. Afterwards, they will be cloned into a vector that includes coding sequences and selectable markers and then further examined to see if the mutant will create a strong phenotype in division plane orientation in comparison to the wild type of *Arabidopsis thaliana.*

Methods

Basic Primer Design

The gene for this experiment is KATANIN P80 SUBUNIT 1. The first step is to visit TAIR and to input the gene ID (AT1G11160) and to select the correct locus that matches. There are several pieces of information on the page about the assigned gene. From TAIR, the CDS and genomic DNA sequences are obtained after clicking "Send to BLAST" to get a text form to be

recorded in a document. In addition, the functions of the gene and an atlas of where it is expressed was collected from the gene's front page.

The genomic DNA and CDS sequences are inputted into Nucleotide BLAST after clicking "Align two or more sequences." The gDNA sequence is the query sequence where the CDS is the subject sequences. The graphic summary was kept to later make a gene map and the alignments of exons, sorted by query position, were recorded. The gene map was based on what was displayed on BLAST and recreated on draw.io. To be able to tell between the cDNA and gDNA sequence, a section with a large intron is preferred for primers to target.

Now in primer BLAST, the primers will be designed. The targeted area was from exons 7-9. The forward primer input was 2270-2300 and the reverse primer was 3405-3435. The primer temperatures were of a minimum of 57.0, an optimum of 60.0, and maximum of 63.0. The maximum product size was 1500. The "Enable search for primer pairs specific to the intended PCR template" option was deselected. After clicking "Get Primers" the graphic summaries of the primers and detailed primer reports should appear. The product length should be displayed as the last item on each primer report and when the length of the introns is subtracted from the product length it will give you the cDNA length. The primer pair chosen is recorded into a table.

Phylogeny

Plaza, an interactive phylogenetics module, will be used to learn more about evolution and relation to other genes in similar species through creating a phylogenetic tree. At the top, the gene ID, AT1G11160, is input and *Arabidopsis thaliana* is selected in the drop down menu where it says "All species." The front page of the gene should show its structure and its functions. Under "Create" the phylogenetic tree can be created. The species selected were *Arabidopsis lyrate, Arabidopsis thaliana, Citrus clementina, Oryza sativa,* and *Zea mayz.* The phylogenetic tree shows various branches and the paralogs and orthologs can be identified in relation to the assigned gene. The phylogenetic tree is screenshotted to keep.

Guide RNA Design

The guide RNAs will be designed in E-CRISP to remove parts of genes to mutant in CRISPR Cas 9. The Gene ID is pasted into the site and “strict” is selected at the bottom. Multiple guide RNAs will appear with their sequences and hits. Pairs of guide RNAs were chosen to remove small sections and the entire gene. If there is more than one hit on the guide RNA it was avoided unless it included part of another gene and had mismatches. Guide RNAs without the PAM sequence (NGG) were avoided because then they would not be spliced correctly by CRISPR. Two pairs of each for small section and entire gene were selected and then a favorite from each category was recorded.

Data/Results

Basic Primer Design

The name of the gene is KATANIN P80 SUBUNIT 1 and its gene ID is AT1G11160. It is a gene in *Arabidopsis thaliana* and according to Figure 1, obtained from TAIR, it is involved in microtubule severing and organization. Figure 2 is an Atlas that displays where the gene is expressed and the absolute areas of expression in read are in the roots, axis of inflorescence, and young flower stam. This makes sense because microtubules are present in division plane orientation for the division of cells, which is crucial for these major areas of growth and development. This is important information as the purpose of the project is to investigate genes and their involvement in division plane orientation. Figure 3 is a gene map of AT1G11160 that shows that there are 18 exons and 17 introns. This graphic summary helped decide where the primers were going to target. The primers for this experiment are over exons 7-9 because there is a large intron between exons 8 and 9. Therefore, it will be easy to distinguish between the gDNA and cDNA if ran through a PCR gel. The input of the forward primer was 2270-2300 while the reverse primer was 3405-3435. There were 10 primer results, as displayed in Figure 4, and the first pair of primers were chosen. Table 1 and Table 2 display detailed information about the primers chosen. The lengths of the primers were 20 and 23 base pairs. The product length of

the genomic DNA is 1159, so the length of the cDNA is 570 so there is a significant difference between the lengths between the two.

Phylogeny

The phylogenetic tree for the gene AT1G11160 in *Arabidopsis thaliana* was created using Plaza with the species *Arabidopsis lyrate, Arabidopsis thaliana, Citrus clementina, Oryza sativa,* and *Zea mayz* as displayed in Figure 6. The *Arabidopsis lyrata* gene, AL1G22180, was an ortholog- as indicated in Figure 7 where the orthologs and paralogs are shown. However, there are several branches that are closely related but are only paralogs. Figure 5 shows a graphic summary of the gene on Plaza and the locations of exons, UTRs, and start/stop codons.

Guide RNA Design

There were four pairs of guide RNAs that were chosen in total through the bioinformatics program, E-CRISP. There were several results as depicted in Figure 8. There were two pairs for small sections of gene: Pair 1: Exons 4-6 (Figure 9) and Pair 2: Exons 11-13 (Figure 10). The two pairs to remove the entire gene were Pair 1: Exons 4-18 (Figure 11) and Pair 2: Exons 6-18 (Figure 12). Unfortunately, there were no guide RNAs that were present for Exons 1-3 with the criteria that was selected for this experiment. The favorite guide RNAs that were chosen for CRISPR Cas9 are the first pairs for both categories.

Conclusion and Future Directions

After creating the guide RNAs, a plasmid needs to be made to remove the assigned gene from the *Arabidopsis thaliana* plant. This plasmid is called the pHEE104E plasmid and it contains conditional promoters that express cloned guide RNAs and prompts the Cas9 protein that splices the gene (Wang et al., 2015). These conditional promoters are a crucial part of the plasmid because it only is driven in the single egg cell (Wang et al., 2015). This ensures that there is a uniform and consistent mutation throughout the plant instead of patches in different tissue, which makes it much easier to identify the resulting phenotype of the plant (Wang et al., 2015). When a successful plasmid is created, it is inserted into *Agrobacterium tumefaciens* by a

restriction enzyme and DNA ligase (Wang et al., 2015). *Agrobacterium tumefaciens* is a type of bacteria created by plant scientists that can infect plants and transfer their genome into them (Wang et al., 2015). This is important for the *Arabidopsis thaliana* plant the seeds produced from a floral dip transformed plant will be grown and the new phenotype with the altered gene can be observed. A plant with a non-functional target protein for the KATAIN P80 SUBUNIT 1 gene would help advance the study on cell division because it could possibly show an asymmetrical division plane orientation in the roots and reproductive parts where it is expressed the most. As a result, if it shows this strong phenotype then it could be verified that it is involved in the air9 background and tan1 pathway which can help further study the synthetic mutant (Rasmussen and Bellinger, 2018).

Works Cited

Guérin, Adrien, et al. “Forces behind Plant Cell Division.” *Proceedings of the National Academy of Sciences*, vol. 113, no. 32, Aug. 2016, p. 8891, doi:10.1073/pnas.1609309113.

Louveaux, Marion, et al. “Cell Division Plane Orientation Based on Tensile Stress in Arabidopsis Thaliana</Em>.” *Proceedings of the National Academy of Sciences*, vol. 113, no. 30, July 2016, p. E4294, doi:10.1073/pnas.1600677113.

Mir, Ricardo, et al. “Division Plane Orientation Defects Revealed by a Synthetic Double Mutant Phenotype.” *Plant Physiology*, vol. 176, no. 1, Jan. 2018, pp. 418–31. *DOI.org (Crossref)*, doi:10.1104/pp.17.01075.

NSF. “Arabidopsis: The Model Plant.” *The Multinational Coordinated Arabidopsis Thaliana Functional Genomics Project*, June 2002, https://www.nsf.gov/pubs/2002/bio0202/model.htm.

Rasmussen, Carolyn G., and Marschal Bellinger. “An Overview of Plant Division-Plane Orientation.” *New Phytologist*, vol. 219, no. 2, 2018, pp. 505–12. *Wiley Online Library*, doi:10.1111/nph.15183.

Reference, Genetics Home. “What Are Genome Editing and CRISPR-Cas9?” *Genetics Home Reference*. *ghr.nlm.nih.gov*, 2017. https://ghr.nlm.nih.gov/primer/genomicresearch/genomeediting.

Sherr, Charles J. “Cancer Cell Cycles.” *Science*, vol. 274, no. 5293, Dec. 1996, p. 1672, doi:10.1126/science.274.5293.1672.

Wang, Z., Xing, H., Dong, L. et al. Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biol 16, 144 (2015). https://doi.org/10.1186/s13059-015-0715-0

https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0715-0

Figures and Tables

Annotations	category	relationship type	keyword
	GO Biological Process	involved in	cortical microtubule organization, microtubule depolymerization, microtubule severing
	GO Cellular Component	is subunit of	katanin complex
	GO Cellular Component	located in	Cul4-RING E3 ubiquitin ligase complex, katanin complex, nucleus
	GO Molecular Function	has	microtubule-severing ATPase activity
	Growth and Developmental Stages	expressed during	LP.04 four leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage
	Plant structure	expressed in	carpel, collective leaf structure, flower, flower pedicel, hypocotyl, petal, petiole, plant embryo, root, seed, sepal, shoot apex, shoot system, stem, vascular leaf

Figure 1: Multiple functions of KATANIN P80 SUBUNIT 1 gene from TAIR. AT1G11160 is involved in microtubule severing and organization in several growth and development stages.

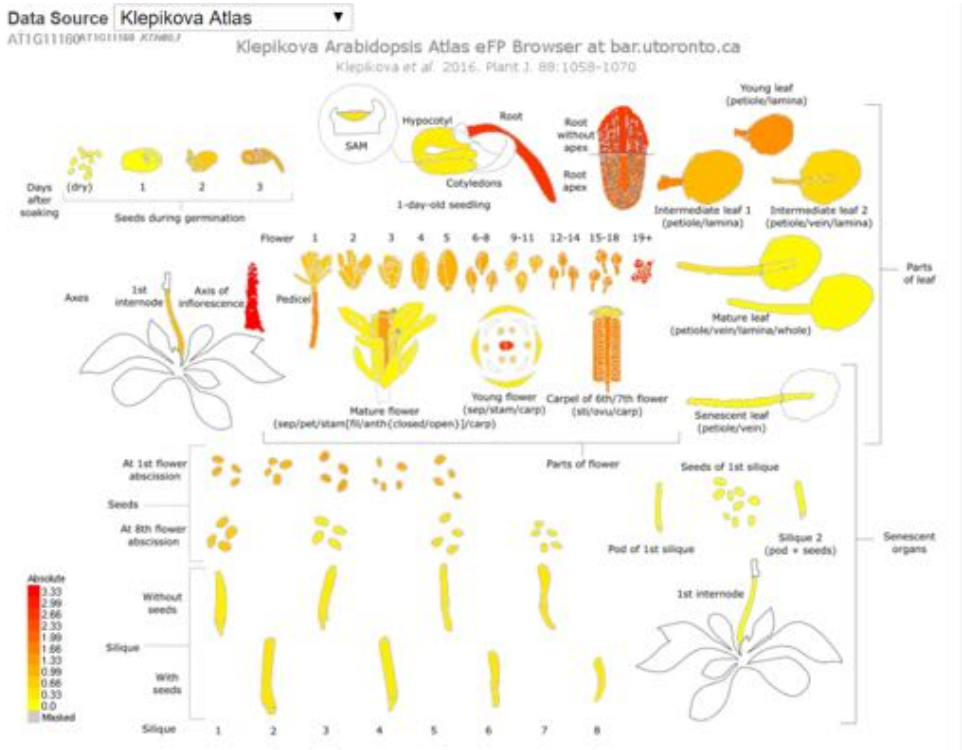


Figure 2: Atlas of where gene is involved in *Arabidopsis thaliana,* where the sections in red are where the gene is absolutely expressed

Figure 3: Gene map of KATANIN P80 SUBUNIT 1, displays 18 exons and 17 introns- In addition to the forward and reverse primers chosen

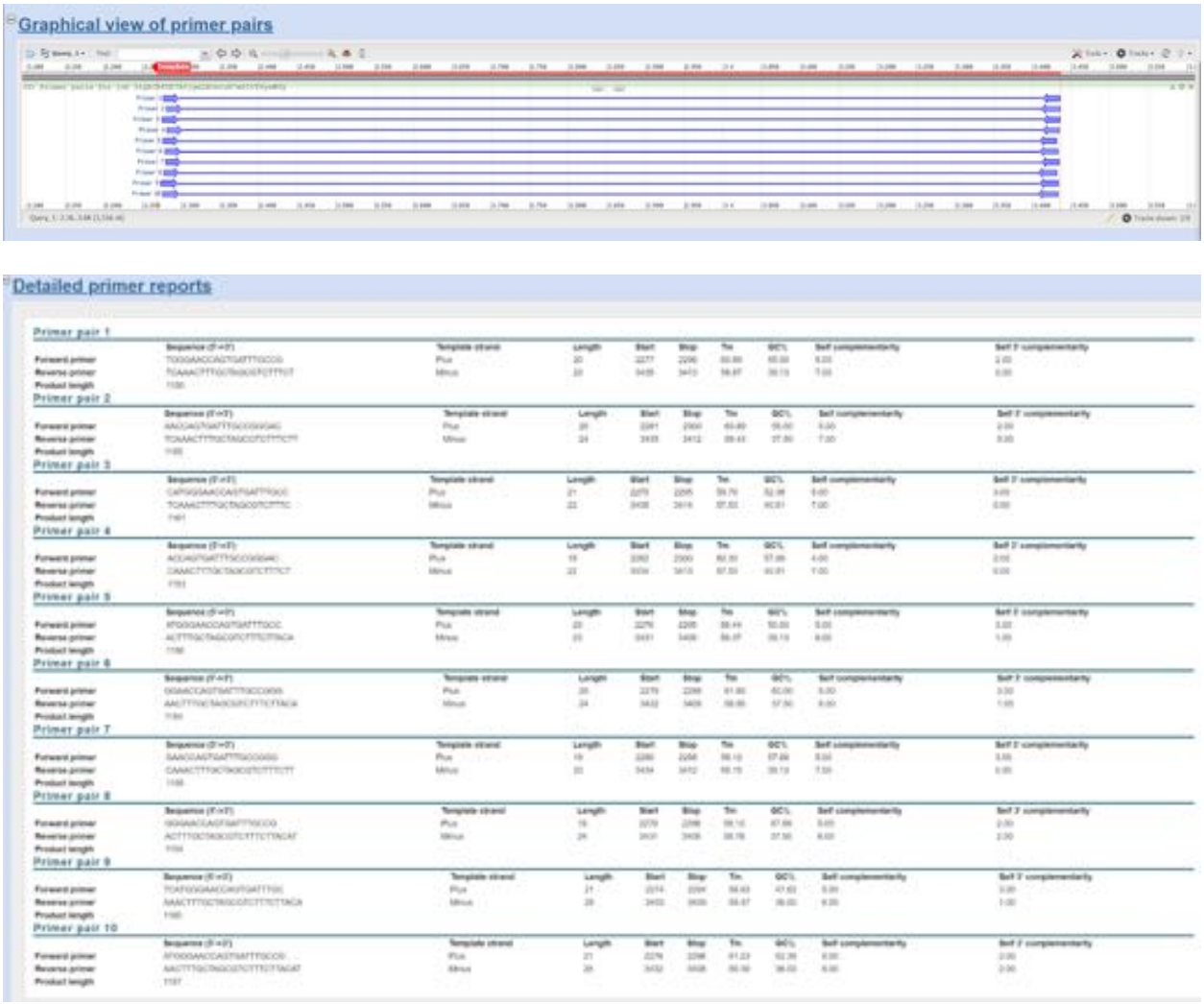


Figure 4: All Primer Results- graphic summary (top) and detailed summary of sequences and length (bottom)

Table 1: Forward primer

Sequence (5'→3')	Length	Start	Stop	Tm	GC%	Self-complementarity	3' complementarity
TGGGAACCAGTGATTTGCCG	20	2277	2296	60.89	55.00	5.00	2.00

Table 2: Reverse primer

Sequence (5'→3')	Length	Start	Stop	Tm	GC%	Self-complementarity	3' complementarity
TCAAACTTTGCTAGCGTCTTTCT	23	3435	3413	58.87	39.13	7.00	0.00

Phylogeny (Plaza)

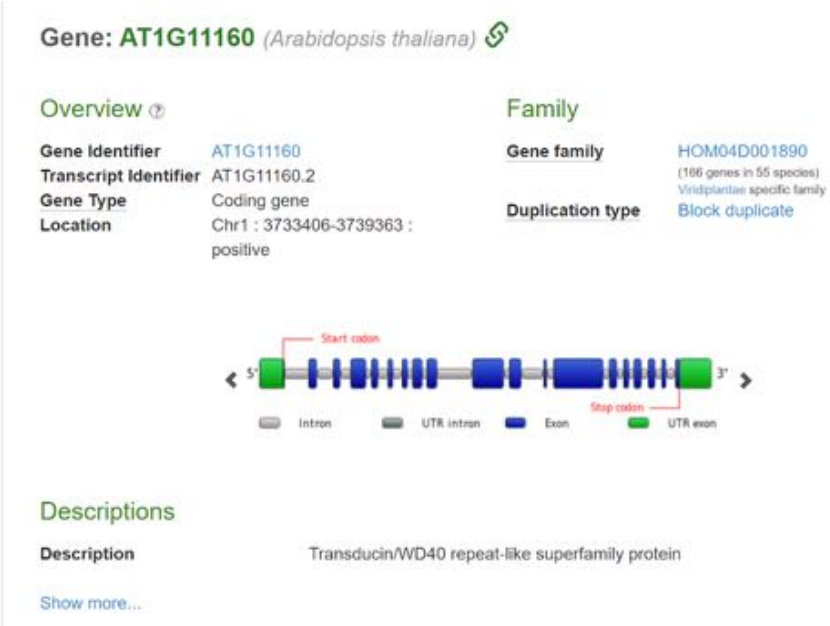


Figure 5: Gene Expression Diagram, also shows the exons, UTRs, and start/stop codons

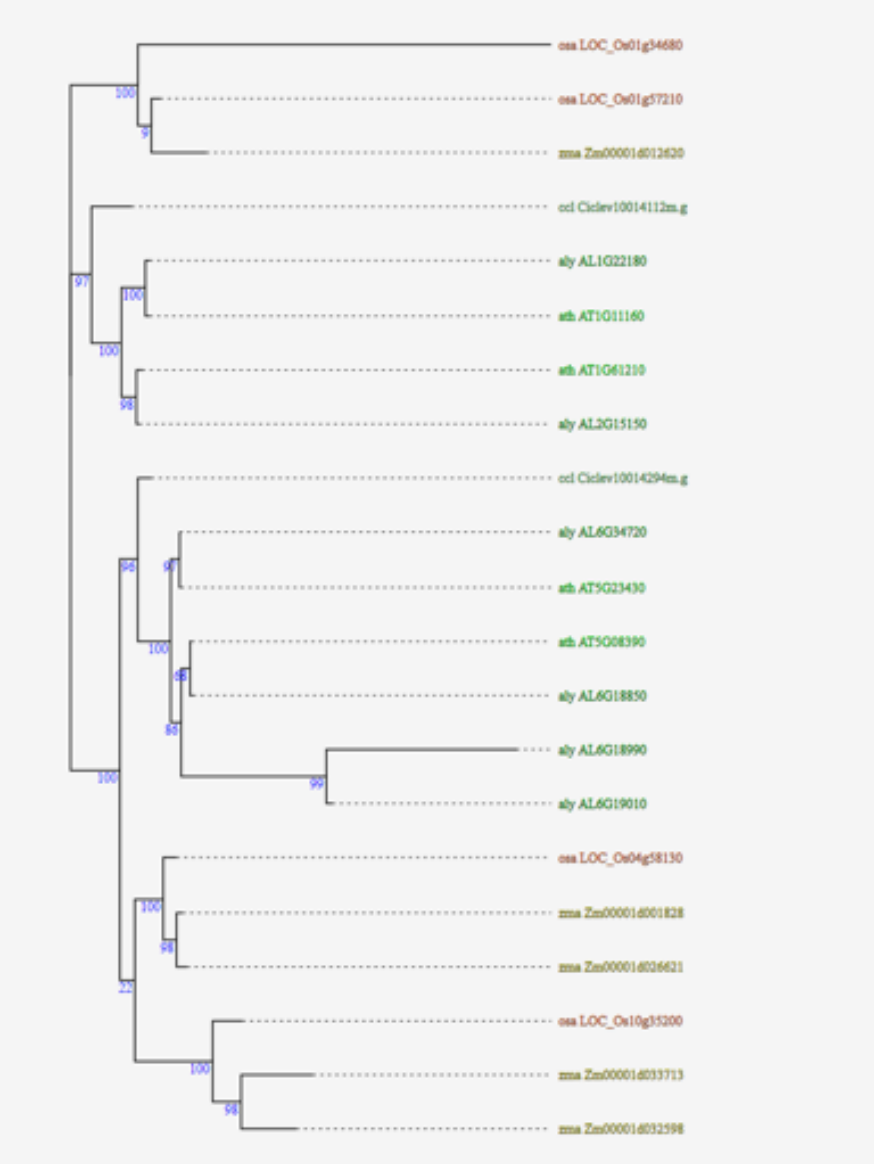


Figure 6: Phylogeny Tree with *Arabidopsis lyrate, Arabidopsis thaliana (*AT1G11160)*, Citrus clementina, Oryza sativa,* and *Zea mayz*

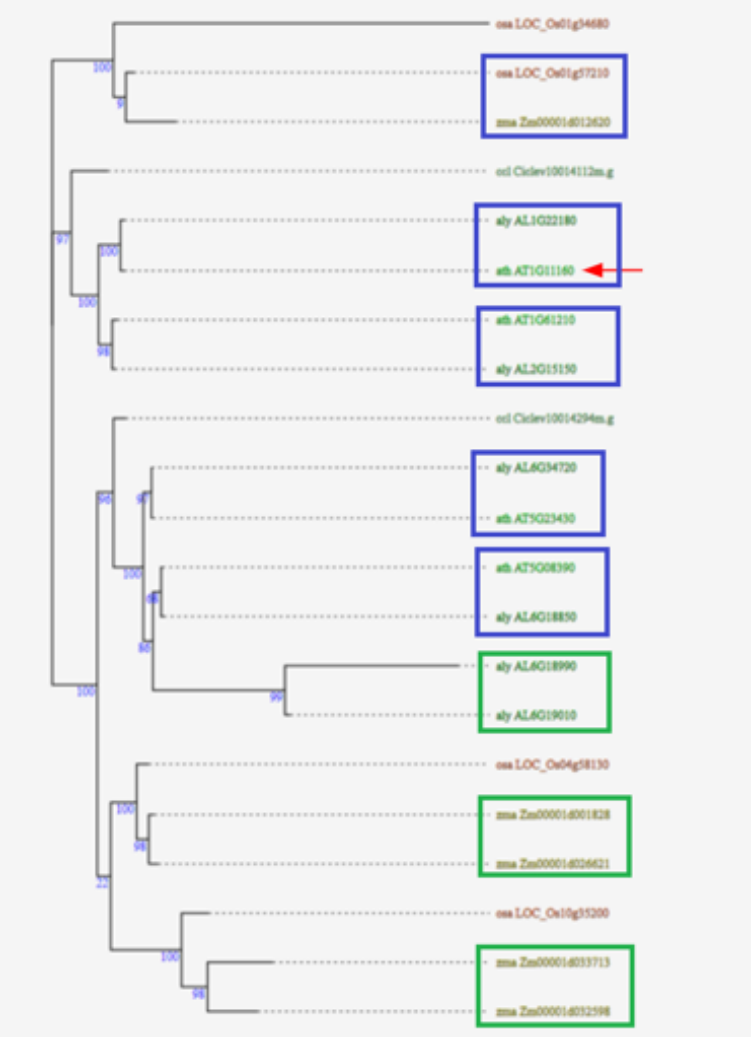


Figure 7: Inked Phylogeny Tree, indicates orthologs (blue), paralogs (green), and AT1G11160 gene (red arrow)

Guide RNA Tables & Figures

Figure 8: E-CRISP guide RNA results for “strict” option

Figure 9: First pair for small section, Exons 4-6 (Favorite pair), if mutated it should change the rest of the gene's expression completely

Figure 10: Second pair for small section for Exons 11-13

Figure 11: First Pair to remove entire gene, Exons 4-18 (Favorite pair), removes almost the entire gene- only leaving out exons 1-3, and 19.

Figure 12: Second pair for removing entire gene, Exons 6-18

Appendix 26: Project Report Rubric

Additional Notes:

-The multi-point categories are not all or none, partial credit can be awarded.
-The information does not to be housed in each of the categories above. Points should be awarded if it is organized differently but the information is included.

Introduction Draft Rubric

Categories	Points	Justification
Format	3	The report is written as specified (not an outline). The file format is appropriate and accessible e.g. doc, docx (hand-written, google slides, pdf, PowerPoint are not).
Title	1	Title should be relevant to the project.
Background	4	Frames biological question/purpose of project. Introduces model organism and research topic.
Experimental Design	3	Student explains what is being tested and addresses how the project attempts to answer the biological question.
References/Citations	2	Includes required primary literature references. Correctly uses in-text citations and reference section.
Cohesive/Clarity	1	The report is written with a logical flow and using appropriate language.
Complete	1	The report contains all the required elements.
Total	**15**	

Sample scoring:

15 points for an outstanding draft.
12 points for a good draft that might need to improve on one or two of the categories.
10 points for an okay draft that is missing major categories or requires significant improvements.
5 points for a weak draft that is missing major elements and needs a complete overhaul.

Methods & Results Draft Rubric

Categories	Points	Justification
Methods	3	Methods used over the course of the project are concisely summarized in complete sentences.
Figures	3	This should include items like gel images, sequences, photos, chromatograms, etc. The figures should be annotated and include a figure legend.
Data analysis	4	At a minimum, students should discuss the results of the included figures and how they relate to the goal of the project.
Accuracy/Relevance	2	The data included is relevant to the project and correctly labeled/identified.
Format	1	The report is written as specified (not an outline). The file format is appropriate and accessible e.g. doc, docx (hand-written, google slides, pdf, PowerPoint are not).
Cohesive/Clarity	1	The report is written with a logical flow and using appropriate language.
Complete	1	The report contains all the required elements.
Total	**15**	

Sample scoring:

15 points for an outstanding draft.
12 points for a good draft that might need to improve on one or two of the categories.
10 points for an okay draft that is missing major categories or requires significant improvements.
5 points for a weak draft that is missing major elements and needs a complete overhaul.

Final Report Rubric

Categories	**Points**	**Justification**
Revised Drafts	10	Student thoroughly revised their drafts and addressed correction/suggestions that were provided on the drafts. Drafts were combined into a cohesive report
Data	10	Figures and legends are included for data acquired during the project (PCR data, sequences, guide RNA sequences, etc.)
Data Analysis/Interpretation	10	Student has analyzed results correctly, and their analysis relates to the goal of the project.
Future Directions & Conclusion	10	Student has provided a conclusion and discusses future directions.
References	5	Student used references and provided citations.
Format	5	Student formatted the report properly.
Cohesive/Clarity	10	The report has a logical flow and makes sense.
Complete	10	Student correctly addressed of the categories included in the document.
Total	**70**	

Sample scoring:
70 points for an outstanding report.
60 points for a good report that might need to improve on one or two of the categories.
50 points for an okay report that is missing major categories or requires significant improvements.
30 points for a weak report that is missing major elements and needs a complete overhaul.

Appendix 27: Final exam

Final Exam

Instructions

This exam is to be completed by you alone. You are not to discuss or receive help of any kind form another person except a Biology 20 instructor or TA. Please type your name here to indicate that you did not receive, nor did you give help on this assignment.

Name:

The is an open note exam. You may not discuss it with another person (except Instructors and TAs). You may use the web.

All questions need to be answered in complete sentences and all calculations much be shown. Please use BLACK font color for answering. All questions must be typed. Equations must be typed, using the Word Equation Editor. Drawings must be done electronically (e.g., Google Draw) and inserted as an image. The answer length maximum is provided for each question. No hand written answers or drawings will be accepted.

Proper formatting of gene and protein names is required.

1. (12 points) A gene from rice (*Oryza sativa*) will be assigned for this exam. In addition to NCBI and others used in class you will need to use these as well:

Oryzabase

Rice Genome Annotation project (RGAP)

Your gene is: LOC_Os07g31450

A. The gene name and chromosomal location. (Oryzabase)
B. Number of exons and introns. (RGAP)
C. Describe the expression pattern. This will require some research as there is no convienent Atlas.

2. (15 points) Design PCR primers to detect the full gene and the mRNA. Draw an image (do not copy and paste the Primer Blast output) showing the location of the primers and the expected PCR product size.

3. (12 points) Generate a phylogenetic tree using *Orzya sativa, Zea mays*, and *Glycine max*. Modify the number of species selected if the number of genes are greater than 6.

A. Circle paralogs in BLUE for the three species. If there are no paralogs state why you concluded that.
B. Circle the orthologs to the rice gene in RED.

4. (12 points) Generate a multiple sequence alignment with 6 different genes and discuss the types of polymorphism you find.

5. (14 points) Design guide RNAs that will remove the whole gene (but make sure your primers will still bind). Draw (not copy and paste) an image with guide locations and show the size of the deletion. Justify why you chose those two guides.

6. (12 points) Design a PCR experiment that would show that the CRISPR guide RNAs were effective in making the deletion. Include unaltered genomic and mRNA in the experiment as well. Set up a PCR table including the proper controls. You have 2 X Taq mix, primers are 21 µM and you need 1 µM for each reaction. All DNA will be used at 5 µl in a 68 µl reaction. Show all of your work.

7. (15 points) Using a computer drawing program or the Word drawing tools, draw the results. Include a ladder lane and band sizes.

8. (8 points) Based on the available information about the gene predict the phenotype if the CRISPR experiment generated the deletion mutant. (Hint: Each gene has an image on Oryzabase.)

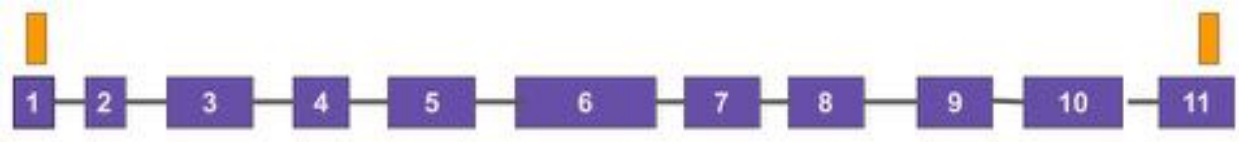
Appendix 28: Final Exam Answer Key

1. A gene from rice (Oryza sativa) will be assigned for this exam. In addition to NCBI and other used in class you will need to use there as well:

   Oryzabase
   RGAP
   Your gene is: LOC_Os07g31450
   A. Rolled Fine Striped Leaf, chromosome 7
   B. 11 exons, 10 introns
   C. Leaves at 20 days, pistil, and anther

2. Design PCR primers to detect the full gene and the mRNA. Draw an image (do not copy and paste the Primer Blast output) showing the location of the primers and the expected PCR product size.

Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC %	Self complementarity	Self 3' complementarity	
Forward primer	TGCAGGCTAATTGTGTTGGTG	Plus	21	56	76	59.39	47.62	4.00	0.00
Reverse primer	TTGCCGCACCTGAATTTTCG	Minus	20	12827	12808	60.04	50.00	5.00	2.00
Product length	12772								



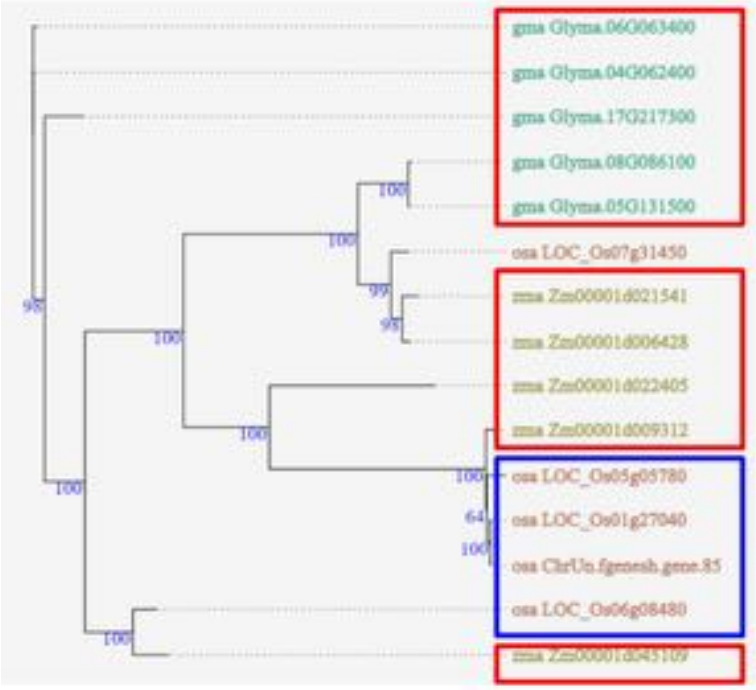
Approximate primer locations indicated by orange boxes.
Exon 1: 1-90
Exon 2: 1698-1800
Exon 3: 1900-2543
Exon 4: 3478-3554
Exon 5: 3727-4410
Exon 6: 4478-6428
Exon 7: 6766-6922
Exon 8: 7639-7758
Exon 9: 9728-9797
Exon 10: 9916-10510

Exon 11: 10787-12895

Genomic: 12772 base pairs
mRNA/CDS: 6456 base pairs

3. Generate a phylogenetic tree using *Oryza sativa, Zea mays*, and *Glycine max*. Modify the number of species selected if the number of genes are greater than 6.

   A) Circle paralogs in BLUE for the three species. If there are no paralogs state why you concluded that.
   B) Circle the orthologs to the rice gene in RED.



6. Generate a multiple sequence alignment with 6 different genes and discuss the types of polymorphisms you find.

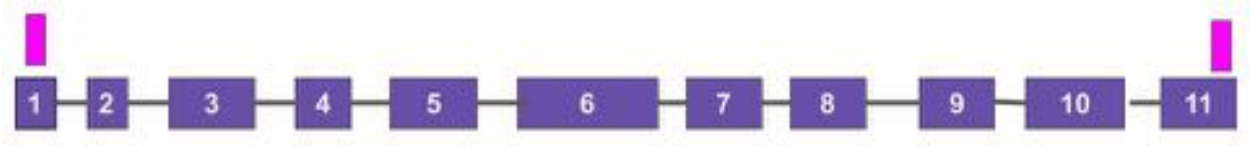
```
LOC_Os03g22900.1    AGCCAGCGATGGAAGACATTACTGAG-TTTTAACT-----GTCGAAATCGCTTGCTTCTC
LOC_Os06g01320.1    TGC------CTTGAACAACTTAAGCA-TCTTTCTA-----CCAATTTTAGGATGGTTCTT
LOC_Os06g08480.1    TCCAAGTTGTTCGGTCAACTTAAAGA-GTATCATA-----CTAAGCATCGTGTTCTCTTA
LOC_Os03g01200.1    AGCGTACTGTATAATGTCCTTGAGCA-ACGCTTTATCATGCCAAGA--CGTCTGCTACTA
LOC_Os07g31450.1    AGCAAACTTTTCAGTTTACTCA-----ATACATTATCAT-TTCAGCATAGAGTTTTGCTG
LOC_Os07g46590.1    GCATCGCTGTATACGACTCTTTTGGA-ATTCAGCA-----CTAAGAATAAGCTTCTCATC
                                 *                                 *  *  *
```

I aligned paralogs of LOC_Os07g31450. The dashed regions “---” indicate indels. Areas with “*” match perfectly across the 6 sequences. If these sequences were more closely related, regions without an “*” that are not indels could be SNPs.

7. Design guide RNAs that will remove the whole gene (but make sure your primers will still bind). Draw (not copy and paste) an image with guide locations and show the size of the deletion. Justify why you chose those two guides.

LOC_Os07G31450.1_0_0	GTGCAGGCTAATTGTGTTGG NGG
LOC_Os07G31450.1_12_5634	GACAGCTGACGAAGATAAGG NGG

The deletion is approximately 6400 base pairs



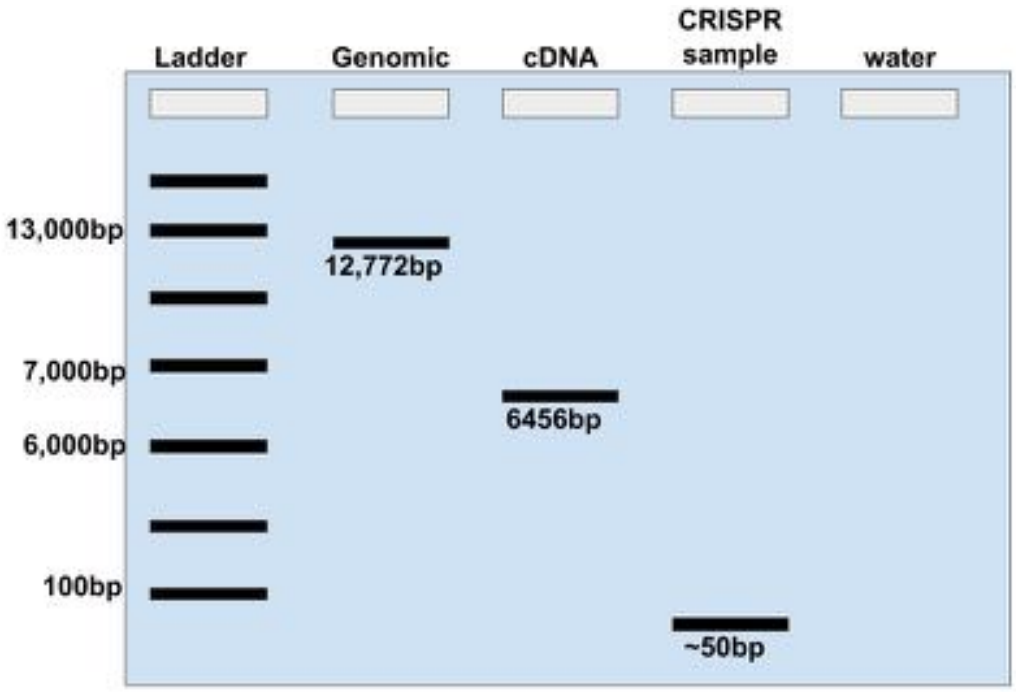
Pink boxes indicate approximate locations of the guide RNAs. I chose them because they were at the very ends of the gene so the whole gene could be cut out, and they were specific to just my gene of interest.

8. Design a PCR experiment that would show that the CRISPR guide RNAs were effective in making the deletion. Include unaltered genomic and mRNA (cDNA) in the experiment as well. Set up a PCR table including the proper controls. You have 2 X Taq mix, primers are 21 µM and you need 1 µM for each reaction. All DNA will be used at 5 µl in a 68 µl reaction. Show all of your work.

Reagent	1 reaction (µl)	5 reactions (µl)
MyTaq 2x	34	170
Forward primer	3.2	16
Reverse primer	3.2	16
water	22.6	--
DNA	5	113
Total	68	315

Number of reactions in master mix: 3 DNA samples + 2 (one water negative control and 1 for pipetting error)
C1V1=C2V2
MyTaq: 2x * V1 = 2x * 68µl
V1 = 34 µl
Primers: 21 µM * V1 = 1 µM * 68µl
V1 = 3.2 µl
Water: 68 µl - 34 µl – 3.2 µl – 3.2 µl - 5 µl = 22.6 µl

9. Using a computer drawing program or the Word drawing tools, draw the results. Include a ladder lane and band sizes.



10. Based on the available information about the gene, predict the phenotype if the CRISPR experiment generated the deletion mutant.

If the gene is deleted, the mutant plant will have rolled up leaves.

Appendix 29: Final Exam Rubrics

Exam 2 Spring 2019
100 Points

Answer all questions in the space provided and within the lined border. Any writing outside of the space provided will not be graded. Write legibly.

You may not use a computer, tablet, or phone during the exam. Violation of any of those conditions will result in a zero (0) grade for this exam. Use only calculators provided by the instructors.

If the question asks for an experimental design do not provide a step-by-step set of instructions for how to set-up a PCR experiment.

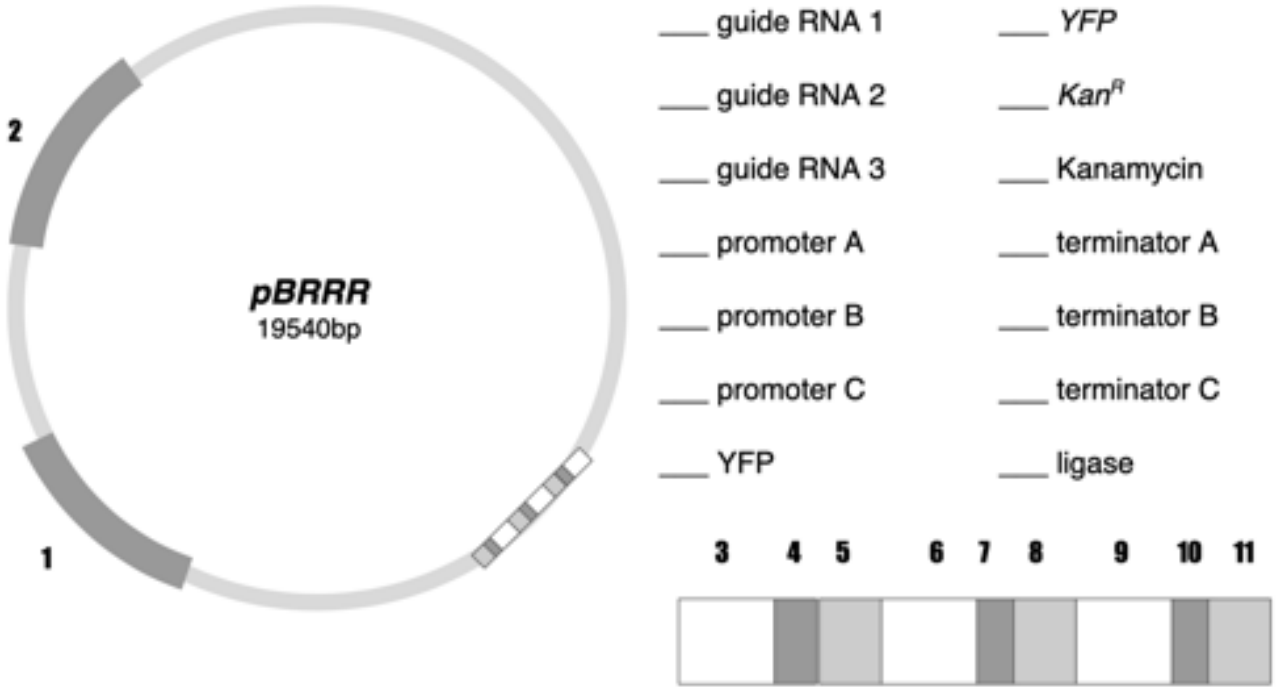
1. Explain why the following statements are wrong:

    a. (3 points) Transforming *A. thaliana* with *A. tumefaciens* in our project will result in plants with tumors.

Outcomes: 7
Rubric:

Points	Answer
3	The tumor genes were removed from the T-DNA Plasmid so no tumors will be made.
1	Answer only mentions no tumors without reason.

2. (15 points) You decide to jump into the biotech industry by starting a custom plasmid design service called Build-pBRRR. A client needs to engineer three guide RNAs to use CRISPR/Cas9 to knock out a gene and two of its homologs.
    a. Match the number on the blank next to the term (some will be left blank)



Outcomes: 9
Rubric: 1 point each blank

a. Which of the genes in that construct is similar to the selectable marker used in Experiment 3?

Outcomes: 8
Rubric:

Points	Answer
4	Exp. 3 used ampicillin selection and had AmpR on the plasmid, here kanamycin is used for selection so the marker is KanR.
2	Either antibiotic or marker mentioned, but not both
1	Answer KanR without explanation.

3. You successfully transformed *A. thaliana* with your plasmid construct. However, the seedlings were mixed up with some of the wild-type (Col-0) plants. The mutation you introduced is not noticeable until a later stage in the plant's life cycle.

a. (5 points) Suppose the transformation works and you have knocked out the target gene. The new mutant gene results in a truncated (much shorter) non-functional protein. Explain how this happens.

Outcomes: 7
Rubric:

Points	Answer
5	Guides direct CAS9 to cut the DNA. Host repair is error prone.
2	Guide directs CAS9 to cut the DNA. Repair not mentioned
1	Weak answer.

4. (10 points) The two “guide RNA” sequences from the cloned plasmids are compared to the two sequences that were designed.
    a. What program (bioinformatics tool) can be used for this comparison?

Outcomes: 7
Rubric:

Points	Answer
1	MUSCLE
0.5	BLAST
0	E-CRISP

b. Use the sequence comparisons in the following table to explain what could account for those results

Sequence	guide RNA 1	guide RNA 2
generic	NNNNNNNNNNNNNNNNNNN	NNNNNNNNNNNNNNNNNNN
hypothetical	AGTTTGCATGGTAACCGAA	CTTAAGCGTGAAAGAGCAA
Plasmid 1	AGTTTGCATGGTAACCGAA	CTTAAGCGTGAAAGAGCAA
Plasmid 2	AGTTTGCATGGTAACCGAA	CT-AAG-GTGAAAGAGCAA
Plasmid 3	ATTACTTACTCCAAAGAGG	ATAGATTGCATCTTCCAAA

i. Plasmid 1:

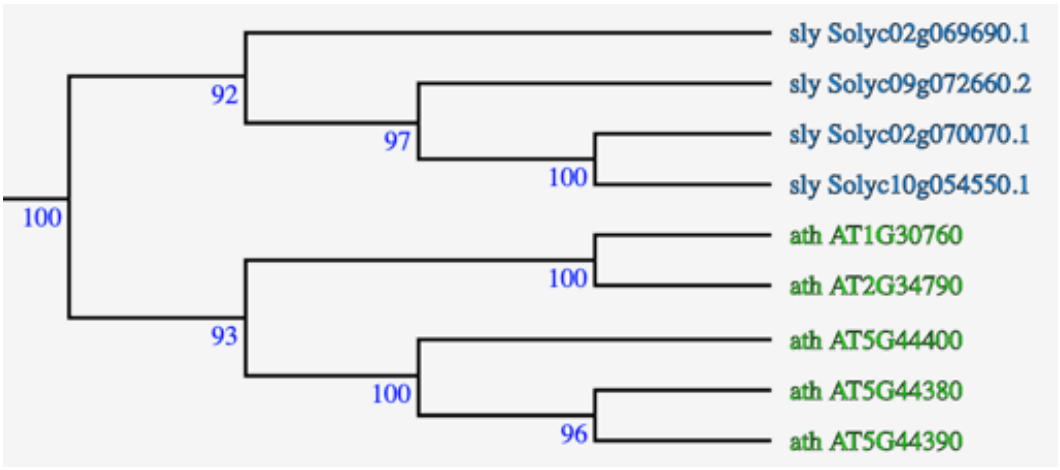
ii. Plasmid 2:

iii. Plasmid 3:

Outcomes: 11
Rubric:

Points	Answer
9	Plasmid 1: Correct guides, Plasmid 2: One guides correct, one with errors. Check chromatogram. Plasmid 3: no guides correct or inserted in plasmid.
5	Doesn’t mention differences or differences mentioned but no analysis
1	2 plasmids described incorrectly.

5. (8 points) The following questions relate to this phylogenetic tree:



a. What are homologs?

Outcomes: 4
Rubric:

Points	Answer
4	Genes similar in sequence and share a common ancestor
2	Genes similar in sequence, no evolutionary connection
1	Weak answer.

b. Which gene is most similar to **ath AT2G34790**? (Circle one from each pair)

ath AT1G30760 or **ath AT5G44400**

a**th AT5G44390** or **sly Solyc10g054550.1**

Outcomes: 4, 5
Rubric:

Points	Answer
2	Selects both correct homolog
1	Only one correct

c. Which paralog of the gene **AT1G30760** is most likely to code for a functionally redundant protein?

Outcomes: 4, 5
Rubric:

Points	Answer
2	Selects correct paralog and correct reason
1	Selects correct paralog only

Spring 2020 Questions

Instructions

This exam is to be completed by you alone. You are not to discuss or receive help of any kind form another person except a Biology 20 instructor or TA. Please type your name here to indicate that you did not receive, nor did you give help on this assignment.

Name:

The is an open note, open iLearn documents exam. You may not discuss it with another person (except Instructors and TAs). You may use the web.

All questions need to be answered in complete sentences and all calculations much be shown. Please use BLACK font color for answering. All questions must be typed. Equations must be typed, using the Word Equation Editor. Drawings must be done electronically (e.g., Google Draw) and inserted as an image. The answer length maximum is provided for each question. No hand written answers or drawings will be accepted.

Proper formatting of gene and protein names is required.

1. (12 points) A gene from rice (*Oryza sativa*) will be assigned for this exam. In addition to NCBI and others used in class you will need to use these as well:

Oryzabase

Rice Genome Annotation project (RGAP)

Your gene is: LOC_Os03g11600

a. The gene name and chromosomal location. (Oryzabase)

Outcomes: 1

Rubric:

Points	Answer varies based on assigned gene
4	Answer is correct and in complete sentences
2	Answer is correct but not complete sentences
1	Answer is partially correct.

b. Number of exons and introns. (RGAP)

Outcomes: 1

Rubric:

Points	Answer varies based on assigned gene
4	Answer is correct and in complete sentences
2	Answer is correct but not complete sentences
1	Answer is partially correct.

c. Describe the expression pattern. This will require some research as there is no convienent Atlas.

Outcomes: 1

Rubric:

Points	Answer varies based on assigned gene
4	Answer fully describes expression and in complete sentences
2	Answer is partially describes or not complete sentences
1	Answer is partially correct.

2. (15 points)Design PCR primers to detect the full gene and the mRNA. Draw an image (do not copy and paste the Primer Blast output) showing the location of the primers and the expected PCR product size.

Outcomes: 2

Rubric:

i. Drawing: 5 points: Drawing is well done, 3 points: Needs work, 1point: hand drawn.
ii. PCR Primers: 5 points: Correctly shown as arrow and bp, 3: Arrow or bps missing, 1: Shown but not correct
iii. PCR Product size: 5: points: shown, correct and includes primers, 3: primers not included in PCR product size, 1: line shown but size missing

3. (12 points) Generate a phylogenetic tree using *Orzya sativa, Zea mays*, and *Glycine max*. Modify the number of species selected if the number of genes are greater than 6.

A. Circle paralogs in BLUE for the three species. If there are no paralogs state why you concluded that.
B. Circle the orthologs to the rice gene in RED.

Outcomes: 4, 5, 6

Rubric:

Tree: 4 points: Correct tree, all species included as necessary, 2 points: Species missing or not enough homologs, 1 point: tree not well done.

Paralogs: 4: points: three paralogs correctly identified or explanation for none is correct. 2: Some paralogs shown, 1: Paralogs indicated but not fully correct

Orthologs: 4: points: three orthologs correctly identified or explanation for none is correct. 2: Some orthologs shown, 1: orthologs indicated but not fully correct

5. (14 points) Design guide RNAs that will remove the whole gene (but make sure your primers will still bind). Draw (not copy and paste) an image with guide locations and show the size of the deletion. Justify why you chose those two guides.

Outcomes: 9, 10
Rubric:

Points	Answer varies based on assigned gene
14	gRNAs shown with location, target whole gene, PCR primers shown, gRNAs are inside the PCR priming sites, Deletion shown and size is correct. Diagram is drawn correctly
7	gRNAs are wrong: outside PCR primers and/or do not target whole gene
1	Much is missing, weak attempt

6. (6 points) Design a PCR experiment that would show that the CRISPR guide RNAs were effective in making the deletion. Include unaltered genomic and mRNA in the experiment as well.

Experimental design
Outcomes: 2, 11
Rubric:

Points	
6	Prediction that the genomic DNA size will be smaller in the deletion plant, gDNA and mRNA from wild type and mutant, mRNA copied to cDNA by RT, PCR with primers, run gel
3	Prediction is missing or RT is missing
1	Description of setting up the reactions is given

7. (15 points) Using a computer drawing program or the Word drawing tools, draw the results. Include a ladder lane and band sizes.

Outcomes: 11
Rubric:

Points	
15	Gel image is completely correct: band sizes correct, band shown for gDNA for WT and mutant and cDNA for WT (maybe mutant), ladder given with band sizes
7	Much is missing or image is not neat.
1	Only minimal gel image provided without predicted results.

8. (8 points) Based on the available information about the gene predict the phenotype if the CRISPR experiment generated the deletion mutant. (Hint: Each gene has an image on Oryzabase.)

Outcomes: 8
Rubric:

Points	
8	Wild type phenotype is described, mutant phenotype is predicted based on gene function
3	Only mutant phenotype is given
1	Only WT phenotype is given.

Exam 2 Fall 2020

Instructions

Answer all the questions. You may use only websites with links in the exam. You may not use Google or your notes during the exam.

1. (20 Points) Use Arabidopsis.org and gene AT1G07200

a. (5 points) Using E-CRISP design two guide RNAs to delete the gene.

b. (5 points) Look up the gene structure. Draw the structure here (no screen shots).

c. (5 points) On the drawing indicate the location of your guide RNAs.

d. (5 points) Draw the location of two PCR primers that you could use to verify the proper deletion was made.
Outcomes: 1, 2, 3, 9, 10, 11
Rubric:

Points	
15	Diagram is correct, primers and gRNAs clearly indicate ad labeled. E-crisp output is shown and discussed. Primers flank the gRNA positions
7	Diagram is correct and most information present. Primers do not flank gRNAs. Diagram is screen shot not a student drawing.
1	Diagram is missing or incomplete.

Appendix 30: Selection of target gene(s)

The goals of today are to 1) find the protein sequence of your target gene, 2) identify genes that are homologous to your target, and 3) learn more about the target and its homolog(s), if any.

Introduction

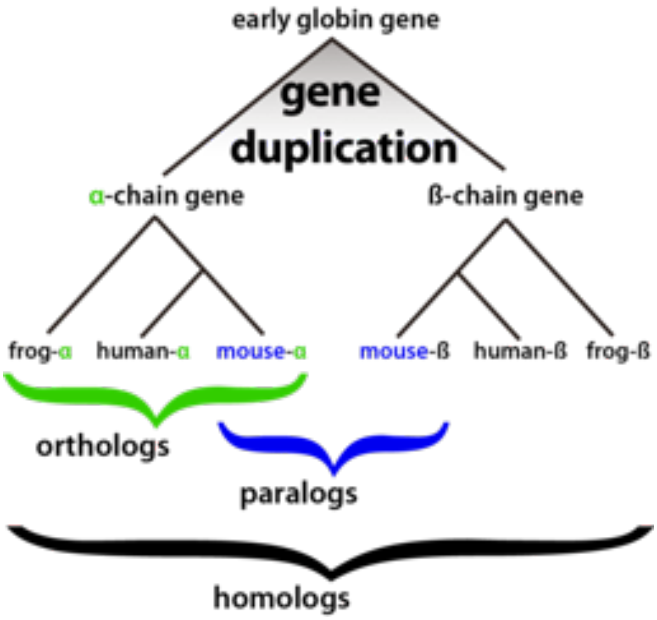
A limitation of genetic approaches is that sometimes loss of a gene does not have an obvious consequence (phenotype). This can be due to **functional redundancy**. In our radio example, imagine that there are two parts that do the same job, so breaking only one doesn't matter. Or, imagine that protein A_1 and protein A_2 are enzymes that can both catalyze the same reaction. Loss of the gene encoding A_1 may not produce a phenotype if A_2 has sufficient activity to keep the enzyme reaction products close to a normal level. Loss of the genes encoding both A_1 and A_2, however, may produce a phenotype.

Therefore, we may need to use CRISPR to knock out (mutate) both the target gene *and* any functionally redundant genes. So, how can we identify genes that may have functional redundancy?

One way to do this is to look for proteins that have similar sequences to the protein encoded by your candidate gene. These are known as **homologs** and can be readily found with BLAST. Generally, similarity in protein sequence implies at least some similarity in protein function. The stronger the sequence similarity, the more likely they have shared functions.

Another factor to consider is the evolutionary history of the target and its homologous genes. The more recently two genes have diverged, the more likely it is that they have similar functions, as there has not been sufficient time for them to acquire mutations that alter their functions. It is common for genes to undergo duplication. Sometimes these duplication events are localized (small portions of a chromosome), and sometimes they are due to whole genome duplications that have occurred in different plant lineages. The pair of genes created from a gene duplication event are called **paralogs**. If these do not acquire inactivating mutations and become pseudogenes (a common fate when there is minimal selection to keep two copies of the same thing active), the paralogs may eventually evolve a subset of functions of the original gene or entirely new functions. For example, two paralogs of a gene that is expressed in shoots and roots may evolve into one that is expressed in shoots only and the other that is expressed in roots only.

Orthologs are another form of homologous genes. These are genes that are separated by speciation rather than duplication. That is, equivalent genes may be found in different species that share a common ancestor because the ancestor carried this gene. Orthologs may often retain similar functions in different species. We will try to identify relatively recent paralogs of the target gene in *Arabidopsis thaliana* by defining an orthologous group found in monocots and dicots.



Finally, when considering potential sources of functional redundancy we may wish to examine the expression patterns of the homologous genes. If two homologous genes have altogether different expression patterns, one may not be able to substitute for the loss of the other and so they may not have redundancy. Also, if you are looking for homologous genes that redundantly control the growth of seedlings with your target gene of interest, it is reasonable to expect that a relevant gene(s) will also be expressed in seedlings (this does not exclude expression in other tissues, too). The trick here is that we don't know where to set the threshold for a level of expression that is developmentally significant. Some genes are expressed and functional at very low levels. Also, loss of a gene may trigger increased expression of a homolog to compensate.

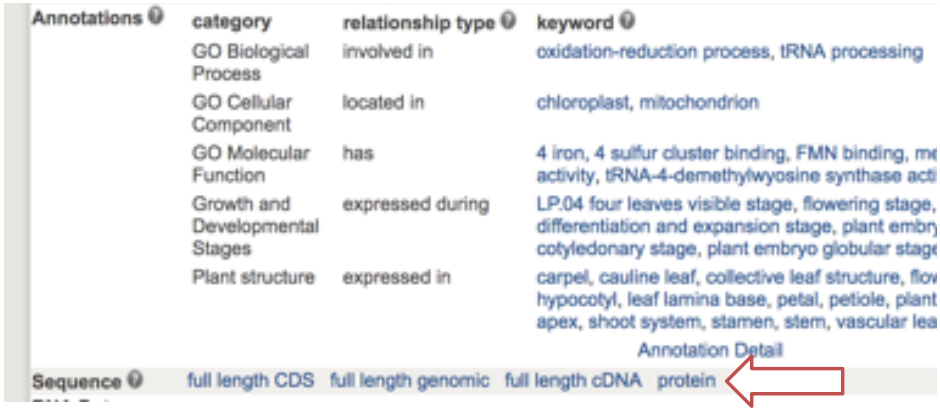
Identify homologs of your candidate target gene

TAIR (The Arabidopsis Information Resource) is a database for *Arabidopsis thaliana* that maintains molecular and genetic data. There are several tools we can use on this website, but the main resource we will use within TAIR is the BLAST function. Also, you will get information about your assigned gene by typing in the gene identifier in the search bar.

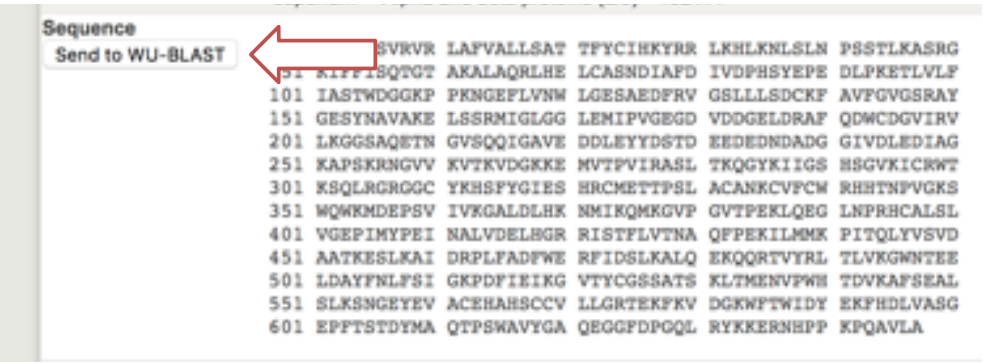
a. Go to the TAIR website- https://www.arabidopsis.org/index.jsp
   i. Type in the gene identifier in the search bar. Example: AT2G18790

   Arabidopsis gene identifier nomenclature is species (***A****rabidopsis* ***t****haliana*), chromosome (**2**), **G** (for gene), **gene number** on chromosome (typically in increments of 10 going from top to bottom of chromosome, as they anticipated adding in some genes that they didn't predict in the original annotations). Decimal places (AT2G18790**.1** and AT2G18790**.2**) indicate different models of the same gene (e.g. differently spliced forms, which would produce slightly different proteins).
   ii. You may see more than one search result appear, so double check that you are moving forward with the correct locus.

   iii. You can read some of the information about your gene, like the gene description, gene model, and some of the biological processes it may be involved in. The annotations are often very speculative and derived from homology to other genes or expression pattern rather than experimental evidence. Below the Annotations section, under **Sequence**, select "**protein**"

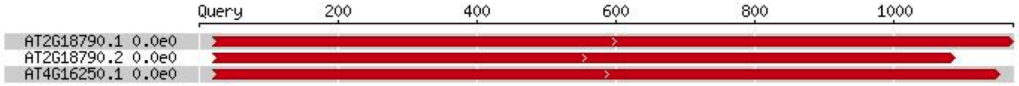


iv. Towards the bottom of the page, you will see your protein's amino acid sequence. Click on “Send to WU-BLAST.” You will probably want to save the protein sequence in another file.



v. At the bottom of the following page, hit “Run BLAST” to see your search results. Make sure that you are using the **BLASTP** program and a protein sequences dataset, e.g. "**Araport11 protein sequences (protein)**". This may take a few minutes. If the search leads you to a broken link, paste your protein sequence in to https://www.arabidopsis.org/Blast/ Note that using BLAST in the Safari browser may not work very well. If the TAIR BLASTs do not work, see the Phytozome addendum.

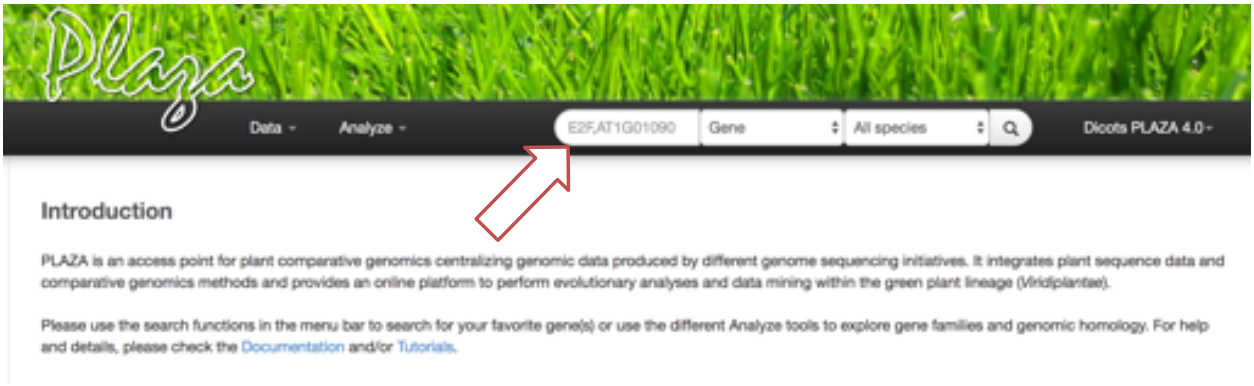
vi. The results will show you the most closely related proteins to your protein of interest. You may see multiple variants of a protein, which are indicated by a “.1” or “.2”, etc., at the end of a gene identifier (example: AT2G18790.1 vs AT2G18790.2).



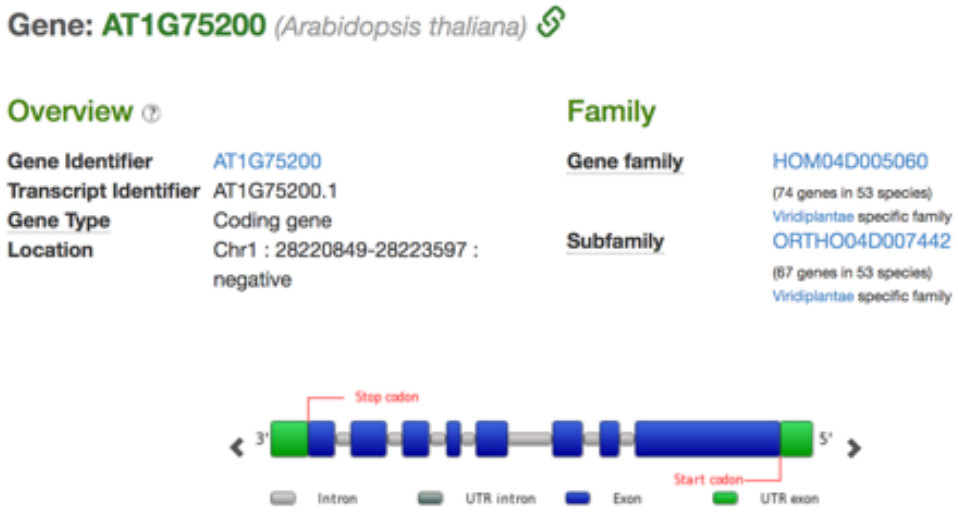
You will need to make a list of the closest homologs, but you only need to refer to the core gene identifier and not all of the variants for a gene (e.g. AT2G18790 only is fine). Make sure to note the E-values for the homologs, which indicates how likely it is for the homologous sequence to be similar to your query sequence by chance (the lower this value, the better). There will be a substantial rise in the E-value with homologs that have weak similarity. Examine the alignments to get a sense of how well the sequence is conserved and whether it is over the full-length of the query sequence or just a small portion. Does your target protein have any likely homologs or is it relatively unique?

Examine the evolutionary history of your candidate target gene and its homologs

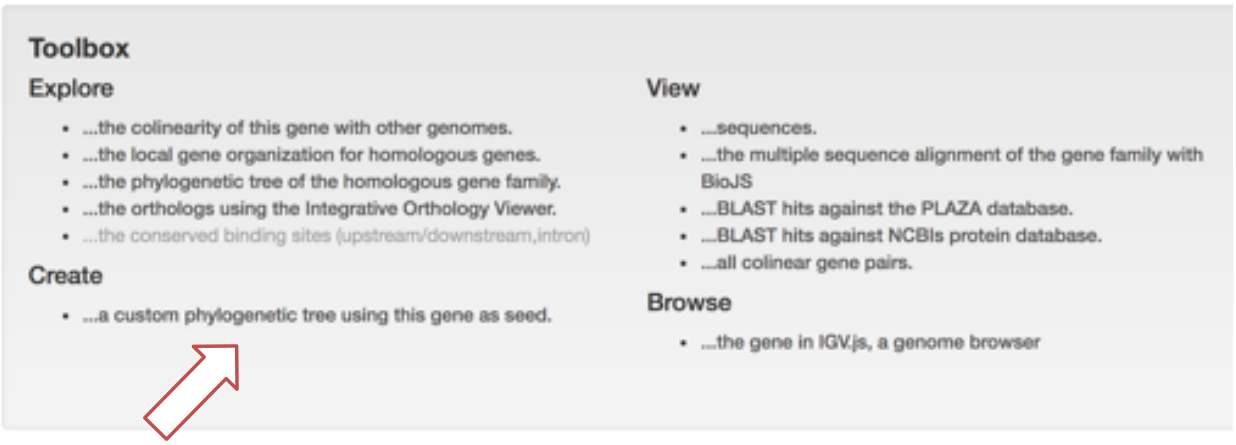
1. Go to PLAZA https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_dicots/
At the top of the page enter your gene identifier.



2. Next you'll find some basic information about your gene. How many exons does it have? Note the orientation of the gene (left-to-right or right-to-left) on the chromosome (but always 5'>3'!!).



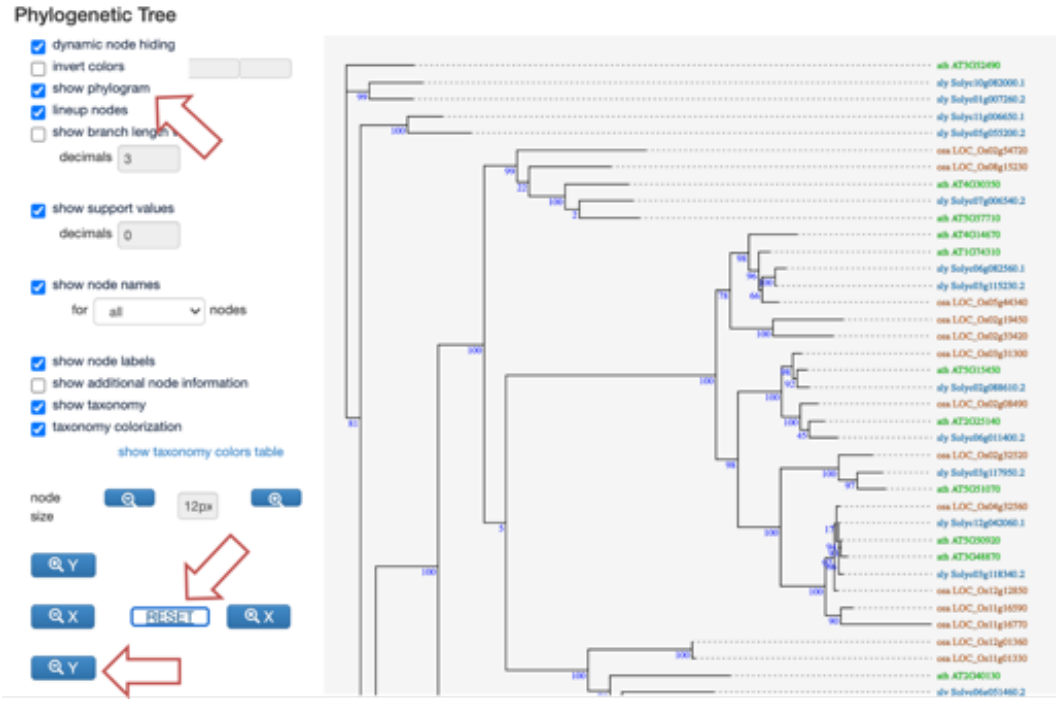
3. Under Toolbox section, select **Create** a custom phylogenetic tree.



4. On the next page, choose **Species Selection**

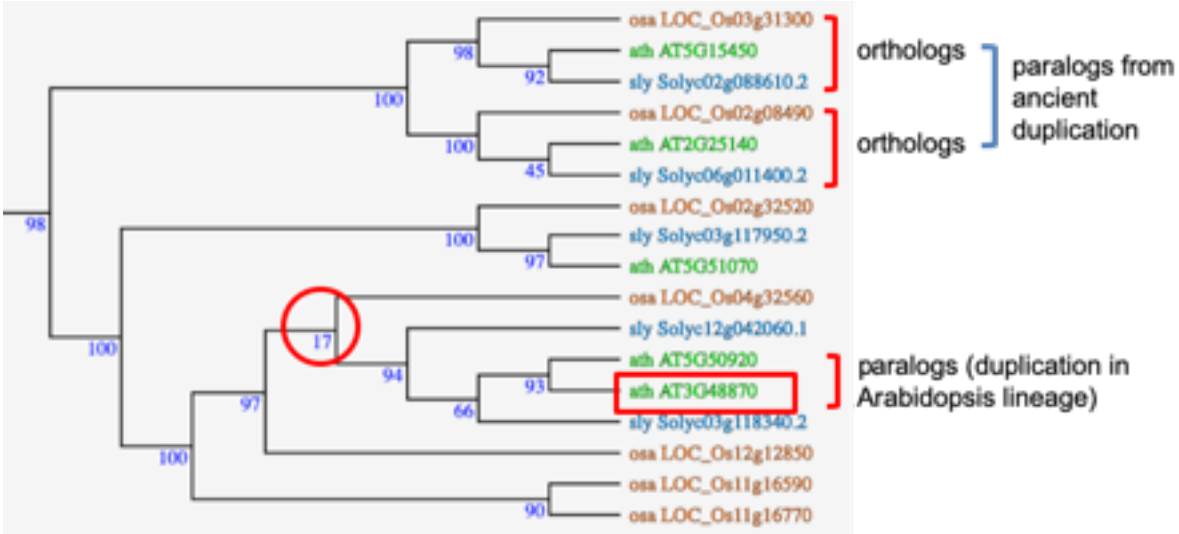
5. Select the boxes for *Arabidopsis thaliana, Solanum lycopersicum* (tomato)*, Oryza sativa* (rice), and *Amborella trichopoda*. The first three species have highly refined diploid genome sequences and are representatives of dicots (*Arabidopsis* is a rosid genus, *Solanum* is an asterid genus) and monocots. *Amborella* is an early-diverging angiosperm. Then, click the blue button at the bottom, **Create MSA/Tree**. Behind the scenes, this takes the homologous sequences from those species and creates a multiple sequence alignment to match up stretches of identical or highly similar amino acids. Then it generates a simple phylogenetic tree based on those alignments. This quick phylogeny is fine for our purposes but would need to be more rigorously done for actual publications. If your phylogeny fails to build, try removing some species. Get rid of *Amborella* first, and if it still doesn't work remove *Solanum*.

Once the tree has been produced you will see a window like this. It may help to unclick "show phylogram". This present the tree with the same structure, but basically ignores the branch lengths (which are indicators of sequence divergence). It may be useful to click "-Y" to make it zoom out a bit, but don't overdo it or branches will start to disappear from view. Also try zooming in to make sure you can whether genes are already collapsed in the original view of the phylogeny (your target may be among them). If things go haywire, just click RESET to get back to the original view.



On your phylogeny, you will need to

- Locate your candidate gene
- Trace back up the branch leading to your target gene until you find a node from which branches that lead to homologous rice genes also connect. This defines a clade of orthologs for monocots and dicots. Your clade might also connect to an *Amborella* sequence, but not necessarily.
- Look within this clade for any other Arabidopsis genes. These are the most closely related genes to your candidate and may need to be knocked out as well.



In the example above, At3g48870 is the target gene of interest. At5g50920 is a close paralog that likely emerged in the lineage leading to Arabidopsis after the divergence of monocots and dicots. It is the strongest candidate for functional redundancy and should be targeted for knockout. The genes that are the next most similar to the target and potential contributors to functional redundancy are At5g51070, followed by At5g15450 and At2g25140. All of these Arabidopsis genes may be considered paralogs, although most arose from duplication events that preceded the separation of dicot and monocot lineages.

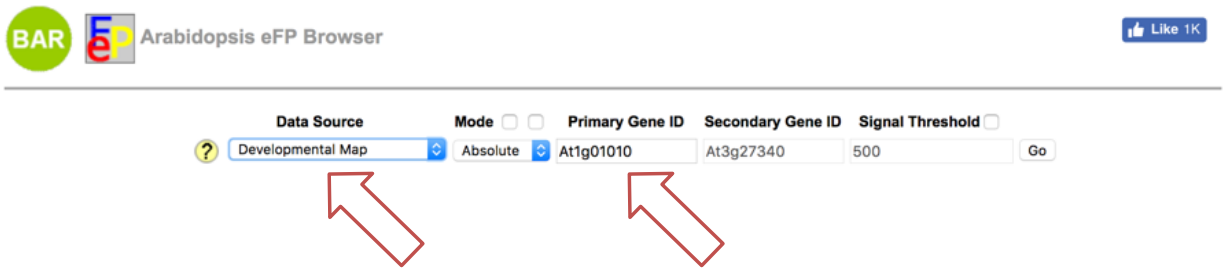
Consider the list of BLAST matches to your target protein that you identified. Where do they appear in the phylogeny? How does the phylogeny add to your understanding of the target gene's evolution compared to a simple similarity comparison?

Examine the expression patterns of your target gene and its homologs

The abundance of tens of thousands of transcripts can be determined in a single experiment through microarrays or next-generation sequencing of DNA derived from RNAs (RNA-seq). By collecting tissue samples from different parts of the plant, or plants that have been treated with different stimuli, you can extract RNA and examine the **transcriptome** to see what genes have higher or lower expression.

An easy viewer of expression data is found here: http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi

At the top of the page, under Primary Gene ID, enter the gene identifier you are interested in. At the left, select as your Data Source either Developmental Map and/or Klepikova Atlas. Press Go and you will see a color-coded representation of the expression levels of your gene.



Note that some genes are not present in the Developmental Map dataset, which is based on Affymetrix microarray data. That is because the microarrays only assay ~22,000 of the genes found in Arabidopsis. If this is the case, the Klepikova Atlas, which was done with RNA-seq, should have data.

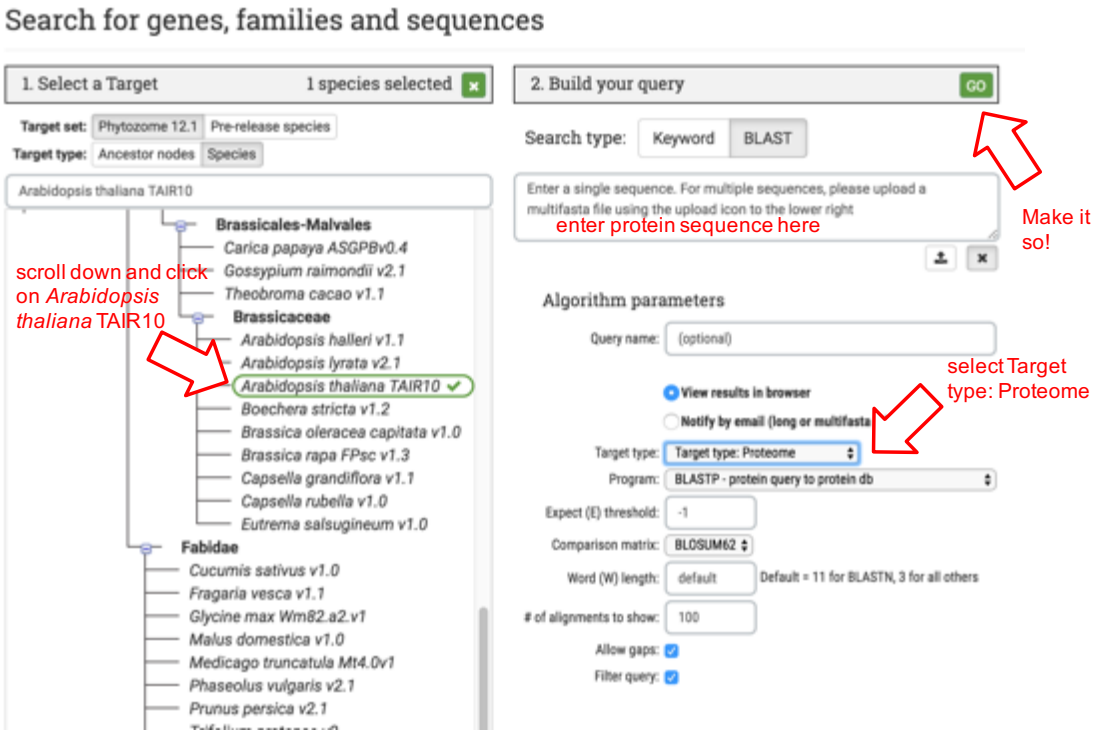
Pay special attention to the seedling data. Is your gene or its homologs expressed in the hypocotyl, roots, or cotyledons of seedlings? What about seeds? How high is the expression? (you can hover over to see numbers) The scale bar is set to the bottom left. Darker red indicates higher expression. Note that the scale bar is relative, and red for one gene does not mean the same level of expression as for another gene.

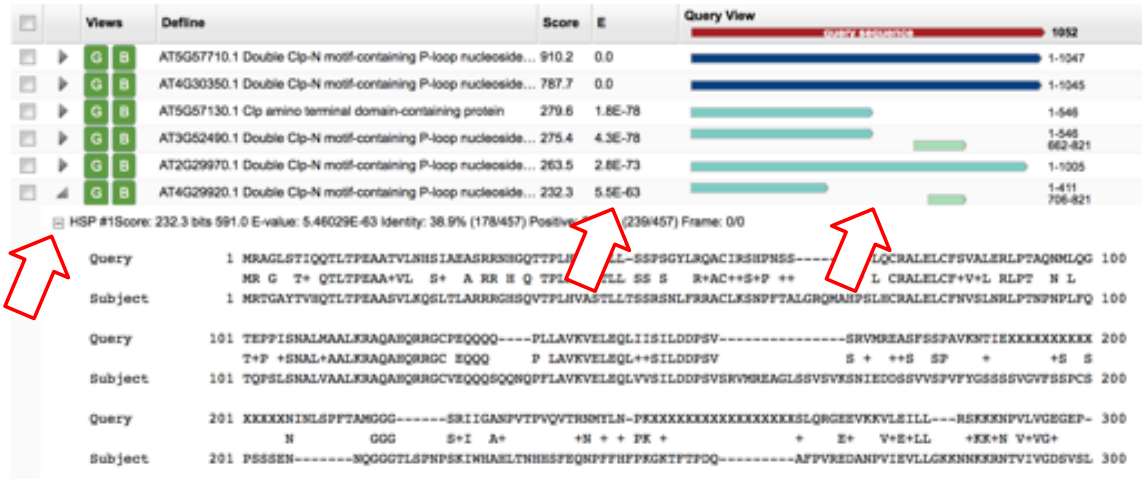
This is another interesting site to explore your genes in different ways
http://bar.utoronto.ca/eplant/

Phytozome BLAST alternative to TAIR's BLAST

Go to phytozome.org (this will direct you to https://phytozome.jgi.doe.gov/pz/portal.html)

At the top left under the Tools dropdown menu, select BLAST.





You can visualize the alignments by clicking the triangle symbol, and then the + below. The top sequence is your query, the bottom sequence is that of the matching protein, and in the middle are the amino acids that are identical or have similar chemical properties (+) between the two sequences.

Note that 2.3E-38 means 2.3 x 10^-38 (this is scientific notation). The digits after the E matter much, much more than those before it.

The results are ranked by their similarity to the query sequence, with the strongest matches on top. Note that extent of homology across the length of the query sequence. Does it cover the entire or most of the query length? Or does the bar only match a fragment of the query?

Appendix 31: Guide RNA design

Selection of guide RNA sequences

In the previous lab you were given a target gene and looked for genes that may have similar function to the target gene. Today's goals are to design two synthetic guide RNAs that will enable Cas9 to target your gene and potentially its close homologs at the same time.

The CRISPR-Cas9 system

Cas9 is a double-stranded nuclease that cleaves DNA at specific locations based upon the guide RNA it is carrying. The guide RNA can be 17-20 nt long and can be any sequence as long as it is followed by the correct protospacer adjacent motif (PAM) in the genome being targeted. The PAM for Cas9 is NGG, where N is any of the four nucleotides. (See figure below for an example of how Cas9 binds to and cleaves its target.) Other enzymes with similar activity to Cas9 have been found that have different PAM requirements.

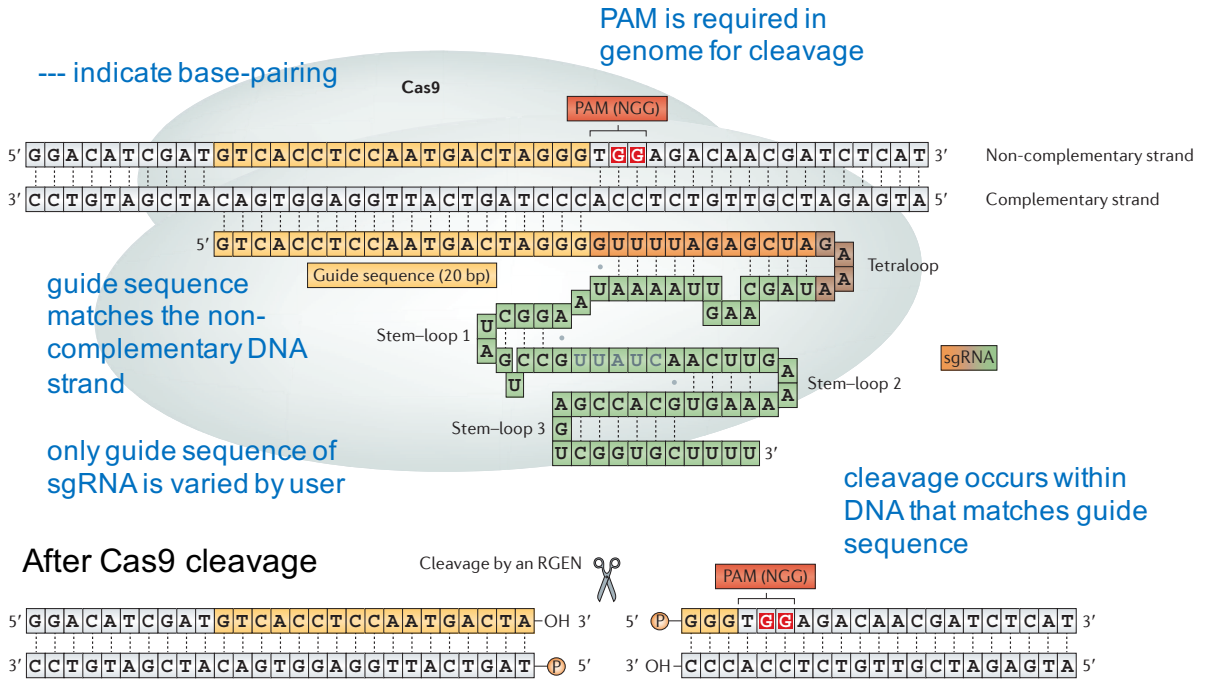


Figure adapted from Kim & Kim, 2014, *Nat Rev Gen*

After Cas9 induces a double-stranded DNA break, the cell will attempt to repair it. Two methods are used. Nonhomologous end-joining is relatively straightforward but prone to errors. In it, the ends of two double-stranded DNA molecules are simply ligated together, but potentially with the insertion or loss of one to hundreds of base pairs at the ends as the broken ends are cleaned up. If the repair perfectly restores the guide sequence, Cas9 can cleave again. Homology-directed repair uses another copy of the broken DNA as a template for repair. This is a high-fidelity DNA repair mechanism.

sgRNA design considerations
If G makes up 25% of the bases in a genome and is equally distributed, on average a guide RNA for Cas9 could be designed every 16 bp. However, this is typically not the case, and further considerations must be made when selecting a gRNA.

First, if the guide sequence (+PAM) of an sgRNA is present at more than one location in the genome, Cas9 will cleave at more than one place. It has no way of distinguishing which location the researcher is interested in. Related to this, a primary concern with CRISPR-Cas9 is **off-targeting**. This is when Cas9 cleaves DNA that is similar to the guide sequence, rather than only cleaving at the desired location. Researchers typically go to great lengths to ensure that their sgRNA is unlikely to have off-targeting effects, or use Cas9 variants that have higher-fidelity. Note that even with some off-targeting, genome editing with Cas9 induces far, far fewer mutations than traditional methods of mutagenesis, which makes it ideal for developing genetically engineered organisms.

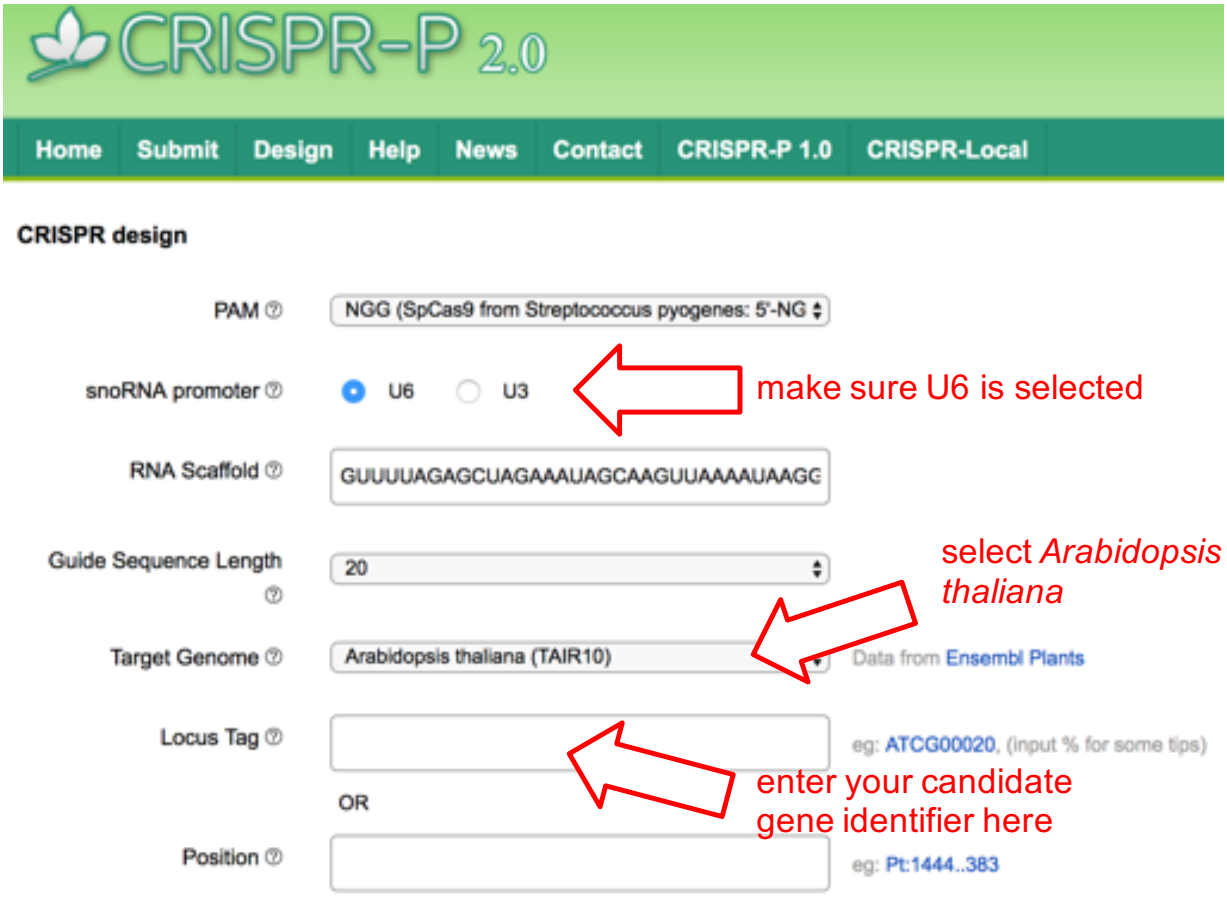
Second, some guide sequences enable more effective target cleavage than others. Accurately predicting which sgRNAs will be most effective is an ongoing challenge that many labs are attempting to decipher. Some potential factors are GC-content, which could influence the strength of nucleic acid binding, and the accessibility of DNA. DNA that is tightly packed into chromatin may be more resistant to cleavage simply because Cas9 cannot scan it easily. In the end, we simply have to try different guide sequences and be a little lucky.

Third, the purposes of the experiment must be taken into consideration. In many cases, the goal is simply to introduce a loss-of-function mutation into a gene. This requires targeting an exon, as the typically small insertion/deletion (**indel**) mutations caused by Cas9 are unlikely to have an effect if they occur in introns. Mutations in non-coding regulatory regions adjacent to the gene may have an effect on gene expression but this is usually unpredictable. In contrast, an indel in an exon will cause a **frameshift** mutation approximately 2/3 of the time. Frameshift mutations cause the ribosome to incorporate the wrong amino acids as it translates an mRNA, effectively stopping translation of the correct protein sequence at the point of the mutation. Loss-of-function mutations will generally work well if placed closer to the beginning of the gene than the end, but caution must be exercised with mutations very near the beginning of the gene as translation could potentially proceed from an alternative downstream start codon. Sometimes it is useful to use two sgRNAs to target a gene. This has the potential to create large deletions between the two target sites, which are effective at knocking out a gene's function and are also very easy to detect. It also improves the likelihood of successfully inducing mutations, as one guide RNA may work poorly.

The sgRNA strategy you use below will be based on the number of homologs you identified that may have functional redundancy with your primary candidate gene.

CRISPR-P 2.0 is a tool that we can use to easily design sgRNAs in plants. The CRISPR-Cas9 vector we will introduce into plants can express two gRNAs. CRISPR-P has a nice feature that lets you see if your sgRNA will potentially target sequences other than that of your gene of interest. In contrast to the typical practice of avoiding off-targeting, we will take advantage of off-targeting to try to simultaneously knock out genes that are homologous to your candidate gene.

1. Go to the CRISPR-P website: http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR



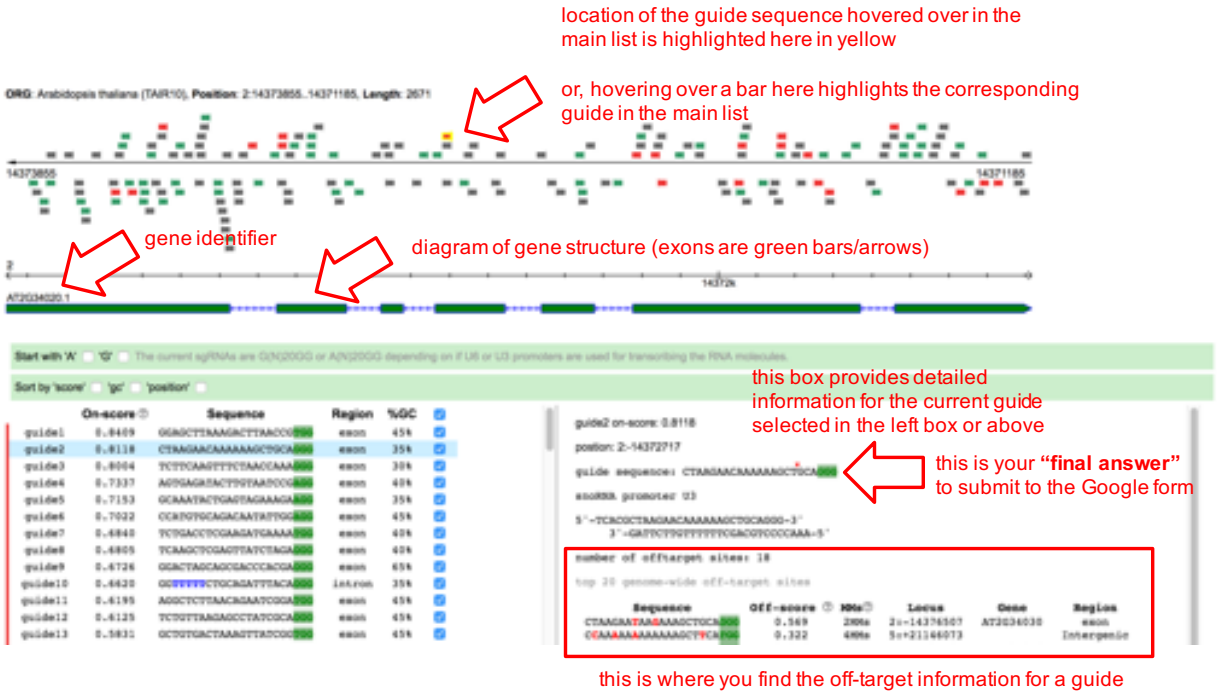
Keep the default settings of

PAM: NGG (SpCas9 from Streptococcus pyogenes..)
RNA Scaffold
Guide Sequence Length: 20

2. **Change** the following settings

snoRNA promoter: U6
Target Genome: Arabidopsis thaliana (TAIR10)
Locus Tag: your candidate gene identifier (given to you at the beginning of Day 1)

3. Click Submit.



General rules and tips you need to follow - READ CAREFULLY

- It is easy to accidentally hover over the wrong guide when moving your mouse to the detail box, which changes everything - beware!
- The check boxes can be used to mark a set of guides you find potentially interesting
- Make sure that the Region targeted is an exon
- Check the orientation of the gene - if the arrow points to the right, the beginning of the gene is on the left side of the screen; if the arrow points to the left, the beginning of the gene is on the right side of the screen. This is critical for identifying the first exons.
- Preferably place a guide sequence toward the middle or first part of the gene rather than close to the end
- **Do not use guides with five or more Ts in a row (these are highlighted in blue), as these will trigger premature termination of sgRNA transcription**
- **Do not use guides that contain the sequence "GGTCTC". This will disrupt the cloning process later.**
- Given a good option, choose a guide sequence that starts with G (for an initial pass, you can select Start with a 'G' in the light green toolbar to see these guides only). This is not necessary but is potentially helpful for increasing guide binding to its target (20 nt will match the genome instead of 19).
- Guides with high On-score are predicted to work well, but this is only a prediction
- Guides with high Off-score are predicted to target the listed gene more efficiently. Higher numbers of mismatches (MMs) should *reduce* the likelihood of cleavage
- Despite what the Off-score may say, mismatches are supposedly not tolerated in the 12 nucleotide "seed" sequence directly before the PAM (which is highlighted in green)
- Most genes have been assigned to two people in each section. Coordinate with the other person to make sure you don't select the same two guides

- **Copy and paste everything in the detail boxes for your chosen guides to your lab notebook.** Copy and paste your two chosen 23 nt guide sequences into the provided Google form

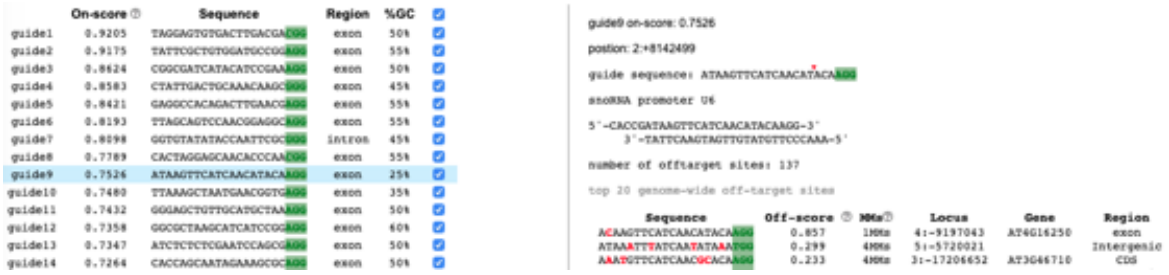
If your candidate gene has no homologs
> Well, aren't you lucky? Design both guide RNAs to target the same gene. Avoid off-targets by choosing guides that have high numbers (3 or more) of mismatches (MMs) to the coding sequences of other loci. Off-target loci that are not in exons (e.g. intergenic or intron) are probably ok, particularly with 3 or more MMs.

If your candidate gene has one close homolog
> Choose one guide that has a high on-score value for the candidate gene, and one guide that has a high on-score value for the homolog. If possible, choose guides that may off-target the other gene in the pair (look for 0 or 1 MMs, avoid MMs in the 12 nt adjacent to the PAM, and look for high Off-score for the other Gene). It is worth sacrificing some on-score value (remember, it's just a prediction) for the sake of beneficial off-targeting. You may want to perform a new CRISPR-P search with the homolog gene identifier to find a good second guide.

If your candidate gene has two or three close homologs
> Prioritize one guide that will work well for your candidate gene. If possible, choose a guide that may off-target another one of the homologs (look for 0 or 1 MMs, avoid MMs in the 12 nt adjacent to the PAM, and look for high Off-score for the other Gene). It is worth sacrificing some on-score value (remember, it's just a prediction) for the sake of beneficial off-targeting. For the second guide, try to prioritize knocking out the gene most likely to have functional redundancy. You'll have to make this decision based on expression (is one much more highly expressed in seedlings compared to the other homologs?) and similarity to the candidate (is one much more similar than the other homologs?). If possible, try to find a guide that may target this gene and one or more of the others. Sub-optimal off-targeting will be acceptable. It may be easiest to start a new CRISPR-P search for this homolog so you can find something that matches it perfectly and potentially off-targets the others.

If your candidate gene has many close homologs
> Well, aren't you unlucky? You must do your best! Try to find guides that will target as many of the homologs as you can, making sure that one of them will target the candidate gene exactly. As in the previous description, you will want to take expression patterns and sequence similarity into account to prioritize your desired off-targets. Also as above, it may help to run additional CRISPR-P searches to find a guide with an exact match to a homolog that may also may target some of the other genes.

Here is an example of a sgRNA (guide9, left box) that targets the second exon of a gene of interest and will also likely target an exon in AT4G16250, which is the gene that is most homologous to the gene of interest (right box). Note the off-score value for this guide to target AT4G16250 is close to 1 (= more likely), and there is only a single mismatch that is outside the seed region of the guide. This guide sequence looks promising to target both genes at once.



	On-score	Sequence	Region	%GC	
guide1	0.9205	TAGGAGTGTGACTTGACGACGG	exon	50%	☑
guide2	0.9175	TATTCGCTGTGGATGCCGCAGG	exon	55%	☑
guide3	0.8624	CGGCGATCATACATCCGAAAGG	exon	50%	☑
guide4	0.8583	CTATTGACTGCAAACAAGCGGG	exon	45%	☑
guide5	0.8421	GAGGCCACAGACTTGAACGAGG	exon	55%	☑
guide6	0.8193	TTAGCAGTCCAACGGAGGCAGG	exon	55%	☑
guide7	0.8098	GGTGTATATACCAATTCGCGGG	intron	45%	☑
guide8	0.7789	CACTAGGAGCAACACCCAACGG	exon	55%	☑
guide9	0.7526	ATAAGTTCATCAACATACAAGG	exon	25%	☑
guide10	0.7480	TTAAAGCTAATGAACGGTGAGG	exon	35%	☑
guide11	0.7432	GGGAGCTGTTGCATGCTAAAGG	exon	50%	☑
guide12	0.7358	GGCGCTAAGCATCATCCGGAGG	exon	60%	☑
guide13	0.7347	ATCTCTCTCGAATCCAGCGAGG	exon	50%	☑
guide14	0.7264	CACCAGCAATAGAAAGCGCAGG	exon	50%	☑

guide9 on-score: 0.7526

postion: 2:+8142499

guide sequence: ATAAGTTCATCAACATACAAGG

snoRNA promoter U6

5'-CACCGATAAGTTCATCAACATACAAGG-3'
3'-TATTCAAGTAGTTGTATGTTCCCAAA-5'

number of offtarget sites: 137

top 20 genome-wide off-target sites

Sequence	Off-score	MMs	Locus	Gene	Region
ACAAGTTCATCAACATACAAGG	0.857	1MMs	4:-9197043	AT4G16250	exon
ATAAATTTATCAATATAAATGG	0.299	4MMs	5:-5720021		Intergenic
AAATGTTCATCAACGCACAAGG	0.233	4MMs	3:-17206652	AT3G46710	CDS

Whatever you do, explain the reasoning for your guide choices in your lab notebook.

Appendix 32: Designing and ordering primers

How to Collect Sequences From Students

Use a Google Form to collect student guide sequence submissions. We collect the following information from each student

- Student name
- Student email
- Target gene identifier (using At_g____ format)
- Homologous genes being targeted, if any (separate Arabidopsis gene identifiers by commas)
- Guide sequence 1 (copy entire 23 nt sequence)
- Does it target an exon? Y/N/M
- Guide sequence 2 (copy entire 23 nt sequence)
- Does it target an exon? Y/N/M
- Comments (e.g. which genes do each of your gRNAs target and how well?)

How to Process and Order Primers

Primer design can be carried out by the students. But, each set of 4 primers will cost about $20-$30, depending on per-base pricing for 162 nt total. To minimize losses due to errors, we recommend at least double-checking their primer sequences with the following process.

Set up formulas (e.g. in Excel) to automatically process the 23 nt sequences into primer sequences to order. For the example formulas below (**bold**), assume A1 = Arabidopsis gene identifier, A2 = 23 nt guide 1, A3 = 23 nt guide 2.

1. Check that the sequence is exactly 23 nt long. **=IF(LEN(A2)=23,"OK","WRONG LENGTH")**
2. Check that the sequence ends in GG.
   =IF(MID(A2,22,2)="GG","OK","BAD PAM")
3. Trim off the first nucleotide (if using N_{19} guide length) and 3' PAM (NGG)
   =MID(A2,2,19)
4. For the second guide, the N_{19} sequence needs to be reverse-complemented. For batch reverse complement preparation it is easiest to create FASTA-formatted sequences of the second guide.
   =">"&A1 & CHAR(10) & MID(A3,2,19)
   Then use the tool at https://www.bioinformatics.org/sms2/rev_comp.html
5. Prepare primer sequences as follows *(*these are not Excel formulas; BsaI recognition sites in red)*
   BsF = "ATATATGGTCTCGATTG" + (guide 1 N_{19}) + "GTT"
   BsF0 = "TG" + (guide 1 N_{19}) + "GTTTTAGAGCTAGAAATAGC"
   BsR0 = "AAC" + (r.c. guide 2 N_{19}) + "CAATCTCTTAGTCGACTCTAC"
   BsR = "ATTATTGGTCTCGAAAC" + (r.c. guide 2 N_{19}) + "CAA"

Appendix 33: Complete materials and equipment list

Note that materials and equipment listed here pertain to the wet-lab protocols described below. Actual materials needed may vary slightly depending on the plasmids needed for the project and reagents/equipment available.

Reagents for making media and solutions:

- TAE – Used at 1X concentration but is useful to create and store as a 50X stock solution. Preparing 2L of 50x stock should be plenty for a CURE of 24 students, although the volume of TAE needed does vary depending on the apparatus.
    - Tris base (Fisher Scientific, catalog number: BP152-1)
    - EDTA (Fisher Scientific, catalog number: BP120500)
    - Acetic acid (glacial) (Fisher Scientific, catalog number: A38212)
    - KOH (for adjusting pH of solutions) (Fisher Scientific, catalog number: P250-1)
    - 50X TAE Recipe for 1L:
        - Combine 242g of Tris base, 57.1mL of acetic acid, 100mL of 0.5 M EDTA pH 8.0
        - Bring to volume with deionized water and store at room temperature.
    - 0.5M EDTA recipe for 500mL:
        - Add 93.05g of EDTA disodium salt (MW=372.24g/mol) to 400mL deionized water. Adjust the pH with solid KOH plates, until the EDTA dissolves at around 8.0 pH. Bring to volume with deionized water. Autoclave if autoclave is available. Store at room temperature.
- SOC and LB media – LB can also be purchased in a prepared powder form (such as Sigma-Aldrich L3022 LB Broth) that just needs water added. Media should be sterilized before storage or use. If an autoclave is unavailable, microwaving until the media boils can be done instead. Make sure so cool media to ~55°C before adding antibiotics. LB can be used instead of SOC.
    - Yeast extract (ACROS Organics, catalog number: AC451120010)
    - NaCl (Fisher Scientific, catalog number: BP358-1)
    - Tryptone (Fisher Scientific, catalog number: BP9726-500)
    - Agar (Teknova, catalog number: A7777)
    - Glucose (Fisher Scientific, catalog number: D16-1)
    - 0.22-µm filter (Genesee Scientific, catalog number: 25-244)
    - Antibiotics: rifampicin (Sigma-Aldrich, catalog number: R3501-1G), gentamicin (Fisher Scientific, catalog number: BP9181), kanamycin (Fisher Scientific, catalog number: BP906-5) (or appropriate antibiotic selection for vector used)
    - For 1L SOC, use 5g yeast extract, 0.5g NaCl, 20g Tryptone, bring volume to 980mL with DI water and autoclave. Allow it to cool to 60°C or less. Add 20 mL of a sterile 1 M solution of glucose.
        - To make 1M sterile glucose: dissolve 18 g of glucose in 90 mL of H_2O. After the sugar has dissolved, adjust the volume of the solution to 100 mL with H_2O and sterilize by passing it through a 0.22-µm filter (Millipore). Store the solution at 4 °C for short term storage.
    - For 1L LB media, dissolve 10g tryptone, 10g NaCl, and 5g yeast extract in 950mL of DI water. Adjust the pH to 7.2 with 1M KOH. Adjust the final volume to 1L with water. Autoclave or microwave to sterilize.

Materials:

- Autoclave
- pCBC-DT1T2 (Can be ordered from Arabidopsis Biological Resource Center (ABRC): https://abrc.osu.edu/stocks/number/CD3-1982)
- Deionized water
- Ice
- Sharpies
- Disposable gloves (such as nitrile disposable gloves) (Fisher Scientific, catalog number: (large) 19169075, (medium) 199169069, (small) 19169068)
- Safety glasses or goggles
- Lab coat
- Micropipettes and disposable tips (P1000, P200, P20, and P2 recommended. Students can share micropipettes if needed.)
    - Micropipettes available from Eppendorf catalog number: 2231300004 or 3123000900
    - Tips available from Genesee Scientific
        - Racked 10µl Catalog number: 23-121R
        - Racked 200µl Catalog number: 24-151R
        - Racked 1000µl Catalog number: 24-161R
        - Tip boxes can be recycled and refilled by ordering inserts to save money
- 1.5mL microcentrifuge tubes (Genesee Scientific, catalog number: 24-282)
- 0.2mL PCR tubes (Thermo Fisher, catalog number: AB0771)
- 6X loading dye (Thermo Fisher, catalog number: R0611)
    - Can also be made with the following recipe:
        - 30% glycerol (v/v) (Fisher Scientific, catalog number: BP2291)
        - 0.25% (w/v) bromophenol blue (Sigma-Aldrich, catalog number: B0126-25G)
        - 0.25% (w/v) xylene cyanol FF (Sigma-Aldrich, catalog number: X4126-10G)
- Agarose (Fisher Scientific, catalog number: BP160500)
- Heat resistant mittens or gloves – oven mitts or potholders work as well
- DNA ladder (100bp or 1kb ladder is recommended.) (New England Biolabs, catalog number: 43501-312)
- Competent *Agrobacterium tumefaciens* (electrocompetent bacteria is ideal because the transformation process is faster, but thermo competent bacteria could be used as well.)
    - Protocols for creating thermocompetent and electrocompetent Agrobacteria are readily available online.
    - See (Hellens, R., et al., 2000)
- PCR product purification kit (Qiagen, catalog number: 27106)
- High fidelity polymerase PCR reagents (Fisher Scientific, catalog number: 80511-386)
- Low fidelity polymerase for colony PCR (Bioline, catalog number: BIO-25044)
- Ethidium bromide (Fisher Scientific, catalog number: BP1302-10)
- 10X T4 DNA ligase buffer (New England Biolabs, catalog number: B0202S)
- NEB Golden Gate Assembly mix (recommended) (New England Biolabs, catalog number: E1602)
    - Also some protocols to create your own mix available online that require:
        - T4 DNA ligase (New England Biolabs, catalog number: M0202T)

        - Bsa-I HF (New England Biolabs, catalog number: R3535S)
- pHEE401E, pYUU, or similar vector
    - Both available from ABRC
        - PYUU: https://abrc.osu.edu/stocks/number/CD3-2854
        - pHEE401e: https://abrc.osu.edu/stocks/number/CD3-1980
- Chemically competent or electrocompetent *E. coli* (Thermo Fisher, catalog number: C404003)
    - Commercial cells are expensive. To conserve reagents the 50µl tubes can be split into 25µl or 16µl aliquots.
    - Protocols are available online to create competent *E. coli*.
- Electroporation cuvette (if using electrocompetent *E. coli*)
    - 0.1cm or 0.2cm cuvettes will work (Thermo Fisher, catalog number: P41050 and P45050)
- Parafilm or regular plastic wrap
- Biohazard bags (Fisher Scientific, catalog number: 01828ELC)
- Glycerol (Fisher Scientific, catalog number: BP2291)
- Sucrose (Fisher Scientific, catalog number: S5-3)
- 50mL tubes (Fisher Scientific, catalog number: 1495949A)
- Bleach
- Masking tape and sharpies (for labeling pots)
- disposable plastic transfer pipette or petri dishes or weigh boats
    - There are two methods to dip Arabidopsis flower buds:
        - Transfer pipettes to drip Agrobacteria onto the flower buds (Fisher Scientific, catalog number: 13-711-9AM)
        - Or placing the Agrobacteria in a shallow dish and dipping the flower buds into the solution. Petri dishes (Fisherbrand, catalog number FB0875711A), weigh boats (VWR, catalog number 10770-448), or any (preferably disposable) shallow container can be used.
- Flats for storing Arabidopsis (Greenhouse Megastore, catalog number: CN-FLHD)
- Plastic domes (that fit the Arabidopsis flats) or plastic wrap
- Silwet L-77 (Lehle Seeds, catalog number: VIS-30)
- Arabidopsis plants of the appropriate mutant background (Many mutant lines are available through ABRC https://abrc.osu.edu)
- Soil for growing Arabidopsis (We recommend SunGro Sunshine Mix #1)
- Primers:
    - U6-26p-F: 5'-TGTCCCAGGATTAGAATGATTAGGC
    - U6-29p-R: 5'-AGCCCTCTTCTTTCGATCCATCAAC
    - We recommend ordering primers from Integrated DNA Technologies (https://www.idtdna.com/)

Equipment:

- Thermocycler
- Microwave
- pH meter – for adjusting pH of media
- Gel electrophoresis system (including combs, gel tray, power supply and leads, and gel box)
- UV transilluminator
- Camera (To take pictures of gel. Cell phone cameras are acceptable.)

- Tabletop centrifuge or minicentrifuge (To spin down PCR tubes.)
- Nanodrop
- -20 °C freezer
- Vortex mixer – vigorous shaking by hand can be used instead
- Incubator with shaking that can be set to 30-37°C
- Heat block or bead bath for 60 °C incubation
- Electroporator (if using electrocompetent cells)
- Refrigerator for 4°C storage
- -80 °C freezer
- Minicentrifuge
- Centrifuge that can hold 50mL tubes and can reach 3214 x g
- Stationary incubator set to 37°C
- Growth chamber (for Arabidopsis)

Appendix 34: PCR amplification of guide RNA cassette

Biosafety Level: None – No hazardous biological agents used

Preparation Notes:

Other types of high-fidelity polymerases can be used for amplification. Adjust PCR cocktail components accordingly but maintain the relative proportions of primers.

Prepare for Students:

- Tris-EDTA (TE) buffer (10mM Tris-HCl, 1mM EDTA•Na_2, pH 8.0)
- 10mM Tris-HCl, pH 8.0
- 1x TAE (40mM Tris, 20mM acetate, 1mM EDTA)
- pCBC-DT1T2 – 1ng per student (Or other appropriate guide RNA expression cassette. pCBC-DT1T2 is used for expression of two guide RNAs.)

Supply for Students:

- Deionized water (sterile)
- Ice
- Sharpies
- Disposable gloves (such as nitrile disposable gloves)
- Micropipettes and disposable tips (P1000, P200, P20, and P2 recommended. Students can share micropipettes if needed.)
- Safety glasses or goggles
- 1.5mL microcentrifuge tubes
- Lab coat
- 0.2mL PCR tubes
- PCR reagents (The protocol below uses Accuzyme DNA polymerase)
- 6X loading dye
- Agarose
- Heat resistant mittens or gloves
- DNA ladder (1kb ladder is recommended such as GeneRuler 1 kb Plus, Thermofisher Scientific. Expected product size for this activity is 676bp)
- PCR product purification kit
- Ethidium bromide (Final concentration in an agarose gel should be approximately 0.5µg/mL. We recommend using a working stock of 10mg/mL).
    - **ADDITIONAL SAFETY RECOMMENDATIONS:** Ethidium bromide is a mutagen and must be handled with care and disposed of correctly according to EH&S guidelines. We highly recommend having designated containers to collect ethidium bromide contaminated micropipette tips and agarose gels. We also recommend having a designated lab space with 1 or 2 P2 micropipettes to add ethidium bromide to agarose gels to minimize the risk of contaminating other lab surfaces with ethidium bromide. Alternatively, other nonhazardous gel stains, such as SYBR Safe gel stain can be substituted for ethidium bromide. Note that equipment needed to image alternative gel stains may vary.

Equipment:

- Thermocycler (Bio-Rad)
- Microwave (**ADDITIONAL SAFETY NOTE:** microwave should be dedicated for lab use only. No food or beverages should be heated in the microwave.)
- Gel electrophoresis system (including combs, gel tray, power supply and leads, and gel box)
- UV transilluminator
- Camera (To take pictures of gel. Cell phone cameras are acceptable.)
- Tabletop centrifuge or minicentrifuge (To spin down PCR tubes.)
- Nanodrop
- -20 °C freezer
- Vortex mixer (Alternatively, reagents can be mixed by gently flicking the tube.)

Student Activity:
Estimated time of activity: 3 hours

The two guide sequences you provided were used to create four long primers that we will use in today's PCR reaction.

Today you will prepare the primers and set up a PCR reaction that will incorporate your guide sequences into the ends of a DNA fragment that we will clone into a Cas9 vector in the next class.

1. **Dilute the primer stock provided by IDT (it has a blue cap).**

Look for the primer yield on the side of the tube that is specified in **nmol** (next to the OD value). Typically, this might be in the 10 to 35 nmol range. Multiple the **nmol** value x 10 to know the volume in **µl** to dissolve that primer with TE buffer to achieve a **100 µM** solution. Every tube is different, so you will need to adjust your volume accordingly.
e.g. 22.5 nmol is diluted with 225 µl TE buffer.

Mix well. Spin down briefly.

2. **Create working stocks of your four primers.**

All tubes need to be labeled with

- primer name
- concentration
- date and initials

Make working stocks of your -**BsF** and -**BsR** primers at final concentration **20 µM**.
For each primer, add into a clean labeled 1.5 mL tube

80 µl of DI water
20 µl of blue-cap primer stock (100 µM)

Vortex very briefly and/or flick the tube 20 times. Spin down a few seconds.

Make working stocks of your -**F0** and -**R0** primers at final concentration **1 µM**.
For each primer, add into a clean labeled 1.5 mL tube

990 µl DI water
10 µl of blue-cap primer stock (100 µM)

Vortex very briefly and/or flick the tube 20 times. Spin down a few seconds.

3. Make a mixture of the four primers.

Into a clean tube labeled with the student's name and primers added

10 uL -BsF primer (20 µM)
10 uL -BsR primer (20 µM)
10 uL -F0 primer (1 µM)
10 uL -R0 primer (1 µM)

Vortex very briefly and/or flick the tube 20 times. Spin down a few seconds.

4. Set up the PCR reaction.

To a clean PCR tube add

21 µL diluted pCBC-DT1T2 template (~1 ng total template)
4 µL primer mixture

Set pipettor to 10 µL and pipette mixture up and down 10 times. Keep on ice.

Then, add

25 µl Accuzyme 2X mix

Set pipettor to 20 µL and gently pipette mixture up and down 10 times. Cap the tube. Keep on ice until ready to move to the thermocycler.

5. Run PCR reaction

Initial denaturation	95 °C 3 minutes
Cycle (repeat 30-35x)	95 °C 15 seconds
	60 °C 30 seconds
	72 °C 1 minute
Final extension	72 °C 5 minutes
Store	10 °C forever

6. Check for a PCR product.

While your PCR is running, set up an agarose gel. You may share a gel with one or more people. When the PCR cycle is complete, run a sample (5 µL) of the PCR product on the gel and check for a 676 bp product. Don't forget to add loading dye to your 5 µL sample and add a ladder to the gel. Save a picture of your gel.

Keep the remainder of the PCR product stored at -20 °C. If you have the correct PCR product, we will purify it later and use it in the cloning reaction.

Example protocol for running a gel:
We will use a 1.5% agarose gel (0.75g agarose for a 50mL gel).

1. Assemble a gel rig to make a cast for your gel with the necessary combs for wells.
2. Weigh 0.75g of agarose.
3. Pour agarose into a round-bottom boiling flask.
4. Add 50mL of 1X TAE, placing int the microwave until it boils and the agarose dissolves (approximately 1 minute).
5. Remove the flask using heat resistant mittens or gloves (Caution: flask is hot!) Gently swirl the flask to ensure all agarose is dissolved.
6. Add 2µl of 10mg/ml ethidium bromide and swirl again to mix (Caution: ethidium bromide is a mutagen. Extra care should be taken to avoid contact with skin. Designated containers should be used to collect ethidium bromide contaminated micropipette tips).
7. Pour the hot liquid solution into the cast.

Example protocol for running a gel:

1. On a clean piece of parafilm or in a clean 0.2mL PCR tube mix 1µl of 6X loading dye with 5µl of PCR reaction. Pipette to mix well.
2. Load 5µl of the PCR reaction into a well in the agarose gel.
3. Load 5µl of DNA ladder in one lane of the gel.
4. Connect the gel apparatus to the power source. Voltage and time may vary. (120V for 30-45 minutes is usually sufficient).
5. Visualize the gel using a UV transilluminator and take a picture.

7. Purify PCR product. *(can be done this day or on the next lab day)*

If a correctly sized PCR product is visible, follow instructions for PCR cleanup using a spin purification kit. After elution in water or 10 mM Tris HCl, pH 8 (TE buffer is not recommended due to EDTA absorbance and short-term needs for storage), quantify DNA concentration with Nanodrop or similar method. Store purified PCR product at -20 °C.

Appendix 35: Golden Gate cloning protocol and *E. coli* transformation

Biosafety Level: BSL-1 – *E. coli* transformation

Preparation Notes:

Golden Gate assembly is done with BsaI and a high-concentration DNA ligase (see instructions from appendix in Wang et al., 2015). For convenience, uniformity, and cost-savings, we recommend NEB® Golden Gate Assembly Kit (BsaI-HF®v2). Costs might be reduced further by attempting half-reactions.

This protocol works with pHEE401E (Wang et al., 2015, *Genome Biology*), pYUU (Angulo et al. 2020 *BioRxiv*), and other similar vectors derived from pHEE401E. pHEE401E (catalog #71287 from Addgene) uses hygromycin selection in plants and an egg cell-specific promoter to express Cas9. pYUU uses a yellow fluorescent protein expressed in seeds as a selectable marker and a ubiquitin promoter to drive expression of Cas9.

Escherichia coli (*E.coli*) used in this experiment has been rendered nonpathogenic. However, follow the EHS guidelines to dispose of waste. It takes two days to prepare electrocompetent *E. coli* cells.

Ensure that all pipette tips and tubes that contacted *E. coli* are disposed of in biohazard bags and autoclaved.

Prepare for Students:

- SOC or 2X YT media
- LB agar plates with 50 µg/mL kanamycin (for pHEE401E or pYUU. Use appropriate antibiotic selection if a different vector is used)
  Optional: *It is usually difficult for students to accurately pipette and distribute 1 µL, so we recommend making a cocktail for the class and aliquoting 15 µL of it to each student for step 2.*

  Each student will need the following per reaction:
  2 µL T4 DNA Ligase Buffer (10x)
  1 µL NEB Golden Gate Assembly Mix
  300 ng pYUU or pHEE401E, etc.
  water to a total volume of 15 µL

Supply for Students:

- Deionized water (sterile)
- Ice
- Sharpies
- Disposable gloves (such as nitrile disposable gloves)
- Micropipettes and disposable tips (P1000, P200, P20, and P2 recommended. Students can share micropipettes if needed.)
- Safety glasses or goggles
- Lab coat

- 10X T4 DNA ligase buffer
- NEB Golden Gate Assembly mix (recommended)
- pHEE401E, pYUU, or similar vector (~300ng per student)
- Purified PCR product from guide RNA cassette PCR (~20ng per student)
- Chemically competent or electrocompetent *E. coli*
- Electroporation cuvette (if using electrocompetent *E. coli*)
- Parafilm or plastic wrap
- Biohazard bags (to dispose of *E. coli* contaminated materials)

Equipment:
- 37°C incubator with shaking
- Stationary incubator set to 37°C
- Heat block or bead bath for 60 °C incubation
- Electroporator (if using electrocompetent cells)
- Refrigerator for 4°C storage
- -80 °C freezer (for storage of competent cells and glycerol stocks)

Student Activity:
Estimated time of activity: 3 hours

<u>Golden Gate cloning into CRISPR/Cas9 vector</u>

1. Dilute a sample of purified PCR product to 4 ng/µL concentration.

2. Assemble Golden Gate reaction mixture or distribute 15µL prepared master mix to each student

   2 µL T4 DNA Ligase Buffer (10X)
   1 µL NEB Golden Gate Assembly Mix
   300 ng pYUU or pHEE401E, etc.
   water to a total volume of 15 µL

3. Add 5 µL of the diluted PCR product (20 ng PCR product total) to the Golden Gate cocktail. Mix gently, but thoroughly, by pipetting.

4. Incubate the 20 µL Golden Gate reaction at 37 °C for 1 hour, then at 60 °C for 5 minutes. If time is limited, the 37 °C step may be shortened.

5. Dilute the Golden Gate reaction with 20 µL of water.

***E. coli* transformation of Golden Gate reaction**

The purpose of this transformation is to introduce foreign DNA into *E. coli*, so that it will replicate the plasmid. Also, the bacteria will be used as a means for storing the plasmid for a long time at -80°C. Most plasmids carry a selection marker and an origin of replication for these purposes.

1. Remove *E. coli* competent cells from the -80°C freezer and thaw on ice.

2. Add 2 μL of diluted Golden Gate reaction to the thawed cells and mix thoroughly with a p200.

3. Pipette cells into sterile, ice-cold 1mm electroporation cuvettes

4. Gently tap the cuvette to remove air bubbles and then electroporate per the device instructions.

5. Immediately add 1 mL of SOC or 2xYT media to resuspend electroporated cells. Mix thoroughly with a P1000 and transfer to a clean 1.5 mL Eppendorf tube.

6. Recover cells for 1 hr at 37°C with shaking at 200 rpm.

7. After 1 hr, pipette 100 μL on a LB agar plate + 50 μg/mL kanamycin (working concentration). Spin down the remaining transformation mixture and resuspend the pellet in 100-200 μL LB media. Plate the resuspension on a second plate.

8. Grow the bacteria at 37°C with the agar side up overnight. Wrap with parafilm and store at 4°C for up to 1 month.

Appendix 36: Colony PCR of *E. coli* Transformants

Biosafety Level: BSL-1 – Handling recombinant *E. coli*

Preparation Notes:

Ensure that all pipette tips that contacted *E. coli* are disposed of in biohazard bags and autoclaved.

Prepare for Students:

- 20μM of the following primers. Each student will need 0.5 μLof each primer.
    - U6-26p-F: 5'-TGTCCCAGGATTAGAATGATTAGGC
    - U6-29p-R: 5'-AGCCCTCTTCTTTCGATCCATCAAC
- 1x TAE (40mM Tris, 20mM acetate, 1mM EDTA)

Supply for Students:

- 1.5mL microcentrifuge tubes
- Sharpies
- 0.2mL PCR tubes
- Disposable gloves (such as nitrile disposable gloves)
- Micropipettes and disposable tips (P1000, P200, P20, and P2 recommended. Students can share micropipettes if needed.)
- Safety goggles or glasses
- Lab coat
- Deionized water (sterile)
- PCR reagents (Will vary depending on polymerase used. A high-fidelity polymerase is not required for colony PCR. A master mix of PCR reagents can also be prepared ahead of time and distributed to students.)
- Agarose
- DNA ladder (GeneRuler 1 kb Plus, Thermofisher Scientific)
- Parafilm or plastic wrap
- Biohazard bags
- Ethidium bromide (Final concentration in an agarose gel should be approximately 0.5μg/mL. We recommend using a working stock of 10mg/mL).
    - **ADDITIONAL SAFETY RECOMMENDATIONS:** Ethidium bromide is a mutagen and must be handled with care and disposed of correctly according to EH&S guidelines. We highly recommend having designated containers to collect ethidium bromide contaminated micropipette tips and agarose gels. We also recommend having a designated lab space with 1 or 2 P2 micropipettes to add ethidium bromide to agarose gels to minimize the risk of contaminating other lab surfaces with ethidium bromide. Alternatively, other nonhazardous gel stains, such as SYBR Safe gel stain can be substituted for ethidium bromide. Note that equipment needed to image alternative gel stains may vary.

Equipment:

- Thermocycler (Bio-Rad)

- Microwave (**ADDITIONAL SAFETY NOTE:** microwave should be dedicated for lab use only. No food or beverages should be heated in the microwave.)
- Gel electrophoresis system (including combs, gel tray, power supply and leads, and gel box)
- UV transilluminator
- Camera (To take pictures of gel. Cell phone cameras are acceptable.)
- Tabletop centrifuge or minicentrifuge (To spin down PCR tubes.)
- Refrigerator for 4°C storage

Student Activity:

Estimated time of activity: 2 hours

Colony PCR is a fast way to genotype bacterial colonies after transformation. The way that we extract DNA from bacteria is very crude, but it's not necessary to have purified DNA for colony PCR. After the DNA extraction, you will incubate your reaction in the thermocycler to amplify your region of interest.

Perform colony PCR with primers U6-26p-F and U6-29p-R.

1. Add 10 μL of sterile water in 0.2 mL PCR tubes for each bacterial colony to be genotyped. 7 colonies per student plus a negative control with no bacteria is recommended. The number of colonies tested can vary though based on transformation efficiency and thermocycler space available.

2. Use a Sharpie to mark and number which colonies you are going to test using colony PCR on the bottom of the plate.

3. With a 2 μL pipette tip, add a small amount of colony to each PCR tube. You can swirl the pipette tip to mix the bacteria in the water.

4. Heat the PCR tubes to 95°C for 10 minutes inside a thermocycler. This will lyse the cells and expose crude DNA that can be used for genotyping.

Reaction setup	25 μL reaction
10x *Taq* buffer	2.5 μL
10 mM dNTPs	0.5 μL
10 uM U6-26p-F	0.5 μL
10 uM U6-29p-R	0.5 μL
DNA	2 μL
Taq polymerase	0.125 μL
Sterile H_2O	18.875 μL

PCR conditions	
Initial Denaturation	95°C 3 min
30 cycles	95°C 20 sec
	52°C 30 sec
	68°C 45 sec
Final extension	68°C 5 min
Hold	4°C forever

5. Analyze PCR products with gel electrophoresis. Successful insertion of the guide RNA cassette during GoldenGate cloning will produce a 726 bp product.

6. Students should clearly indicate on the bottom of their plates with Sharpie which colonies had successful amplification of the guide RNA cassette. Plates should be wrapped in parafilm of plastic wrap and stored at 4°C.

Appendix 37: DNA minipreps of CRISPR/Cas9 constructs

Biosafety Level: BSL-1 – Handling recombinant *E. coli*
Preparation Notes:

The day before students meet for lab, the instructor or lab assistant should inoculate 15 mL culture tubes containing ~2-3 mL of LB media containing 50 µg/mL kanamycin (for pHEE401E or pYUU. Use appropriate antibiotic selection if a different vector is used) with *E. coli* colonies that had successful amplification by colony PCR the previous lab day. Shake overnight at 37°C at 200-250 rpm.

Ensure that all pipette tips and tubes that contacted *E. coli* are disposed of in biohazard bags and autoclaved.

Prepare for Students:

- Overnight *E. coli* cultures for candidate colonies (as described in the Preparation Notes above)
- Sterile 30% (v/v) glycerol
- Miniprep kit or plasmid extraction reagents

If using alkaline-lysis for plasmid extraction we recommend the following or a similar protocol:

He, F. (2011). Plasmid DNA Extraction from *E. coli* Using Alkaline Lysis Method. *Bio-101*: e30. DOI: 10.21769/BioProtoc.30.

If using a miniprep kit we recommend NucleoSpin Plasmid EasyPure from Takara or a similar kit.

Supply for Students:

- 1.5 mL microcentrifuge tubes
- 2.0 mL sterile cryogenic vials for glycerol stock
- Disposable gloves (such as nitrile disposable gloves)
- Micropipettes and disposable tips (P1000, P200, P20, and P2 recommended. Students can share micropipettes if needed.)
- Safety goggles or glasses
- Lab coat
- Biohazard bags

Equipment:

- Tabletop centrifuge (it may be helpful to have more than one for if there are many students)
- Nanodrop
- -80 °C freezer

Student Activity:

Estimated time of activity: 2 hours

1. It is advisable to create a glycerol stock from each culture by combining 700µL with 700µL sterile 30% (v/v) glycerol in a sterile 1.5 mL microcentrifuge tube. Store at -80°C.

2. Spin down the remaining culture and perform a plasmid miniprep extraction. This can be done with an inexpensive alkaline-lysis method, or if time is limited with a spin miniprep kit.

3. Quantify plasmid miniprep yield, using a Nanodrop.

4. Send plasmid samples for sequencing. A good Sanger sequencing reaction can cover both guides with just the U6-26p-F primer. But because the second guide sequence is ~900 bp away from the U6-26p-F primer, using a second primer, U6-29p-F, may also be useful. Students can help set up sequencing reactions if time allows otherwise sequencing reactions can be prepared by the instructor or lab assistant.

   U6-26p-F: 5'-TGTCCCAGGATTAGAATGATTAGGC
   U6-29p-F: 5'-TTAATCCAAACTACTGCAGCCTGAC

Appendix 38: Transformation of *Agrobacterium tumefaciens*

Biosafety Level: BSL-1 – Transforming *Agrobacterium tumefaciens*

Preparation Notes:

It takes three days to prepare electrocompetent *A. tumefaciens* (GV3101) cells. Instructor or lab assistant will prepare the cells.

Ensure that all pipette tips and tubes that contacted *A. tumefaciens* are disposed of in biohazard bags and autoclaved.

Prepare for Students:

- SOC media enough for ~1mL per student
    - 5g yeast extract, 0.5g NaCl, 20g Tryptone, bring volume to 980mL with DI water and autoclave
    - allow it to cool to 60°C or less. Add 20 mL of a sterile 1 M solution of glucose
        - To make 1M sterile glucose: dissolve 18 g of glucose in 90 mL of H_2O. After the sugar has dissolved, adjust the volume of the solution to 100 mL with H_2O and sterilize by passing it through a 0.22-µm filter (Millipore). Store the solution at 4 °C for short term storage.
- LB plates with rifampicin (20 µg/mL), gentamicin (50 µg/mL), and plasmid selection (ex. kanamycin 50 µg/mL for pHEE401E or pYUU). Enough for 1-2 plates per student.

Supply for Students:

- Plasmid DNA for CRISPR-Cas9 construct (quality confirmed by colony PCR and/or sequencing)
- Competent *Agrobacterium tumefaciens* (electrocompetent bacteria is ideal because the transformation process is faster, but thermo competent bacteria could be used as well. The protocol below is for transforming electrocompetent *A. tumefaciens.*)
- 1mm electroporation cuvette (if electrocompetent cells are used)
- Ice
- Parafilm or plastic wrap
- Disposable gloves (such as nitrile disposable gloves)
- Micropipettes and disposable tips (P1000, P200, P20, and P2 recommended. Students can share micropipettes if needed.)
- Safety goggles or glasses
- Lab coat
- Biohazard bags

Equipment:

- Electroporator (if electrocompetent cells are used)
- Shaking incubator set to 30°C
- Stationary incubator set to 30°C (*Agrobacterium tumefaciens* can grow at room temperature~28°C in the dark if an incubator is unavailable)
- Refrigerator (4°C storage)

Student Activity:

Estimated time of activity: 3 hours

Agrobacterium tumefaciens is a soil bacteria that causes crown gall disease in plants by introducing its own DNA into the host genome. The infection process involves the transfer of DNA.

1. Thaw *Agrobacterium tumefaciens* electrocompetent cells on ice.

2. Add 100 ng of plasmid DNA to the cells and mix with a pipette or by flicking the tube.

3. Transfer the cells into sterile, ice-cold 1mm electroporation cuvettes. Gently tap the cuvette on the bench to remove bubbles and then electroporate.

4. Immediately add ~1 mL of SOC media.

5. Recover cells at 30 °C with shaking for 3 hrs.

6. Plate cells on LB + rif + gent + plasmid selection marker. It is advisable to plate different amounts of the transformation (e.g. 10 μL and 100 μL). Typically, transformation is very high.

7. After ~2 days, when colonies are visible, remove plates from the incubator, wrap with parafilm, and store upside-down at 4 °C. Optional: perform colony PCR to validate plasmid uptake in select colonies.

Appendix 39: Sequence Analysis

Compare sequence data for individual clones to the following template. This template sequence shows N_{19} for each of the two guide sequences (bold, highlighted). If 20-nt guide sequences were used, extend the N stretches by one base. Sequencing primers are underlined. Guide RNA transcription begins at the bold red G nucleotides.

CGACTTGCCTTCCGCACAATACATCATTTCTTCTTAGCTTTTTTTCTTCTTCTTCGTTCA
TACAGTTTTTTTTTGTTTATCAGCTTACATTTTCTTGAACCGTAGCTTTCGTTTTCTTCT
TTTTAACTTTCCATTCGGAGTTTTTGTATCTTGTTTCATAGTTTGTCCCAGGATTAGAAT
GATTAGGCATCGAACCTTCAAGAATTTGATTGAATAAAACATCTTCATTCTTAAGATATG
AAGATAATCTTCAAAAGGCCCCTGGGAATCTGAAAGAAGAGAAGCAGGCCCATTTATATG
GGAAAGAACAATAGTATTTCTTATATAGGCCCATTTAAGTTGAAAACAATCTTCAAAAGT
CCCACATCGCTTAGATAAGAAAACGAAGCTGAGTTTATATACAGCTAGAGTCGAAGTAGT
GATTGNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT
AGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTTGCAAAATTTTCC
AGATCGATTTCTTCTTCCTCTGTTCTTCGGCGTTCAATTTCTGGGGTTTTCTCTTCGTTT
TCTGTAACTGAAACCTAAAATTTGACCTAAAAAAAAATCTCAAATAATATGATTCAGTGGT
TTTGTACTTTTCAGTTAGTTGAGTTTTGCAGTTCCGATGAGATAAACCAATATTAATCCA
AACTACTGCAGCCTGACAGACAAATGAGGATGCAAACAATTTTAAAGTTTATCTAACGCT
AGCTGTTTTGTTTCTTCTCTCTGGTGCACCAACGACGGCGTTTTCTCAATCATAAAGAGG
CTTGTTTTACTTAAGGCCAATAATGTTGATGGATCGAAAGAAGAGGGCTTTTAATAAACG
AGCCCGTTTAAGCTGTAAACGATGTCAAAAACATCCCACATCGTTCAGTTGAAAATAGAA
GCTCTGTTTATATATTGGTAGAGTCGACTAAGAGATTGNNNNNNNNNNNNNNNNNNGTT
TTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGC
ACCGAGTCGGTGCTTTTTTTTGCAAAATTTTCCAGATCGATTTCTTCTTCCTCTGTTCTT
CGGCGTTCAATTTCTGGGGTTTTCTCTTCGTTTTCTGTAACTGAAACCTAAAATTTGACC
TAAAAAAAATCTCAAATAATATGATTCAGTGGTTTTGTACTTTTCAGTTAGTTGAGTTTT
GCAGTTCCGATGAGATAAACCAATC

Appendix 40: Transformation of *Arabidopsis thaliana*

Biosafety Level: BSL-1 – Handling transgenic *Agrobacterium tumefaciens*

Preparation Notes:

Typically wild type *Arabidopsis thaliana* plants take 3-4 weeks under long-day conditions (16h light/8h dark, 20°C) to reach an ideal age for dipping. However, consideration should be given if the mutant being dipped has an altered growth phenotype. Plants should be flowering with lots of unopened buds and few siliques. Make sure plants are well-watered the day before transformation.

Ensure that all pipette tips and tubes that contacted *A. tumefaciens* are disposed of in biohazard bags and autoclaved. Liquids solutions contaminated with *A. tumefaciens* should have 10% of the total volume in bleach added to the solution. Let the bleach and bacteria solution sit for at least 30 minutes at room temperature before rinsing down the drain with plenty of water.

Prepare for Students:

- Plants of the appropriate mutant background (These should be planted enough time in advance so they will be ready to dip.)
- *Agrobacterium tumefaciens* overnight cultures
    - The afternoon before class meets, set up 20 mL cultures of transgenic *Agrobacterium tumefaciens* in LB media with rifampicin (20 µg/mL), gentamicin (50 µg/mL), and plasmid selection in 50 ml centrifuge tubes. Pick a single colony for each CRISPR-Cas9 construct. Grow overnight at 28-30 °C with shaking. A wide range of culture densities are tolerated.
- Sterile 50% (v/v) glycerol (Optional: 500 µL can be aliquoted into cryogenic vials beforehand and distributed to students in lab)
- 5% (w/v) sucrose

Supply for Students:

- 1.8-2 mL cryogenic vial (1.5 mL microcentrifuge tubes can be used as well)
- Disposable gloves (such as nitrile disposable gloves)
- Micropipettes and disposable tips (P1000, P200, P20, and P2 recommended. Students can share micropipettes if needed.)
- Safety glasses or goggles
- Lab coat
- Bleach
- Biohazard bags
- Masking tape and sharpies (for labeling pots)
- Large flasks or beakers (to pour off excess media into and later dispose of after adding bleach)
- Silwet L-77
- disposable plastic transfer pipette or petri dishes
- Flats for storing Arabidopsis
- Plastic domes (that fit the Arabidopsis flats) or plastic wrap

Equipment:

- -80 °C freezer
- Centrifuge that can hold 50mL tubes and can reach 3214 x g
- Growth chamber to grow and maintain *A. thaliana* plants used

Student Activity:

Estimated time of activity: 2 hours

1. Create a glycerol stock of each culture. Add 500 μL of culture to a sterile, labeled cryogenic vial containing 500 μL of sterile 50% (v/v) glycerol. Mix well and store at -80 °C.

2. Spin remaining culture for 10 minutes at 3900 rpm (3214 x g) (max speed for 50 mL centrifuge 5810R in Eppendorf) at room temperature. Make sure buckets and tubes are adequately balanced first.

3. Pour off supernatants from cultures into a large flask. Later, add bleach to 10% concentration and dispose per biosafety regulations.

4. Add 10 mL of 5% (w/v) sucrose to the pellet. Shake the centrifuge tube vigorously for the cells to resuspend. The sucrose solution can be made the same day and does not need to be sterilized.

5. Make sure each pot of plants is labeled with correct construct identifiers (e.g. student's name, construct name, and date).

6. Immediately before transformation, add 5 μL of Silwet L-77 to the resuspension (final concentration is 0.025% v/v). Mix well but not to the point of excessive foaminess.

7. Use a disposable plastic transfer pipette to add generous drops of the resuspension to any flowers and bud tissues. Add to the base of cauline leaves on the primary shoot and the base of the rosette leaves where axillary buds may not be fully visible. Drops can be applied repeatedly. Alternative methods include pouring *Agrobacterium* solution into a Petri dish and soaking flower buds in it for 10-30 seconds with gentle movement.

8. Lay plants on their side in a clean flat. Make sure there is adequate space from other transformed plants to prevent cross-contamination. Cover with a plastic dome and another flat to keep plants under high humidity in the dark overnight.

9. Dispose of the remaining solution, if any, in the flask with bleach. Dispose of gloves, centrifuge tube, and transfer pipette in biohazardous waste bin for autoclaving.

The day after – to be done by instructor or lab assistant

10. Remove covers the next day and grow to maturation as normal. Transgenic seeds will be ready to harvest and undergo selection in ~3 to 4 weeks.

Appendix 41: Cumulative Exam Scores

Cumulative exam scores for all students that participated in the in-person CURE (Spring 2019) and remote learning CURE (Spring 2020 and Fall 2020).

	Spring 2019; n =116	**Spring 2020; n=157**	**Fall 2020; n=184**
Average	72.61	85.83	81.62
Standard Deviation	13.91	11.31	11.53
Median	72	87	81

Appendix 42: Mapping Exam Questions to Learning Outcomes

Assessment of learning outcomes from 10 random tests each from the in-person CURE (Spring 2019) and the remote learning CURE (Spring 2020 and Fall 2020).

(1) Use online resources and databases to research *A. thaliana* gene information.

	S2020 – 1a	S2020 – 1b	S2020 – 1c	F2020 - 1
Points possible	4	4	4	20
Mean	4	4	4	19.1
Standard dev.	0	0	0	2.02

(2) Design PCR primers with the aid of online tools and databases to amplify their assigned *A. thaliana* gene.

	S2020 - 2	S2020 - 6	F2020 - 1
Points possible	15	6	20
Mean	13.2	5.1	19.1
Standard dev.	2.25	0.32	2.02

3) Identify homologs of their gene and design guide RNAs targeting their gene and any close homologs.

	F2020 - 1
Points possible	20
Mean	19.1
Standard dev.	2.02

4) Define homology and identify gene homologs through DNA database searches.

	S2019 – 5a	S2019 – 5b	S2019 – 5c	S2020 - 3
Points possible	4	2	2	12
Mean	3.56	1.9	1.8	11.5
Standard dev.	1.01	0.32	0.63	1.08

5) Discuss the evolutionary relationships between paralogs and orthologs.

	S2019 – 5b	S2019 – 5c	S2020 - 3
Points possible	2	2	12
Mean	1.9	1.8	11.5
Standard dev.	0.32	0.63	1.08

6) Create a phylogenetic tree for a gene family using online tools and identify paralogs and orthologs for their assigned gene.

	S2020 - 3
Points possible	12
Mean	11.5
Standard dev.	1.08

(7) Discuss the mechanism of CRISPR-Cas9 mutagenesis.

	S2019 – 1a	S2019 – 3a	S2019 – 4a	S2020 – 3
Points possible	3	5	1	12
Mean	2.3	1.6	1	11.5
Standard dev.	1.25	1.82	0	1.08

8) Form a hypothesis about the potential consequences of mutating a gene and what that suggests about the gene's function.

	S2020 - 8
Points possible	8
Mean	7.3
Standard dev.	1.16

9) Design guide RNAs to specifically target different regions of a gene to create a null mutation.

	S2019 - 2	S2020 - 5	F2020 – 1
Points possible	15	14	20
Mean	11.3	13.6	19.1
Standard dev.	4.35	1.26	2.02

10) Evaluate and select guide RNAs based on a given set of criteria to selectively target a single gene.

	S2020 - 5	F2020 - 1
Points possible	14	20
Mean	13.6	19.1
Standard dev.	1.26	2.02

11) Design PCR experiment to detect successful gRNA guided deletions in *A. thaliana*.

	S2019 – 4b	S2020 - 6	S2020 - 7	F2020 - 1
Points possible	9	6	15	20
Mean	9	5.1	13.2	19.1
Standard dev.	0	0.32	2.53	2.02

Appendix 43. Pre and post CURE Student Survey Data

Likert-style surveys designed by the course instructors administered before and after the CURE for in-person students (Spring 2019) and remote learning students (Spring 2020). Students scored their responses on a scale of 1-7 where 1=strongly disagree and 7=strongly agree. “NA” indicates that the question was not included in the pre CURE survey. Percentages in the “Research Involvement” subcategory indicate the percentage of students that responded “Yes”.

	Spring 2019 (in-person)		Spring 2020 (remote)	
	Pre CURE	Post Test	Pre CURE	Post CURE
Number of Respondents	84	108	123	147
Instructor Characteristics				
The instructor presented the course material with enthusiasm	NA	6.25	NA	5.69
The instructor encouraged me to participate in class discussions	NA	6.01	NA	5.67
The instructor encouraged me to interact with other students in class	NA	6.08	NA	5.06
The instructor was informative when responding to student's questions in class	NA	6.31	NA	5.53
The instructor could be contacted outside of class time	NA	6.21	NA	6.06
The instructor promoted a comfortable learning atmosphere	NA	6.38	NA	5.82
The instructor encouraged me to ask questions during class	NA	6.13	NA	5.95
The instructor's explanations of course content was clear	NA	6.05	NA	5.13
The instructor clarified course material by reviewing concepts	NA	6.24	NA	5.45
The instructor used a variety of approaches/strategies when presenting material	NA	5.91	NA	5.22
The instructor presented the course material at an appropriately paced sequence	NA	6.23	NA	5.54
The instructor helped me with my individual learning needs	NA	6.15	NA	5.5
The instructor provided me with feedback on my learning progress.	NA	5.9	NA	5.5
The instructor’s feedback on my learning progress was valuable.	NA	6.04	NA	5.54
Instructor Characteristics Composite Score	**NA**	**6.13**	**NA**	**5.55**

Teamwork				
I feel that learning how to work as a team with your colleagues is important in the Dynamic Genome Lab.	6.31	6.5	6.06	5.73
I get personal satisfaction when I solve a scientific problem by figuring it out with a team in the Dynamic Genome Lab.	5.92	6.12	5.82	5.78
Teamwork Composite Score	**6.11**	**6.31**	**5.94**	**5.76**
Critical Thinking				
My decisions are based on facts rather than on general impressions and feelings.	4.96	6.39	5.1	6.11
I wish science instructors would just tell us what we need to know so we can learn it.	2.94	3.19	2.87	2.76
Scientists know what the results of their experiments will be before they start.	5.34	4.16	5.06	3.9
There is too much emphasis in science classes on figuring things out for yourself.	4.07	3.86	4.11	3.14
My decisions are based on careful thinking and reasoning.	5.7	6.49	5.67	6.28
Even if I forget the facts, I'll still be able to use the thinking skills I learn in science.	5.43	6.27	5.54	5.89
Critical Thinking Composite Score	**4.98**	**5.04**	**4.94**	**4.68**
Lab Technique				
I feel comfortable running gels.	5.21	5.64	4.47	4.47
I understand how to properly pipette.	6.24	5.67	5.48	5.03
I feel comfortable using all the lab machines.	4.92	5.14	4.4	3.96
I feel comfortable with extracting DNA.	4.57	6.41	3.65	4.56
I understand how to analyze genomic DNA and cDNA.	3.54	6.21	2.54	5.68
I understand how PCR is used in lab.	4.02	6.49	2.59	5.77
I am confident in my ability to make good master mixes.	4.86	6.06	3.29	4.46
Lab Techniques Composite Score	**4.77**	**5.94**	**3.77**	**4.85**
Purposeful Work				
The instructor explained the situations under which course content could be applied.	NA	6.2	NA	5.68
The instructor increased my interest in the course subject matter.	NA	6.08	NA	5.32
The instructor provided me with hands-on activities with the course subject matter.	NA	6.46	NA	4.9
The research I do in this class is helpful to people in the real world.	NA	6.19	NA	5.56

Purposeful Work Composite Score	**NA**	**6.23**	**NA**	**5.37**
STEM Understanding				
When it comes to the STEM field, I really know a lot.	4.32	4.42	4.44	4.28
Compared to most other people, I know more about the STEM field.	4.65	4.58	4.47	4.7
I feel very knowledgeable about the STEM field.	4.65	4.66	4.39	4.59
STEM Understanding Composite Score	**4.54**	**4.55**	**4.43**	**4.53**
STEM Interest				
I know the variety of career options for science majors.	5.06	5.44	4.95	5.1
I have a good understanding of what a career in scientific research entails.	4.79	5.32	4.67	4.95
STEM Interest Composite Score	**4.92**	**5.37**	**4.81**	**5.02**
Future Plans				
My future plans include: Biology related graduate school	3.95	4.6	4.24	4.62
My future plans include: Medical school	4.76	5.05	5.65	5.29
My future plans include: Pursuing a career in the medical field, but NOT medical school	4.12	4.7	4.01	4.22
My future plans include: Teaching K-12 science	1.51	2.55	1.49	2.1
My future plans include: - Pharmacy school	2.17	2.86	2.56	2.61
My future plans include: - Biotech	2.46	3.15	2.46	2.94
My future plans include: - I am undecided about my future plans.	2.54	3.06	2.37	2.75
Research Involvement				
I am currently in an independent faculty research lab at UCR.	7.10%	8.30%	6.50%	5.50%
I am interested in working with a specific professor.	23.80%	88.90%	24.60%	90.30%