JOURNAL OF MICROBIOLOGY & BIOLOGY EDUCATION
DOI: https://doi.org/10.1128/jmbe.v22i1.2331

JMBE Journal of Microbiology & Biology Education

Supplemental Materials

for

Student Annotations of Published Data as a Collaboration between an Online Laboratory Course and the C. elegans Database, WormBase

Caroline L. Dahlberg[1]*, Christian A. Grove[2], Heino Hulsey-Vincent[1], and Samiya Ismail[1]

[1]Biology Department, Western Washington University, Bellingham, WA 98225;
[2]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

Table of Contents

(Total pages 11)

*Corresponding author. Mailing address: Western Washington University, Biology Department, MS 9160, 516 High Street, Bellingham, WA 98225.
Phone: 360-650-4671. E-mail: lina.dahlberg@wwu.edu.
Received: 30 September 2020, Accepted: 8 November 2020, Published: 31 March 2021

Appendix 1. Protocol for Student Submission of Annotations to WormBase

This protocol outlines how to provide information about genotype/phenotypes to the scientific community by describing individual published data points. For this assignment, you will use your instructor's WormBase Person ID to enter your annotations; you will use the Comments box (see below) to make sure that we know what individuals annotated which data. An annotation for a piece of data is made up of the collected information that is assembled by filling out the appropriate text boxes using gene/allele names, short phrases, suggestions from drop down menus, and quoted text from the primary literature (scientific manuscript) that is being annotated.

Open your paper and Wormbase.
On the Wormbase front page navigate to:
Wormbase->submit data (under search bar)-> Phenotype data submission (under community curation)>submit online form

For your name: enter InstructorsLastName
Email: (will autofill)
PubMed ID: Pull from the paper (PMID)
[Alternatively: use the last author's name in the "your name" box; then click on the link to "review your publications and see which are in need of phenotype curation" and find your paper in the pop-up table. Once you have the PubMed ID, don't forget to switch "my name" back to InstructorsLastName!]

GENETIC PERTURBATION(S)

- RNAi target gene: N/A
- Allele name: Type in the name of the allele you are working with.
    - When entering allele name, click on the dropdown suggestion so the website recognizes it as more than plain text (if no suggestion pops up that's fine, continue with submission)
    - Only enter one allele per submission

and/or

- Overexpression transgene: Type in the transgene allele (use the *Ex* or *Is* nomenclature) you are working with. Important: transcriptional reporters and background genotypes (for example fusion proteins like GLR-1::GFP) that do not cause a phenotype should NOT be entered here. That information can go in the "Control genotype" field (see OPTIONAL, below).
    - Click on the dropdown suggestion so the website recognizes it as more than plain text (if no suggestion pops up that's fine, continue with submission)
    - Only enter one transgene per submission

Note from WormBase: If the strain with the phenotype had both an allele (e.g. a deletion allele) and a transgene and BOTH were causative, both will get simultaneously submitted as Multiple Perturbations, which appears after entering 2 or more genetic perturbations. This would be a complex genotype experiment.

PHENOTYPES

- Observed Phenotypes: this is what is described in the published data. There is an entire ontology (vocabulary) of phenotypes, and WormBase would like to fit your description to that. Try out some key words and see what pops up!
    - To enter the phenotype(s), click the autofill suggestions after you've entered a bit of text or a key word.
    - *After the phenotype is entered, a new box will open called 'phenotype remarks'*
    - Copy and paste the quote you pulled the phenotype from (please add quotation marks)
    - If the quote refers to multiple phenotypes, fill in another phenotype box, and copy and paste the same quote into the new phenotype remarks box

- Not observed Phenotypes: For our purposes, these are phenotypes that might have been expected but were not observed, or "negative results". This is often low priority, and you can ignore for the time being unless you are *confident* that a phenotype would have been expected or if the authors specifically note its absence.

For the "OPTIONAL" section

Click on the Optional header to enter other information *that you feel is important*.

- Add Control Genotype/strain details and any other information you feel confident about
- Add your name into the comments box at the bottom (this information is kept internal)

Press preview to check the information is entered properly (allele should have a WB# next to it)
****Take a screenshot of your page for your records and to turn in as a Canvas assignment) ****

Press submit, then reset to start a new entry

Appendix 2. Examples of student submissions and instructor comments.

Names and WormBaseIDs have been removed; used with permission.

A. Image of compiled student screen shots for the submission. These were turned in by the student via Canvas.
B. Image of instructor comments on the same student submission. These were uploaded for student review.
C. Example of student revision for the third submission. The student addresses the instructor comments and describes what changes were made. This student re-submitted the annotation, which was noted by the WormBase administrator.

A.

Contribute phenotype connections

Your Name	[redacted]		
Your E-mail Address	[redacted]		
PubMed ID	pmid24356955		
Overexpression Transgene	gk173034	Overexpressed Gene that causes Phenotype	wdr-48 -- WBGene00009441
Not Observed Phenotype	transgene subcellular localization variant -- WBPhenotype:0000679	Phenotype Remark	quoted from paper: "Similarly, we foundno alteration in GLR-1::GFP punctum intensity in anotherindependentwdr-48mutant,wdr-48 (gk173034) (Fig. 6,JandK). "
Comment	curated by [redacted]		

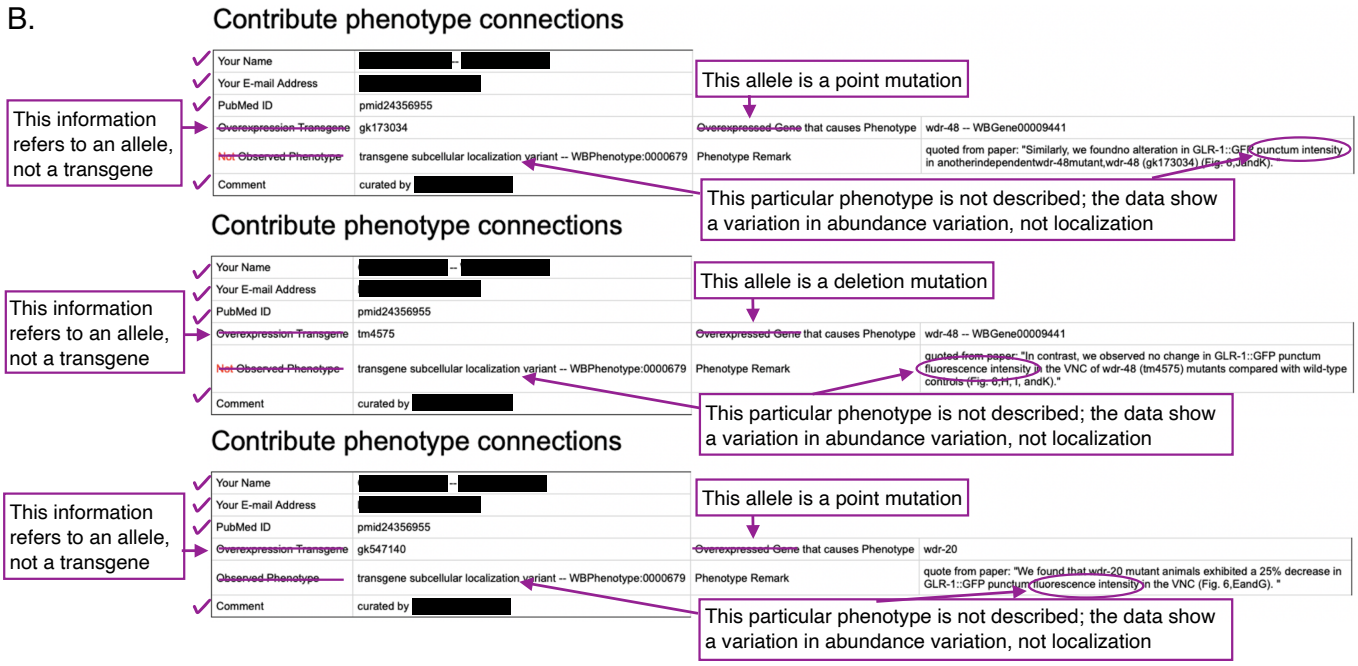
Contribute phenotype connections

Your Name	[redacted]		
Your E-mail Address	[redacted]		
PubMed ID	pmid24356955		
Overexpression Transgene	tm4575	Overexpressed Gene that causes Phenotype	wdr-48 -- WBGene00009441
Not Observed Phenotype	transgene subcellular localization variant -- WBPhenotype:0000679	Phenotype Remark	quoted from paper: "In contrast, we observed no change in GLR-1::GFP punctum fluorescence intensity in the VNC of wdr-48 (tm4575) mutants compared with wild-type controls (Fig. 6,H, I, andK)."
Comment	curated by [redacted]		

Contribute phenotype connections

Your Name	[redacted]		
Your E-mail Address	[redacted]		
PubMed ID	pmid24356955		
Overexpression Transgene	gk547140	Overexpressed Gene that causes Phenotype	wdr-20
Observed Phenotype	transgene subcellular localization variant -- WBPhenotype:0000679	Phenotype Remark	quote from paper: "We found that wdr-20 mutant animals exhibited a 25% decrease in GLR-1::GFP punctum fluorescence intensity in the VNC (Fig. 6,EandG). "
Comment	curated by [redacted]		

B.



C. Contribute phenotype connections

Your Name	[redacted]		
Your E-mail Address	[redacted]		
PubMed ID	24356955		
Allele	tm4575 -- WBVar00252973		
Not Observed Phenotype	transgene expression reduced -- WBPhenotype:0001278	Phenotype Remark	quoted from paper:" In contrast, we observed no change in GLR-1::GFP punctum fluorescence intensity in the VNC of wdr-48 (tm4575) mutants compared with wild-type controls (Fig. 6, H, I, and K)."
Comment	Curated by [redacted] tm4575 is a deletion mutation (quoted from paper): "The wdr-48 (tm4575) allele deletes 132 amino acids from the WDR-48 coding region, disrupting WD40 repeat domains 5–7, and is predicted to represent a null allele because of the introduction of a premature stop codon downstream of the deletion (Fig. 6A). "		

Here I made the correction to add what kind of mutation the allele was in the comments (a deletion allele), instead of including it as a transgene. The allele is a mutation not a transgene so this is why I changed it, and adding it in the comments allows a reader to know exactly what type of mutation it is and where it is. I also changed the Not Observed Phenotype from transgene subcellular localization variant to transgene expression reduced. This study was trying to determine if *wdr-48* was required in the interneurons for normal GLR-1 levels in the ventral nerve chord. A mutation is expected to causes a decrease in GLR-1::GFP, but no change is seen with the deletion, indicating it is not necessary for proper GLR-1 expression.

Contribute phenotype connections

Your Name	[redacted]		
Your E-mail Address	[redacted]		
PubMed ID	24356955		
Allele	gk547140 -- WBVar01040901		
Observed Phenotype	transgene expression reduced -- WBPhenotype:0001278	Phenotype Remark	quoted from paper: "We found that wdr-20 mutant animals exhibited a 25% decrease in GLR-1::GFP punctum fluorescence intensity in the VNC (Fig. 6, E and G)."
Comment	Curated by [redacted] gk547140 is a point mutation (quoted from paper): "The wdr-20 (gk547140) allele contains a non-sense mutation (G>A) that results in a premature stop codon at residue 288 of the WDR-20 protein."		

Here I made the correction to add what kind of mutation the allele was in the comments (a point mutation), instead of including it as a transgene. The allele is a mutation not a transgene so this is why I changed it, and adding it in the comments allows a reader to know exactly what type of mutation it is and where it is. I also changed the Observed Phenotype from transgene subcellular localization variant to transgene expression reduced. This study was trying to determine if *wdr-20* was required in the interneurons for normal GLR-1 levels in the ventral nerve chord. A mutation is expected to causes a decrease in GLR-1::GFP, and a 25% decrease is seen, indicating it is necessary for proper GLR-1 expression.

Contribute phenotype connections

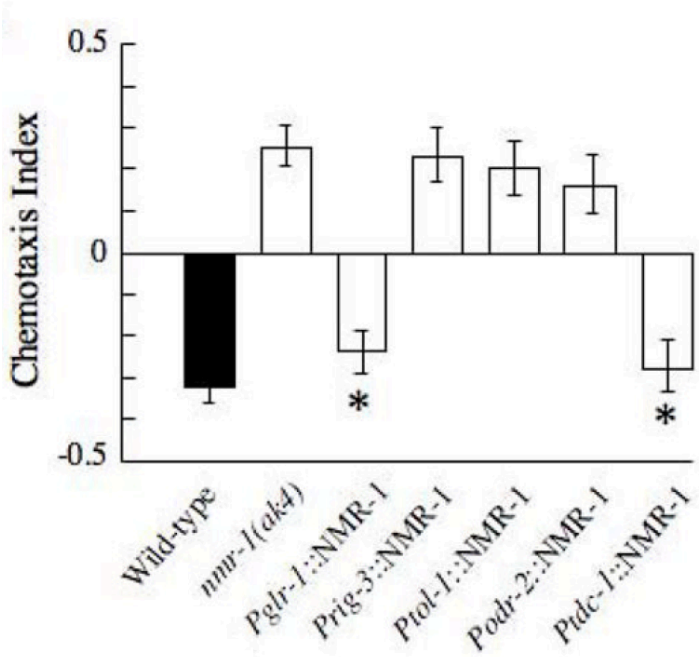
Your Name	[redacted]		
Your E-mail Address	[redacted]		
PubMed ID	24356955		
Allele	gk173034		
Not Observed Phenotype	transgene expression reduced -- WBPhenotype:0001278	Phenotype Remark	quoted from paper:"Similarly, we found no alteration in GLR-1::GFP punctum intensity in another independent wdr-48 mutant, wdr-48 (gk173034) (Fig. 6, J and K)."
Comment	Curated by [redacted] gk173034 has a point mutation (quoted from paper): "The wdr-48 (gk173034) allele contains a missense mutation (G>A) at a putative splice donor site of the first intron (Fig. 6A), creating alternative splice variants that ultimately result in premature stop codons (data not shown)."		

Here I made the correction to add what kind of mutation the allele was in the comments (a point mutation), instead of including it as a transgene. The allele is a mutation not a transgene so this is why I changed it, and adding it in the comments allows a reader to know exactly what type of mutation it is and where it is. I also changed the Not Observed Phenotype from transgene subcellular localization variant to transgene expression reduced. This study was trying to determine if *wdr-48* was required in the interneurons for normal GLR-1 levels in the ventral nerve chord. A mutation is expected to causes a decrease in GLR-1::GFP, but no change is seen with the point mutation, indicating it is not necessary for proper GLR-1 expression.

Appendix 3. Quiz to assess student learning.

This three-question quiz was administered using the Canvas LMS. Multiple choice (question 1) and fill in the blank (question 3) answers were provided as selections or drop-down options, respectively. The short answer question (question 2) was graded by the instructor.

The graph below shows that wild type *C. elegans* move away from high salt (a negative chemotaxis index) after training with an aversive stimulus; however animals lacking *nmr-1* (*ak4*) are attracted to salt (a positive chemotaxis index) even after the same training. (From Kano, et al., 2008).



1. *Which of the following is most likely to be true?*

   A. *rig-3* is expressed in the same cells as *nmr-1*.

   B. The *nmr-1* promoter drives *glr-1* expression.

   C. Both *tdc-1* and *glr-1* are expressed in at least some of the cells that express *nmr-1*.

   D. Both *tol-1* and *odr-2* are expressed in at least some of the cells that express *nmr-1*.

   (Correct answer: C)

2. Briefly, explain your choice.

   The nmr-1 mutant has a distinct chemotaxis defect that is only rescued when NMR-1 is expressed under the *glr-1* or *tdc-1* promoters.

3. Choose from the following choices to Fill in the blanks to complete a Wormbase data annotation for *nmr-1 (ak4)* based on the data from Kano, et al, 2008, above . The image below is for reference.

A. Allele.
   1. *nmr-1*
   2. *ak4*
   (Correct answer: 2)

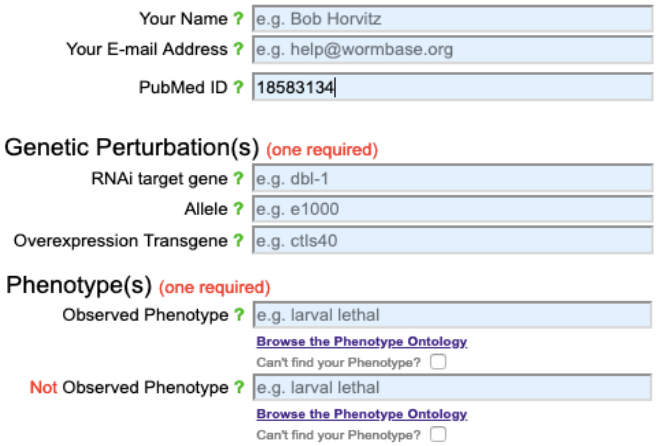
B. Overexpression transgene
   1. No transgene (you would leave this blank)
   2. Pglr-1::NMR-1
   (Correct answer: 1)

C. Observed phenotype
   1. Thermotaxis variant
   2. Sodium chloride chemotaxis variant

   (Correct answer: 2)

D. Not observed phenotype
   1. Positive chemotaxis variant
   2. No phenotype (you would leave this blank)

   (Correct answer: 2)



Appendix 4. Papers annotated by students and URLs of student annotations

A. Papers annotated by students are shown here by WormBase Paper ID and by PubMed ID

WBPaper ID	PubMed ID
WBPaper00002341	8601480
WBPaper00003257	9789046
WBPaper00005349	12123612
WBPaper00005404	12075001
WBPaper00006159	14573539
WBPaper00030893	17671168
WBPaper00031151	17972877
WBPaper00035944	20105303
WBPaper00041586	23013276
WBPaper00044602	24356955
WBPaper00046107	25469499
WBPaper00046459	25688864
WBPaper00049647	27223098
WBPaper00049891	27462879
WBPaper00050092	27593554
WBPaper00052970	28844202
WBPaper00055101	30147641
WBPaper00056441	30908491
WBPaper00057204	31413197
WBPaper00058677	31570707
WBPaper00058974	31740450
WBPaper00059155	31950452
WBPaper00059641	32350283

B. URLs of accepted student annotations.

These URLs are currently not shown for purposes of blinding. 23 unique papers were represented; the URLs will be provided following blinded review.

Appendix 5. Example of WormBase curation spreadsheet for student annotations.

Key to terminology:
"good" = validated annotation; will go into WormBase if it is not redundant
"dump" = will get 'dumped' out of the curation database and into the official WormBase database
"no dump" = will not get 'dumped' out of the curation database and into the official WormBase database; this annotation will remain internally in the curation database.

For this iteration of the CURE, 126 individual annotations were received and validated by WormBase.

Submission Date	WB Paper	PMID	ID	Action
4-24-2020	WBPaper00044602	pmid24356955	59244	good; will get dumped
4-26-2020	WBPaper00044602	pmid24356955	59275	dump from PGID 59244)
4-26-2020	WBPaper00044602	pmid24356955	59276	good; will get dumped
4-26-2020	WBPaper00044602	pmid24356955	59277	good; will get dumped
4-26-2020	WBPaper00044602	pmid24356955	59278	dump from PGID 59290)
4-26-2020	WBPaper00044602	pmid24356955	59279	good; will get dumped
4-26-2020	WBPaper00044602	pmid24356955	59281	dump from PGID 59244)
4-27-2020	WBPaper00044602	pmid24356955	59284	dump from PGID 59244)
4-27-2020	WBPaper00044602	pmid24356955	59288	dump from PGID 59423)
4-27-2020	WBPaper00044602	pmid24356955	59289	NO DUMP
4-27-2020	WBPaper00044602	pmid24356955	59290	good; will get dumped
4-27-2020	WBPaper00044602	pmid24356955	59291	with same transgene
4-27-2020	WBPaper00044602	pmid24356955	59295	good; will get dumped
4-27-2020	WBPaper00044602	pmid24356955	59301	dump from PGID 59290)
4-27-2020	WBPaper00044602	pmid24356955	59306	dump from PGID 59244)
4-27-2020	WBPaper00044602	pmid24356955	59307	good; will get dumped
4-27-2020	WBPaper00044602	pmid24356955	59308	dump from PGID 59244)
4-27-2020	WBPaper00044602	pmid24356955	59309	good; will get dumped
4-27-2020	WBPaper00044602	pmid24356955	59310	modified; will get dumped
4-27-2020	WBPaper00044602	pmid24356955	59311	dump from PGID 59295)
4-27-2020	WBPaper00044602	pmid24356955	59312	dump from PGID 59279)
4-27-2020	WBPaper00044602	pmid24356955	59314	dump from PGID 59279)
4-28-2020	WBPaper00044602	pmid24356955	59315	good; will get dumped
4-28-2020	WBPaper00044602	pmid24356955	59316	dump from PGID 59295)
5-17-2020	WBPaper00044602	pmid24356955	59420	dump from PGID 59244)
5-17-2020	WBPaper00044602	pmid24356955	59423	good; will get dumped
5-17-2020	WBPaper00044602	pmid24356955	59425	NO DUMP
4-24-2020	WBPaper00058677	pmid31570707	59245	good; will get dumped
4-24-2020	WBPaper00058677	pmid31570707	59246	dump from PGID 59245)
4-24-2020	WBPaper00058677	pmid31570707	59247	dump from PGID 59245)
4-24-2020	WBPaper00058677	pmid31570707	59248	dump from PGID 59245)
4-24-2020	WBPaper00058677	pmid31570707	59249	dump from PGID 59245)
4-24-2020	WBPaper00058677	pmid31570707	59250	good; will get dumped
4-24-2020	WBPaper00058677	pmid31570707	59251	good; will get dumped
4-24-2020	WBPaper00058677	pmid31570707	59252	dump from PGID 59300)
4-24-2020	WBPaper00058677	pmid31570707	59253	dump from PGID 59274)
4-26-2020	WBPaper00058677	pmid31570707	59273	good; will get dumped
4-26-2020	WBPaper00058677	pmid31570707	59274	good; will get dumped
4-26-2020	WBPaper00058677	pmid31570707	59280	dump from PGID 59251)
4-26-2020	WBPaper00058677	pmid31570707	59282	dump from PGID 59245)
4-27-2020	WBPaper00058677	pmid31570707	59292	experiment; marked NO DUMP
4-27-2020	WBPaper00058677	pmid31570707	59293	dump from PGID 59251)
4-27-2020	WBPaper00058677	pmid31570707	59294	dump from PGID 59273)
4-27-2020	WBPaper00058677	pmid31570707	59296	dump from PGID 59245)

Appendix 6. Prompts for student survey following the three annotation assignments.

Students were asked to answer the following questions using an anonymous Qualtrics survey. No student identifiers were collected.

Q1. Compared with experiences in other Biology courses, did annotating data in Wormbase help you distinguish between genotype and phenotype? In what way(s)?

Q2. Compared with experiences in other Biology courses, did annotating data in Wormbase help you distinguish between control and experimental conditions? In what way(s)?

Q3. Compared with experiences in other Biology courses, did annotating data in Wormbase help you identify different alleles? In what way(s)?

Q4. Compared with experiences in other Biology courses, did annotating data in Wormbase help you distinguish between transgenic organisms and genetic mutants? In what way(s)?

Q5. Is there anything else that the annotating data in Wormbase helped you learn? Anything that became more difficult for you?

Appendix 7. Code book for analysis of student descriptions of the activity

Examples for each code described in the codebook, Table 2.

A.1. *Making connections*
"… you are understanding the experiment that is being performed as well as concepts in genetics and C. elegans biology."
"[C]omparing the experimental conditions to the control and seeing the data that the different procedures produced helped me connect the two."

A.2. *Deeper or better understanding*
"[D]oing annotations helped me create a deeper understanding of what alleles are and how to identify them within a publication."
"When annotating, I had to determine if the genotypes were transgenic or mutants which is something I've never had to consider in any of my classes."

A.3. *Proficiency*
"I think I had a pretty good understanding of control and experimental conditions, but annotating definitely solidified my knowledge on the subject."
"I think in other courses I always assumed there was only one important allele for each gene, but this experience has helped my learning and understanding of alleles."

A.4. *Skills and Competency*
"It helped me learn to quickly span many papers searching for the "needs annotation" quote on wb [WormBase]. After that it is just comprehension and followup [sic]."
"This was a little tricky for me, but I think that over time and doing multiple annotations I was able to distinguish between [transgenics and genetic mutations]. Learning in the annotations beforehand helped me distinguish between the two.'

B.1. *Different perspective*.
"[The activity] helped me realize what I did and did not know about the paper I had just read."
"I think WormBase primarily helped me understand how to find important information rather than just reading through an article and not actually understanding what was being said."

B.2. *New use of resources*.
"I never knew where to look in the paper for controls, but after annotating data for Wormbase I was able to learn to look at supplementary data in order to learn more about the control used."
"It was very difficult at times to find papers with PMIDs and not just PMCs or DOIs. It was also difficult to find worm papers that were not already annotated and were digesible at our level of knowledge."

B.3. *Emotion*
"No, I honestly had a really hard time with this. It often felt unclear and like I was left to my own devices with this set of assignment."
"We did not touch on [differentiating alleles] much in class, but I figured it out."

C.1. *Community*
"I was able to learn more about specific nomenclature [notation] within the *C. elegans* community."
"It gives a deeper connection with the scientific community because you can be annotating a gene of interest from a paper, then when you see that author or gene of interest again you will be able to remember what the paper is talking about."

C.2. *Real-world application*
"[I]t required me to look and find and understand in papers what someones [*sic*] control and experimental conditions were and why they were like that."
"[W]orking with [J.] Rose's data/paper was especially illuminating [regarding] associated conditioning, and experimental conditions in the real world.