Supplemental Material

CBE—Life Sciences Education

Erdmann and Stains

Supplementary Materials to:

Classroom as genome: Using the tools of genomics and bioinformatics to illuminate classroom observation data

Robert M. Erdmann and Marilyne Stains

Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68588-0304

Contact: rerdmann@unl.edu,mstains2@unl.edu

Table of Contents

Selected classroom observation protocols with a real-time component

Name	Description/purpose	Citation
Classroom Observation Protocol for Undergraduate STEM (COPUS)	Allows users "to reliably characterize how faculty and students are spending their time in the classroom;" uses 25 codes applied within two minute blocks.	Smith *et al.*, 2013
Decibel Analysis for Research in Teaching (DART)	"DART analyzes the volume and variance of classroom recordings to predict the quantity of time spent on single voice (e.g., lecture), multiple voice (e.g., pair discussion), and no voice (e.g., clicker question thinking) activities."	Owens *et al.*, 2017
Flanders Interaction Analysis (FIA)	A system for characterizing verbal interactions between instructors and students; uses 10 codes applied within three second blocks.	Amidon and Flanders, 1967
Laboratory Observation Protocol for Undergraduate STEM (LOPUS)	"LOPUS captures both students' and TAs' behaviors every 2 minutes as well as initiators of verbal interactions and the nature of these verbal interactions (e.g., data analysis, explanation of concepts)."	Velasco *et al.*, 2016
Observation Protocol for Active Learning (OPAL)	OPAL is "a visual approach to integrating observational data into self-evaluation and peer review of teaching."	Frey *et al.*, 2016
Practical Observation Rubric To Assess Active Learning (PORTAAL)	PORTAAL "identifies 21 readily implemented elements that have been shown to increase student outcomes related to achievement, logic development, or other relevant learning goals with college-age students;" the observation log contains more real-time information than the output summary scores.	Eddy *et al.*, 2015
Real-time Instructor Observing Tool (RIOT)	RIOT is "a computerized real-time instructor observation tool to take data of student-instructor interactions."	West *et al.*, 2013
Real-time Professional Development Observation Protocol (RPDOT)	RPDOT is designed "to document the form and focus of faculty engagement during workshops."	Olmstead and Turpen, 2016
Teaching Dimensions Observation Protocol (TDOP)	TDOP is intended to "capture the dynamics among teaching methods (e.g., lecture, small-group discussion), use of instructional technology (e.g., clickers, PowerPoint), and students' cognitive engagement (i.e., the types of student thinking evoked by the instruction)."	Hora and Ferrare, 2013
VaNTH Observation System (VOS)	VOS is a four part system, of which two parts (the Classroom Interaction Observation (CIO) and the Student Engagement Observation (SEO)) are linked to the progression of time within the classroom.	Harris and Cox, 2003

COPUS coding

The COPUS data utilized in this essay originated from the data set described in Lund et al. (2015), where complete details of coding and collection can be found. In summary, 269 classroom observations were coded, with sampled instructors derived from two populations: attendees of the Cottrell Scholars Collaborative New Faculty Workshop (CSC NFW) for early-career chemistry faculty at research-intensive institutions, and STEM faculty from a range of disciplines working at a single Midwestern research-intensive university. Between the two groups, 73 instructors were observed. All 25 COPUS codes were coded for each of the 269 videos. A test of interrater reliability provided an average Cohen's kappa score of 0.852 for the instructor codes and 0.908 for the student codes. While the analysis of Lund et al. (2015) relied on the percentage of each observation'stwo-minute time blocks that were coded for each code, our current analyses utilize the original, uncompressed coding output, with information for each code's presence or absence at each two minute interval.

FIA coding

The Flanders Interaction Analysis(FIA; Amidon and Flanders, 1967) observation protocol was applied to a subset of the videos that had been coded for COPUS within the data set described above. Unlike COPUS, FIA only allows for the coding of a single code within any given three second time block. In cases where code transitions did not correspond exactly with time block boundaries, the code occupying the majority of the time block was recognized. However, if a code would be ignored under this rule but would be essential to understanding the sequencing of verbal interactions (for example, a quick "very good" response given by an instructor in response to a student's answer, taking less than half of the time block), that code was given a single three second time block before coding the following interaction type.

Five videos were randomly selected from the 18 videos categorized as "Socratic (at board)" in Lund et al. (2015). "Socratic (at board)" was specifically selected due to the fact that a relative abundance of questions might be expected within the observations, which fit well with the intent of the related case study. One of the five videos, composed of 977 three second time blocks in total, was tested for intrarater reliability, with a resulting Fleiss' kappa value of 0.929.

Transforming COPUS and FIA data into bed file form

The method for converting code information into a bed file format can vary depending on how the source data is formatted. Our FIA coding spreadsheets were formatted in the style displayed below:

0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60
0	5	5	5	5	5	10	10	5	5	5	5	5	5	5	5	5	5	5	5	5
1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	2	2	2
2	5	5	2	5	5	5	4	2	5	5	5	5	5	5	5	5	5	5	5	5
3	5	5	5	5	5	2	5	5	5	5	5	5	5	5	5	5	5	5	5	5
4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5
5	5	5	5	5	5	5	10	10	10	10	10	10	10	10	10	10	10	10	5	5

…with additional rows stretching below. There are no text labels here, as we would soon convert the spreadsheet file into a tab delimited text file, and the numbers were the important information needed.

The top row indicates the time within a single minute in seconds, with the number at the top of a column indicating the end point for a coded time span. Meanwhile, the left-most column indicates the minute within the observation, starting with the zero of the second row. As an example, the "4" found in the fourth row and the eighth column corresponds to an FIA code of an instructor posing a non-rhetorical question to students within the span from 2 minutes 18 seconds to 2 minutes 21 seconds into the recording.

Once this information was saved as a tab delimited text file, we utilized the statistical program R (R Core Team, 2016) in order to automate much of the remaining conversion process. Again, our approach was specific to the format seen above, and would need to be modified for other arrangements. This is an annotated example of a segment of the code that could be used to perform this transformation:

```
bedchr59 <- read.table("FIAgrid_32jJ4av1_01_test.txt", sep =
"\t", header = FALSE)
```

Above, we are reading our text file into R so that we can transform it. "bedchr59" refers to the fact that this was the 59th observation within our dataset.

```
bedcolumn2 <- numeric()
bedcolumn3 <- numeric()
```

Above, we are creating a pair of empty vectors that will be filled with the start and end times for each segment coded as FIA code 2 (instructor praise or encouragement).

```
# 59minute0 ----------------------------------------------------
------------

#ONEROWBLOCK - minute 0

#1
chr59startat0 <- bedchr59[2,2]
if (chr59startat0 == "2") {
  bedcolumn2 <- c(bedcolumn2,(bedchr59[1,1]+(bedchr59[2,1]*60)))
  bedcolumn3 <-
c(bedcolumn3,(3+bedchr59[1,1]+(bedchr59[2,1]*60)))
}
```

#Above, we examine the intersection of the second row and second column, and if that position is filled with a "2," we insert the start time and the end time in seconds in each of the respective vectors.

```
#2
chr59startat0 <- bedchr59[2,3]
if (chr59startat0 == "1") {
  bedcolumn2 <- c(bedcolumn2,(bedchr59[1,2]+(bedchr59[2,1]*60)))
  bedcolumn3 <-
c(bedcolumn3,(3+bedchr59[1,2]+(bedchr59[2,1]*60)))
}
```

…

The same process continues at the next position over. We can repeat this process across each row and for each column (not shown here) until we've checked the entire document and extracted all of the encoded 2's.

…

```
if (length(bedcolumn2) != 0) {
  chr197dataframe <- data.frame(chr = c("chr197"), start =
bedcolumn2, end = bedcolumn3)
}
```

Here, we see if there is anything in the start position vector (it is common to have certain codes never show up in a classroom observation) for the 197th observation, and if there is something there, we create a data frame that puts together the three columns that define the basic bed file structure.

```
generaldataframe<-
data.frame(chr=double(),start=factor(),end=factor())

if (exists("chr59dataframe") != FALSE) {
generaldataframe<- rbind(generaldataframe,chr59dataframe)
}
```

```
if (exists("chr197dataframe") != FALSE) {
generaldataframe<- rbind(generaldataframe,chr197dataframe)
}
```

Here, we work tocombine the data frames from multiple observations into a single, unified data frame.

```
completedataframe<- generaldataframe
write.table(completedataframe, file =
"FIAbedfile_2_unmerged.bed", sep = "\t", row.names = FALSE)
```

Finally, we export our data frame into a bed file.
Astute readers will notice that the export file name includes “unmerged.” This derives from the fact that even if the same code occurs in multiple adjacent time slots, the bed file will contain separate entries for each. Sometimes this can be useful, but in general, when things occur in a block we would like our bed file to reflect that.

Our sample bed unmerged bed file from above begins:

"chr"	"start"	"end"
"chr59"	111	114
"chr59"	114	117
"chr59"	117	120
"chr59"	126	129
"chr59"	141	144

The first three entries are all back-to-back. If we would like to merge them, we can use the “merge” functionality of the bedtools suite (Quinlan & Hall, 2010), after using a text editor find and replace tool to remove the quotation marks around “chr59” and the top row of headers (hence, the input file labeled “unmerged2”):

```
bedtools merge -i FIAbedfile_2_unmerged2.bed -c 4 -o max >
FIAbedfile_2_merged.bed
```

This output bed file is what we would actually use in our downstream analyses:

Chr59	111	120	1
Chr59	126	129	1
Chr59	141	144	1
Chr59	195	198	1
Chr59	471	477	1

A column of ones was added to denote that these positions “have” the code – these may or may not be necessary, depending on the situation (see “Methodology for Implementation” section in the main essay for related discussion).

This is a general overview that shows just one possible approach – a variety of scripts can be written in R or in other programming languages to accomplish this same task.

Performing ends analysis with genomic scripts

The scripts used to perform ends analysis are derived from scripts available at https://github.com/clp90/Genome_Biology_2017 (Picard & Gehring, 2017). Scripts of interest include:

Primary script:
ends_analysis.sh

Helper scripts:
ends_analysis_get_intervals.py
ends_analysis_make_matrix.py
ends_analysis_make_plot.R
ends_analysis_process_intersect.py

For the purposes of our essay, modifications were made to the scripts in order to adjust for differences in the data. For instance, the ends_analysis_make_plot.R script was adjusted to change the x-axis labels (seconds are more appropriate than kilobases in our case), as well as the colors. Further modifications to output plots were made using Adobe Illustrator – for instance, displaying only the "ending" half of the ends analysis output for the three parts of Figure 3.

Interpreting ends analysis plots

When reading an ends analysis plot, it is important to connect the way in which the lines of data are drawn with the structure of the source data. The bends in each line, which correspond to the actual input data values, are found at the midpoint of a timespan for a code. For instance, in Figure 2, the first span of time following the start of the clicker question code runs from 0 seconds to +120 seconds, while the lines bend at +60 seconds. The value of the line at +60 seconds represents the entire 0 to 120 second span. Therefore, the points along the line that fall between two data points, while showing the connection between the surrounding values, are not meaningful in and of themselves (i.e. the value at the 0 second point is merely the midpoint between the values at -60 seconds and +60 seconds, rather than reflecting an independent data reading). The same idea holds for Figure 3, with data plotted every 3 seconds and bends in the line occurring at the 1.5 second midpoint of each time span.

Construction of Figure 4

In order to calculate relative distance distributions, the Bedtools command “reldist” was used.For example:

Command

```
bedtoolsreldist -a FIAbedfile_4_merged_no97.bed -b
COPUSbedfile_PQ_merged_5chrsforFIAcomp.txt
```

Output

reldist	count	total	fraction
0.00	17	203	0.084
0.01	7	203	0.034
0.02	1	203	0.005
0.03	2	203	0.010
0.04	5	203	0.025
0.05	4	203	0.020
0.06	5	203	0.025
0.07	6	203	0.030
0.08	4	203	0.020
0.09	6	203	0.030
0.10	10	203	0.049
0.11	6	203	0.030
0.12	3	203	0.015
0.13	6	203	0.030
0.14	4	203	0.020
0.15	4	203	0.020
0.16	3	203	0.015
0.17	5	203	0.025
0.18	7	203	0.034
0.19	8	203	0.039
0.20	3	203	0.015
0.21	7	203	0.034
0.22	3	203	0.015
0.23	3	203	0.015
0.24	5	203	0.025
0.25	6	203	0.030
0.26	5	203	0.025
0.27	4	203	0.020
0.28	4	203	0.020
0.29	2	203	0.010
0.30	5	203	0.025
0.31	5	203	0.025
0.32	2	203	0.010
0.33	2	203	0.010
0.34	3	203	0.015
0.35	2	203	0.010
0.36	1	203	0.005
0.37	3	203	0.015
0.38	1	203	0.005
0.39	2	203	0.010
0.40	1	203	0.005

```
0.41     3    203  0.015
0.42     2    203  0.010
0.43     1    203  0.005
0.44     6    203  0.030
0.45     2    203  0.010
0.46     2    203  0.010
0.47     1    203  0.005
0.48     3    203  0.015
0.49  1     2030.005
```

This was repeated for each of the 24 combinations displayed in Figure 4. The fraction values from the command output, along with the relative distance values, were input into the graphing program Graphpad Prism 7. Linear regression was performed for each set of data, using the least squares method. The results of this linear regression analysis are the source for the lines displayed in Figure 4.

Using Circos to visualize classroom observations

Circos is a software tool originally developed for displaying genomic information in a circular motif (Krzywinski*et al.*, 2009). Extensive documentation is available at circos.ca. Below we provide the configuration file used to create Figure 5 in the essay:

```
<colors>
<<include colors.conf>>
</colors>

<fonts>
<<include fonts.conf>>
</fonts>

<ideogram>

<spacing>

default = 0.0020r
break   = 0.0050r
<pairwise Chr197 Chr59>
spacing = 16r
</pairwise>

</spacing>

radius           = 0.88r
thickness        = 70p
fill             = yes
fill_color       = black
stroke_thickness = 2
stroke_color     = black
show_label       = yes
label_font       = default
label_radius     = 1.07r
label_with_tag   = yes
label_size       = 20
label_parallel   = yes
label_case       = upper


label_format     = eval(sprintf(var(label)))

show_bands            = yes
fill_bands            = yes
band_stroke_thickness = 2
band_stroke_color     = white
band_transparency     = 4
```

```
</ideogram>


show_ticks          = yes
show_tick_labels    = yes

<ticks>
skip_first_label = no
skip_last_label  = no
radius           = dims(ideogram,radius_outer)
tick_separation  = 2p
label_separation = 5p
multiplier       = 1
color            = black
thickness        = 4p
size             = 20p

<tick>
spacing        = 1u
show_label     = no
thickness      = 2p
color          = dgrey
</tick>

<tick>
spacing        = 2u
show_label     = yes
label_size     = 40p
label_offset   = 5p
format         = %d
grid           = yes
grid_color     = dgrey
grid_thickness = 1p
grid_start     = 0.5r
grid_end       = 0.999r
</tick>

</ticks>

<image>
angle_offset* = -90
<<include image.conf>>
</image>

karyotype = COPUS269karyotype.txt

chromosomes_units = 600
chromosomes_display_default = no
chromosomes = "Chr156;Chr104;Chr107;Chr59;Chr197"

<plots>
```

```
type = heatmap
color = black,paired-12-qual
stroke_thickness = 0.00
stroke_color = black

#
# Instructor plots
#

<plot> #1A
file = 269karyotype_inbedforbackground.bed
r1  = 0.96r
r0  = 0.93r
color = vlpurple
</plot>

<plot> #1B
file = FIAbedfile_2_merged.bed
r1  = 0.96r
r0  = 0.93r
color = vdpurple
</plot>

<plot> #2A
file = 269karyotype_inbedforbackground.bed
r1  = 0.93r
r0  = 0.90r
color = vlred
</plot>

<plot> #2B
file = FIAbedfile_3_merged.bed
r1  = 0.93r
r0  = 0.90r
color = vdred
</plot>

<plot> #3A
file = 269karyotype_inbedforbackground.bed
r1  = 0.90r
r0  = 0.87r
color = vlblue
</plot>

<plot> #3B
file = FIAbedfile_4_merged.bed
r1  = 0.90r
r0  = 0.87r
color = vdblue
</plot>

<plot> #4A
file = 269karyotype_inbedforbackground.bed
r1  = 0.87r
```

```
r0  = 0.84r
color = vlorange
</plot>

<plot> #4B
file = FIAbedfile_5_merged.bed
r1  = 0.87r
r0  = 0.84r
color = vdorange
</plot>

<plot> #5A
file = 269karyotype_inbedforbackground.bed
r1  = 0.84r
r0  = 0.81r
color = vlgreen
</plot>

<plot> #5B
file = FIAbedfile_7_merged.bed
r1  = 0.84r
r0  = 0.81r
color = vdgreen
</plot>

<plot> #6A
file = 269karyotype_inbedforbackground.bed
r1  = 0.81r
r0  = 0.78r
color = vlpurple
</plot>

<plot> #6B
file = FIAbedfile_8_merged.bed
r1  = 0.81r
r0  = 0.78r
color = vdpurple
</plot>

<plot> #7A
file = 269karyotype_inbedforbackground.bed
r1  = 0.78r
r0  = 0.75r
color = vlred
</plot>

<plot> #7B
file = FIAbedfile_9_merged.bed
r1  = 0.78r
r0  = 0.75r
color = vdred
</plot>

<plot> #8A
```

```
file = 269karyotype_inbedforbackground.bed
r1  = 0.75r
r0  = 0.72r
color = vlblue
</plot>

<plot> #8B
file = FIAbedfile_10_merged.bed
r1  = 0.75r
r0  = 0.72r
color = vdblue
</plot>

#

<plot> #9A
file = 269karyotype_inbedforbackground.bed
r1  = 0.67r
r0  = 0.64r
color = vlpurple
</plot>


<plot> #9B
file = COPUSbedfile_L_merged.bed
r1  = 0.67r
r0  = 0.64r
color = vdpurple
</plot>

<plot> #10A
file = 269karyotype_inbedforbackground.bed
r1  = 0.64r
r0  = 0.61r
color = vlred
</plot>

<plot> #10B
file = COPUSbedfile_Ind_merged.bed
r1  = 0.64r
r0  = 0.61r
color = vdred
</plot>

<plot> #11A
file = 269karyotype_inbedforbackground.bed
r1  = 0.61r
r0  = 0.58r
color = vlblue
</plot>

<plot> #11B
file = COPUSbedfile_SAnQ_merged.bed
r1  = 0.61r
```

```
r0  = 0.58r
color = vdblue
</plot>

<plot> #12A
file = 269karyotype_inbedforbackground.bed
r1  = 0.58r
r0  = 0.55r
color = vlorange
</plot>

<plot> #12B
file = COPUSbedfile_SQ_merged.bed
r1  = 0.58r
r0  = 0.55r
color = vdorange
</plot>


<plot> #13A
file = 269karyotype_inbedforbackground.bed
r1  = 0.55r
r0  = 0.52r
color = vlgreen
</plot>

<plot> #13B
file = COPUSbedfile_SW_merged.bed
r1  = 0.55r
r0  = 0.52r
color = vdgreen
</plot>

<plot> #14A
file = 269karyotype_inbedforbackground.bed
r1  = 0.52r
r0  = 0.49r
color = vlpurple
</plot>

<plot> #14B
file = COPUSbedfile_Lec_merged.bed
r1  = 0.52r
r0  = 0.49r
color = vdpurple
</plot>

<plot> #15A
file = 269karyotype_inbedforbackground.bed
r1  = 0.49r
r0  = 0.46r
color = vlred
</plot>
```

```
<plot> #15B
file = COPUSbedfile_RtW_merged.bed
r1  = 0.49r
r0  = 0.46r
color = vdred
</plot>

<plot> #16A
file = 269karyotype_inbedforbackground.bed
r1  = 0.46r
r0  = 0.43r
color = vlblue
</plot>

<plot> #16B
file = COPUSbedfile_PQ_merged.bed
r1  = 0.46r
r0  = 0.43r
color = vdblue
</plot>

<plot> #17A
file = 269karyotype_inbedforbackground.bed
r1  = 0.43r
r0  = 0.40r
color = vlorange
</plot>

<plot> #17B
file = COPUSbedfile_IAnQ_merged.bed
r1  = 0.43r
r0  = 0.40r
color = vdorange
</plot>

<plot> #18A
file = 269karyotype_inbedforbackground.bed
r1  = 0.40r
r0  = 0.37r
color = vlgreen
</plot>

<plot> #18B
file = COPUSbedfile_Adm_merged.bed
r1  = 0.40r
r0  = 0.37r
color = vdgreen
</plot>

<plot> #19A
file = 269karyotype_inbedforbackground.bed
r1  = 0.37r
r0  = 0.34r
color = vlpurple
```

```
</plot>

<plot> #19B
file = COPUSbedfile_MG_merged.bed
r1  = 0.37r
r0  = 0.34r
color = vdpurple
</plot>

<plot> #20A
file = 269karyotype_inbedforbackground.bed
r1  = 0.34r
r0  = 0.31r
color = vlred
</plot>

<plot> #20B
file = COPUSbedfile_1o1_merged.bed
r1  = 0.34r
r0  = 0.31r
color = vdred
</plot>

<plot> #21A
file = 269karyotype_inbedforbackground.bed
r1  = 0.31r
r0  = 0.28r
color = vlblue
</plot>

<plot> #21B
file = COPUSbedfile_DV_merged.bed
r1  = 0.31r
r0  = 0.28r
color = vdblue
</plot>

<plot> #22A
file = 269karyotype_inbedforbackground.bed
r1  = 0.28r
r0  = 0.25r
color = vlorange
</plot>

<plot> #22B
file = COPUSbedfile_IW_merged.bed
r1  = 0.28r
r0  = 0.25r
color = vdorange
</plot>

<plot> #23A
file = 269karyotype_inbedforbackground.bed
r1  = 0.25r
```

```
r0  = 0.22r
color = vlgreen
</plot>

<plot> #23B
file = COPUSbedfile_IO_merged.bed
r1  = 0.25r
r0  = 0.22r
color = vdgreen
</plot>

</plots>
<<include housekeeping.conf>>
```

Supplemental Material References

Amidon, E.J., & Flanders, N.A. (1967). *The Role of the Teacher in the Classroom*. Minneapolis, Minnesota: Association for Productive Teaching.

Eddy, S.L., Converse, M., &Wenderoth, M. (2015). PORTAAL: A Classroom Observation Tool Assessing Evidence-Based Teaching Practices for Active Learning in Large Science, Technology, Engineering, and Mathematics Classes. *CBE-Life Sciences Education*, *14*(2), ar23. https://doi.org/10.1187/cbe.14-06-0095

Harris, A.H., & Cox, M.F. (2003). Developing an Observation System to Capture Instructional Differences in Engineering Classrooms. *Journal of Engineering Education*, *92*(4), 329–336. https://doi.org/10.1002/j.2168-9830.2003.tb00777.x

Hora, M.T., &Ferrare, J.J. (2013). Instructional Systems of Practice: A Multidimensional Analysis of Math and Science Undergraduate Course Planning and Classroom Teaching. *Journal of the Learning Sciences*, *22*(2), 212–257. https://doi.org/10.1080/10508406.2012.729767

Krzywinski, M., Schein, J., Birol, İ., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., &Marra, M.A. (2009). Circos: An information aesthetic for comparative genomics. *Genome Research*, *19*(9), 1639–1645. https://doi.org/10.1101/gr.092759.109

Lund, T.J., Pilarz, M., Velasco, J.B., Chakraverty, D., Rosploch, K., Undersander, M., & Stains, M. (2015). The Best of Both Worlds: Building on the COPUS and RTOP Observation Protocols to Easily and Reliably Measure Various Levels of Reformed Instructional Practice. *CBE-Life Sciences Education, 14*(2), ar18. https://doi.org/10.1187/cbe.14-10-0168

Owens, M.T., Seidel, S.B., Wong, M., Bejines, T.E., Lietz, S., Perez, J.R., … & Tanner, K.D. (2017). Classroom sound can be used to classify teaching practices in college science courses. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(12), 3085–3090. https://doi.org/10.1073/pnas.1618693114

Picard, C.L., & Gehring, M. (2017). Proximal methylation features associated with nonrandom changes in gene body methylation. *Genome Biology*, *18*(1), 73. https://doi.org/10.1186/s13059-017-1206-2

Quinlan, A.R., & Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team (2016). R: A language and environment for statistical computing, Version 3.3.2. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Velasco, J.B., Knedeisen, A., Xue, D., Vickrey, T.L., Abebe, M., & Stains, M. (2016). Characterizing Instructional Practices in the Laboratory: The Laboratory Observation Protocol for Undergraduate STEM. *Journal of Chemical Education*, *93*(7), 1191–1203. https://doi.org/10.1021/acs.jchemed.6b00062