

BST281

Lab Session 5

Announcement:

- Problems2 has been posted. Deadline is March 10th (before spring break).
 - Reminder: March 1, 6, 8 are for journal club presentations.
 - Today's lab focus on
 - Get presentation group numbers & paper titles
1. Journal club guidelines
 2. Journal club tips
 3. Some python practices
 4. Group discussion (optional)

TA: Xue Zou Email: xuz943@mail.harvard.edu

Journal club groups

Group #	Members	Date	Article title
1	Ben Hong, Olivia Dai, Max Jentsch	Wed, Mar 1, 2017	No evidence for extensive horizontal gene transfer in the genome of the tardigrade <i>Hypsibius dujardini</i>
2	Molly McNamara, Michael Macarthur, Ilana Kelsey, Jason Lin	Wed, Mar 1, 2017	Bronchoalveolar Lavage Proteomics in Patients with Suspected Lung Cancer
3	Hector Perez, Taylor Sehein, Katia Chadaideh	Wed, Mar 1, 2017	Diet Dominates Host Genotype in Shaping the Murine Gut Microbiota
4	Ibrahim Diakite, Jonathan Batty	Mon, Mar 6, 2017	Meta-analysis of host response networks identifies a common core in tuberculosis
5	Jiwon Lee, Min-Jung Wang, Mutong Zhao, Qianyu Yuan, You Wu	Mon, Mar 6, 2017	Analysis of protein-coding genetic variation in 60,706 humans
6	Samantha Giffen, Eddie Irvine, Joanna Olivas, Greg Babunovic	Mon, Mar 6, 2017	Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets
7	Conor Foley, Isabelle Mieling, Alex Wu, Yongling Ye, Nancy Zhang	Mon, Mar 6, 2017	A network of conserved synthetic lethal interactions for exploration of precision cancer therapy
8	Alberta Wang, Kathy Lee, Amit Joshi	Wed, Mar 8, 2017	Association of host genome with intestinal microbial composition in a large healthy cohort
9	Qiu hao Zhang, Xinan Wang, Fangyuan Hong, Yuxi Liu	Wed, Mar 8, 2017	Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer
10	Kaylyn Williamson, Andrew Goldfarb, Soun Lee, Marsha Wibowo	Wed, Mar 8, 2017	Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin
11	Ibrahim Kurt, Hazreen Abdul Majid	Wed, Mar 8, 2017	HiChIP: efficient and sensitive analysis of protein-directed genome architecture

Journal club presentation guideline

1. Article selection (done)
2. Article presentation
 - **Study Background:** This section provides your audience with the necessary information and context for a thoughtful and critical evaluation of the article's **significance**. The goals are
 - 1) to describe the rationale for and clinical relevance of the study question, and
 - 2) to highlight the preclinical and clinical research that led to the current trial.Review the papers referenced in the study's "Background" section as well as previous work by the study's authors. It also may be helpful to discuss data supporting the current standard of care against which the study intervention is being measured.

Journal club presentation guideline

1. Article selection (done)
2. Article presentation
 - **Study Methodology and Results:** Clearly describe the **study population, including inclusion/exclusion criteria**. A diagrammatic schema is easy to construct using PowerPoint software and will help to clearly illustrate treatment arms in complex trials. Explain the statistical methods, obtaining assistance from a statistician if needed. Take this opportunity to verbally and graphically highlight key results from the study, with plans to expand on their significance later in your presentation.
 - **Author's Discussion:** Present the authors' **conclusions and their perspective** on the study results, including **explanations** of inconsistent or unexpected results. Consider whether the conclusions drawn are supported by the data presented.

Journal club presentation guideline

3. Article critique

This component of your presentation will define the success of your journal club.

While a comprehensive discussion of scientific literature appraisal is beyond the scope of this discussion, several helpful tips warrant mention here. In assessing the validity of the study, it is important to assess for **potential sources of bias**, including the funding sources and authors' affiliations. It is also helpful to look for accompanying editorial commentary, which can provide a unique perspective on the article and highlight controversial issues. You should plan to discuss the trade-offs between potential benefits of the study intervention versus potential risks and the cost. By utilizing the concept of number needed to treat (NNT), one can assess the true impact of the study intervention on clinical practice. Furthermore, by incorporating the incidence rates of clinically significant toxicities with the financial costs into the NNT, you can generate a rather sophisticated analysis of the study's impact on practice.

Journal club presentation guideline

4. Conclusions, implications, and future directions:

Restate the authors' take-home message followed by your own **interpretation** of the study. Provide a **personal perspective**, detailing why you find this paper interesting or important. Then, look forward and use this opportunity to "think outside the box." Do you envision these study results changing the landscape of clinical practice or redirecting research in this field? If so, how? In articles about therapy, future directions may include moving the therapy up to first-line setting, assessing the drug in combination regimens or other disease states, or developing same-class novel compounds in the pipeline. Searching for related clinical trials on the NIH Web site⁶ can prove helpful, as can consultation with an expert in this field.

Journal club presentation tips

❖ Good journal club discussions are integral to the educational experience of hematology trainees. Following the above approach, while utilizing the resources available, will lay the groundwork for an outstanding presentation.

- **WEB BASED REFERENCES**

- www.acpjc.org
- hiru.mcmaster.ca/more/InclusionCriteria.htm
- hiru.mcmaster.ca/more/RatingFormSample.htm
- www.cche.net/main.asp
- www.hematology.org/Trainees
- www.cancer.gov/clinicaltrials

Journal club presentation tips

- I. Preparation
 - A. Choose an interesting paper with presentable data
 - 1. Only one paper
 - 2. Research paper, not a review
 - B. Make sure you understand the background
 - 1. Look up references
 - C. Make sure you understand the experimental details
 - 1. Look up references
 - D. Make sure visuals are easy to see - enlarge the figures and tables.
 - 1. Preview your slides in a real auditorium from the back of the room. If you can't read them easily, change them.
 - E. Beware of Power Point
 - 1. Concentrate on content, not gloss.
 - 2. Use a minimum of slides.
 - 3. Avoid low contrast colors. White backgrounds work best.
 - 4. Avoid excessive text and avoid reading text to the audience. It is better to use minimal bullet points where necessary and elaborate by speaking directly to the audience.
 - 5. Avoid excessive "cute" animations or graphics. It's distracting.
 - 6. In general, concentrate on the audience, not on the slides. (See II-G)

Journal club presentation tips

II. Presentation

- A. Practice your presentation with a timer. 50 minutes for a one hour slot is optimum.
- B. No one understands anything. Start from the beginning, keep it as simple as possible, remember the big picture. It is far better to insult an arrogant know-it-all than to leave someone behind.
- C. At least half of your responsibility is to tell the audience what you are doing during the presentation, to guide them through the presentation.
- D. Present an outline of the talk.
 - 1. Better yet, put it on the board, or use it repeatedly during the talk to track your progress.
- E. Presentation drill
 - 1. Tell 'em what you're gonna tell 'em (outline)
 - 2. Tell 'em (body of the presentation)
 - 3. Tell 'em what you told them (concluding summary)
- F. Figure drill
 - 1. Hypothesis
 - 2. Experimental design & details - use written outlines
 - 3. Experimental data
 - 4. Interpretation
 - 5. Conclusions
 - 6. New Hypothesis - how it leads to the next experiment

Journal club presentation tips

7. CAUTION

- a. Don't show a figure unless you intend to actually address specific elements of the figure in detail
- b. If you are not going to discuss all of the data in a figure, tell the audience this and direct their attention to the things you *are* going to discuss.
- c. It takes the audience a long time to assimilate a figure. If you carefully point out all of the features of a figure as it is presented (abscissa, ordinate, symbols, columns, etc.) it helps them to understand it and it forces you to keep a reasonable pace.
- d. Make sure you understand the difference between data, interpretation and conclusions, and that you make these distinctions during the presentation.

G. Communicate with the audience

1. Make eye contact with everyone; concentrate as much as possible on the audience rather than the slides.
2. Ask for questions
3. Listen carefully when a question is asked. Often there is a lot of confusing wasted discussion that revolves around a misunderstood question. If you don't understand the question, ask for clarification.

H. Don't try to be humorous

1. Most likely you will just distract the audience and embarrass yourself
2. Usually the most humorous stuff is spontaneous or unintentional

❖ Important Reminder

- 1) Start by **enumerating the 1-2 major "claims" of the paper**. This is especially important for papers in "big" journals, which must have made a substantial finding to have been published there. At the end of the presentation, students can then evaluate whether they think the claims are justified.
- 2) **Make sure you practice your timing**. Everyone in the group must speak in the allotted ~20 minutes. It is very easy to spend too much time on introduction and then have to rush the conclusions.
- 3) Similarly, students should **leave time to answer questions from the instructors / classmates** at the end of class. Ideally all team members will participate in the Q&A, but sometimes this is not feasible given that there will only be time for a few questions.

Python programming practice problems

Warm up problems

Note: the next problems are taken from the webpage <http://www.practicepython.org>

1. List remove duplicates.

Write a function called `noDups` that takes a list and returns a new list that contains all the elements of the first list minus all the duplicates.

2. Reverse word order.

Write a function called `reverseWord` that takes long string containing multiple words. Print back to the user the same string, except with the words in backwards order.

Example: it takes "The house is yellow" and it returns "yellow is house The".

Hint:

`append()`

`set()`

`join()`

`reversed()`

Example answers for 1:

```
# Write a function that takes a list and returns a new list that contains
# all the elements of the first list minus duplicates.

# this one uses a for loop
def dedupe_v1(x):
    y = []
    for i in x:
        if i not in y:
            y.append(i)
    return y

#this one uses sets
def dedupe_v2(x):
    return list(set(x))

a = [1,2,3,4,3,2,1]
print a
print dedupe_v1(a)
print dedupe_v2(a)
```

Example answers for 2:

```
# method 1: loop through the words and insert each word at the begining of the result list
def reverse_v1(x):
    y = x.split()
    result = []
    for word in y:
        result.insert(0,word)
    return " ".join(result)

# method 2
def reverse_v2(x):
    y = x.split()
    return " ".join(y[::-1])

# method 3
def reverse_v3(x):
    y = x.split()
    return " ".join(reversed(y))

# method 4
def reverse_v4(x):
    y = x.split()
    y.reverse()
    return " ".join(y)

# test code
test1 = raw_input("Enter a sentence: ")
print reverse_v1(test1)
print reverse_v2(test1)
print reverse_v3(test1)
print reverse_v4(test1)
```

Python programming practice problems

3. File with only odd line numbers.

Write a program that reads a file and returns only the odd lines, write the results to a txt file. For example, Input file has the following text:

```
Hello world!  
Today is a great day!  
My house is blue.  
Yabba dabba doo  
Goodbye!
```

Output file has the will have:

```
Hello world!  
My house is blue.  
Goodbye!
```

Hint: to test your code you can try to make your own file, or you can use the file `HMP_trunc.txt`, then do `cat -n HMP_trunc.txt > HMP_trunc_numbers.txt` . The new file you generated contains in the first column the line numbers.

Hint: how does an even / odd number react differently when divided by 2?

(pseudo code)

`mod=num%2`

If `mod > 0`, we have odd number

Python programming practice problems

4. Read from File.

Given a file `nameslist.txt` that has a list of a bunch of names, count how many of each name there are in the file, and print out the results to the screen.

Note: the next problems are taken from Rosalind project (<http://rosalind.info/about/>)

5. Problem 3 from Rosalind

Given: A string of length at most 200 letters and four integers a , b , c and d .

Return: The slice of this string from indices a through b and c through d (with space in between), inclusively.

Sample Dataset:

```
aStr="HumptyDumptysatonaWallHumptyDumptyhadagreatfallAlltheKingshorsesandallt  
heKingsmenCouldntputHumptyDumptyinhisplaceagain."  
a=22; b=27; c=97; d=102
```

Sample Output:

```
Humpty Dumpty
```


Python programming practice problems

6. Problem 4 from Rosalind

Given: Two positive integers a and b ($a < b < 10000$).

Return: The sum of all odd integers from a through b , inclusively.

Sample Dataset:

```
100 200
```

Sample Output:

```
7500
```

Hint: you can use `a%2==1` to test if a is odd.

Python programming practice problems

7. Problem 7 from Rosalind

Given: A DNA string s .

Return: Four integers counting the respective number of times that the symbols 'A', 'C', 'G' and 'T'.

Sample Dataset:

```
AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGATTAAAAAAGAGTGTCTGATAGCAGC
```

Sample Output:

```
20 12 17 21
```

Python programming practice problems

8. Problem 8 from Rosalind

An RNA string is a string formed from the alphabet containing 'A', 'C', 'G', and 'U'. Given a DNA string t corresponding to a coding strand, its transcribed RNA string u is formed by replacing all occurrences of 'T' in t with 'U' in u .

Given: A DNA string t having length at most 1000 nt.

Return: The transcribed RNA string t .

Sample Dataset:

```
GATGGAACCTTGACTACGTAAATT
```

Sample Output:

```
GAUGGAACUUGACUACGUAAAUU
```