M1D8:
Phylogenetic analysis

3/6/15

Lab business

- Lab treat...

- Extra office hours
  - Noreen: Sunday 1-3pm (16-429b)
  - Shannon: Sunday 3-5pm (Simmons Hall 528)
- Office hours
  - Noreen: Monday and Thursday 2-4pm (16-429b)
  - Leslie: Monday and Thursday 1-2pm (16-429b)
  - Shannon: Monday and Friday 3-4pm (16-319)

Assignment reminders

- Journal club II
  - Meet in 16-336 at 1:15p (speakers at 1:00p)
- Microbiome abstract and data summary
  - Due Saturday, March 14 at 5:00p
  - Data posted on M1D8 Talk page
    (by Saturday at 12:00pm)
- Primer design memo
  - Due Thursday, March 20 at 10:00pm
  - Data posted on M1D7 Talk page

Microbiome abstract and data summary

- Read the assignment description on the wiki!

guidelines for formatting

example
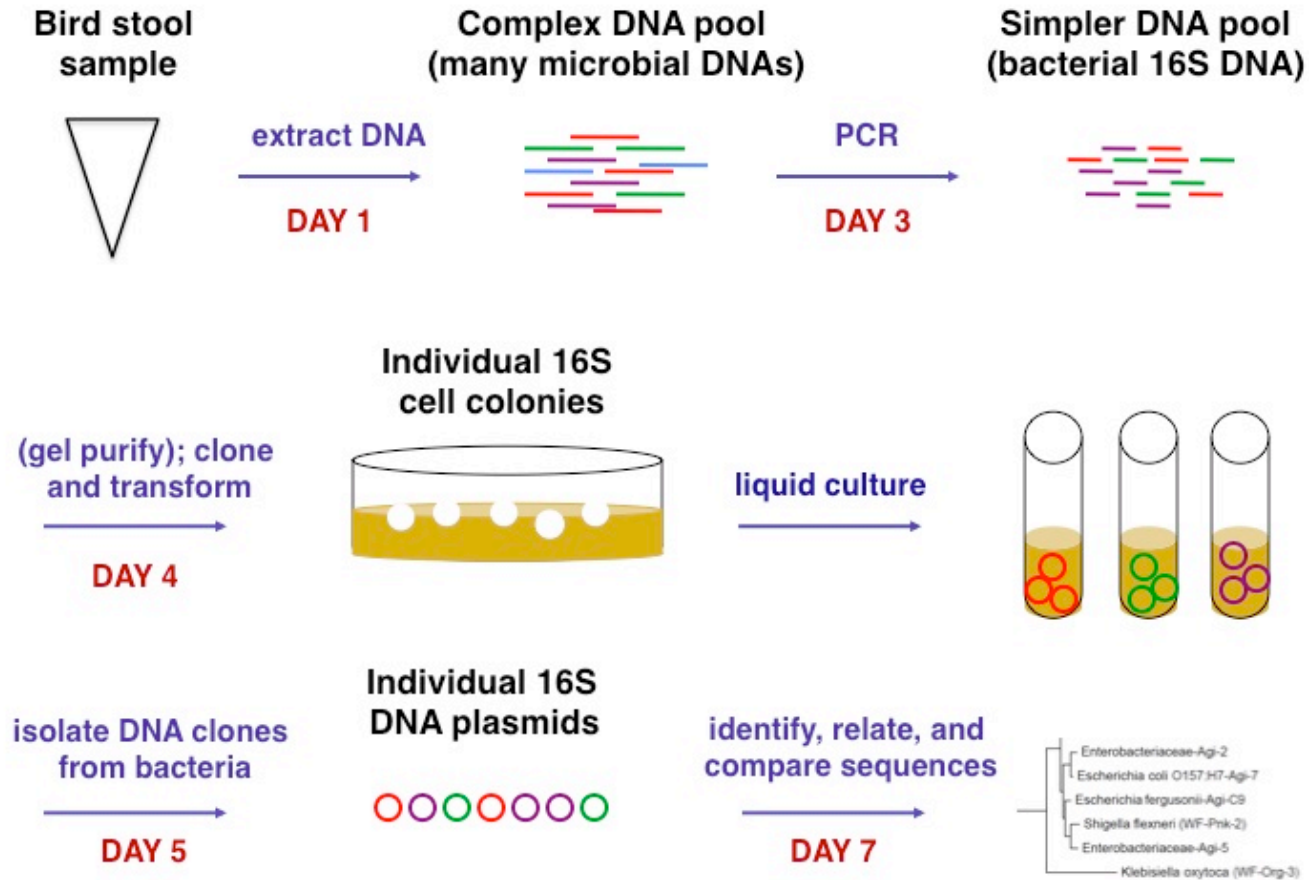
figures 1 pg.

multipaneled

up to 3 additional figures
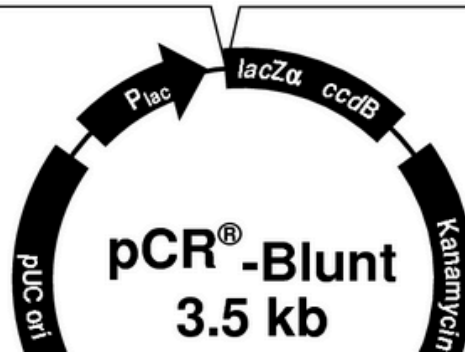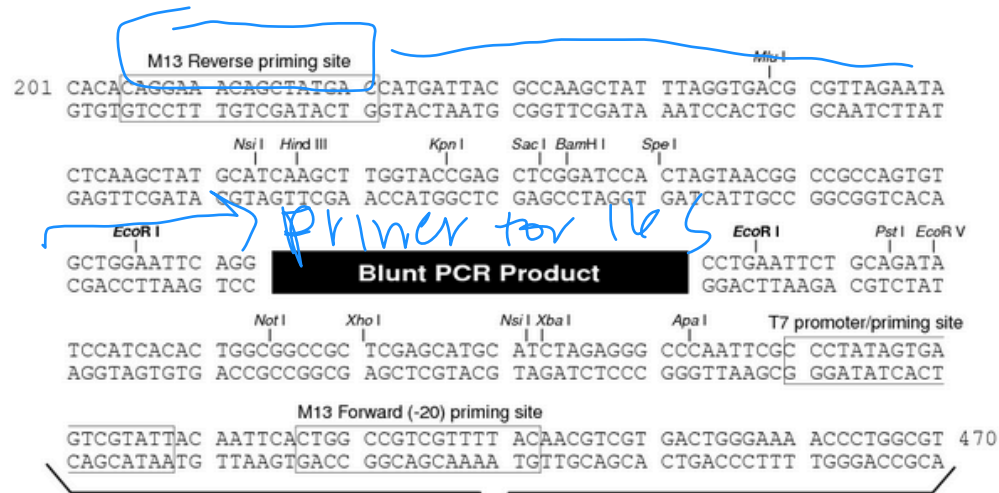
guidelines for content

NOT a checklist

Figures, captions, and results/discussion

1. What was the overall goal of these data? State concisely in introductory sentence.
2. What was your expected result? State this concisely in a bullet point.
   - What are the expected band sizes on your gel?
3. What was the result? State this concisely in one or two bullets.
   - Were bands of the expected size present? Why or why not?
4. What evidence do you have that your result is correct or incorrect?
   - What controls did you perform and did they work as expected?
5. In sum, what do these data suggest or indicate? Think about how the data were used.
6. What does this motivate you to do next?

Thank you, Shannon

Experimental overview

Let's review our cloning procedure

Thinking about your data…

- Bootstrapping

Handwritten annotations: 5, 24, 31, Strength of the clades, *random Sampling w/ replacement

```
           Site: 1  2  3  4  5  6  7  8  9  10 11 12 13 14 15 16 17 18 19 20
  Species       ............................................
 1 Pre (Chimp)   C  T  T  G  A  G  A  A  A  A  T  T  C  T  T  A  G  A  T  A
 2 Pme (Lizard)  T  C  T  A  A  A  A  G  A  T  T  A  T  A  T  A  G  A  T  A
 3 Pma (Human)   T  T  T  A  A  G  G  A  A  A  T  T  C  T  T  A  A  A  T  T
 4 Pfa (Human)   T  T  T  G  A  G  A  A  A  A  T  T  C  T  T  A  G  A  T  A
 5 Pbe (Rodent)  T  T  T  A  A  G  A  A  A  A  T  T  T  A  T  A  A  A  T  A
 6 Plo (Bird)    T  T  T  A  A  G  A  A  A  A  C  T  C  A  C  A  A  A  T  C
 7 Pfr (Monkey)  C  T  T  A  A  G  A  A  G  A  T  T  C  T  T  A  G  G  A  A
 8 Pkn (Monkey)  C  T  T  A  A  G  A  A  A  G  T  T  C  T  T  A  G  A  T  A
 9 Pcy (Monkey)  C  T  C  A  T  G  A  A  A  A  T  T  C  T  T  A  G  A  T  A
10 Pv (Human)    C  T  T  A  T  G  A  A  A  A  T  T  C  T  C  G  G  A  T  A
11 Pga(Bird)     T  T  T  A  A  G  A  A  A  A  T  T  T  T  C  A  A  A  T  C
```

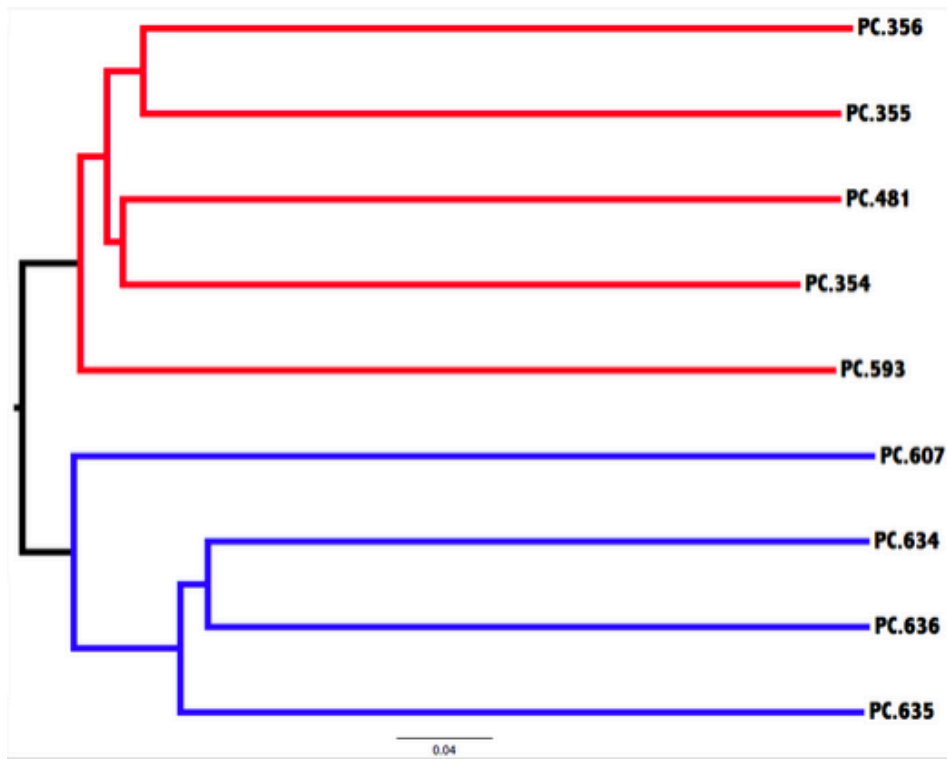Bradley Efron et al. PNAS 1996; 93:13429
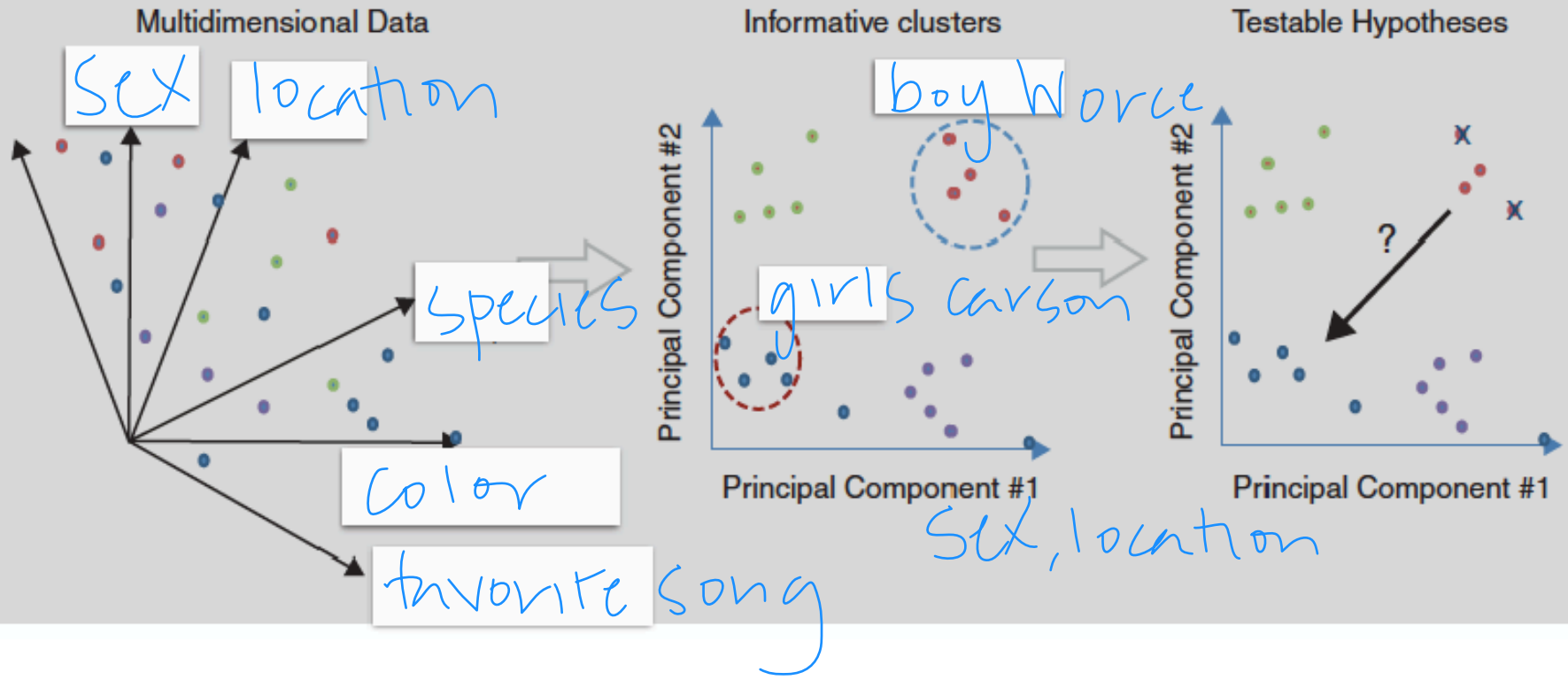
Thinking about your data…

Sample = bird

- Jackknife sample clusters

robustness / evenness

Thinking about your data…

Visualize highly dimensional data

- Principle component analysis (PCoA)

Multidimensional Data — Sex, location, species, color, favorite song

Informative clusters — boy Worce, girls Carson

Testable Hypotheses — Sex, location

Overview of Unifrac files

1. Tree (.nwk) with all of the class sequences

   Upload MEGA alignments
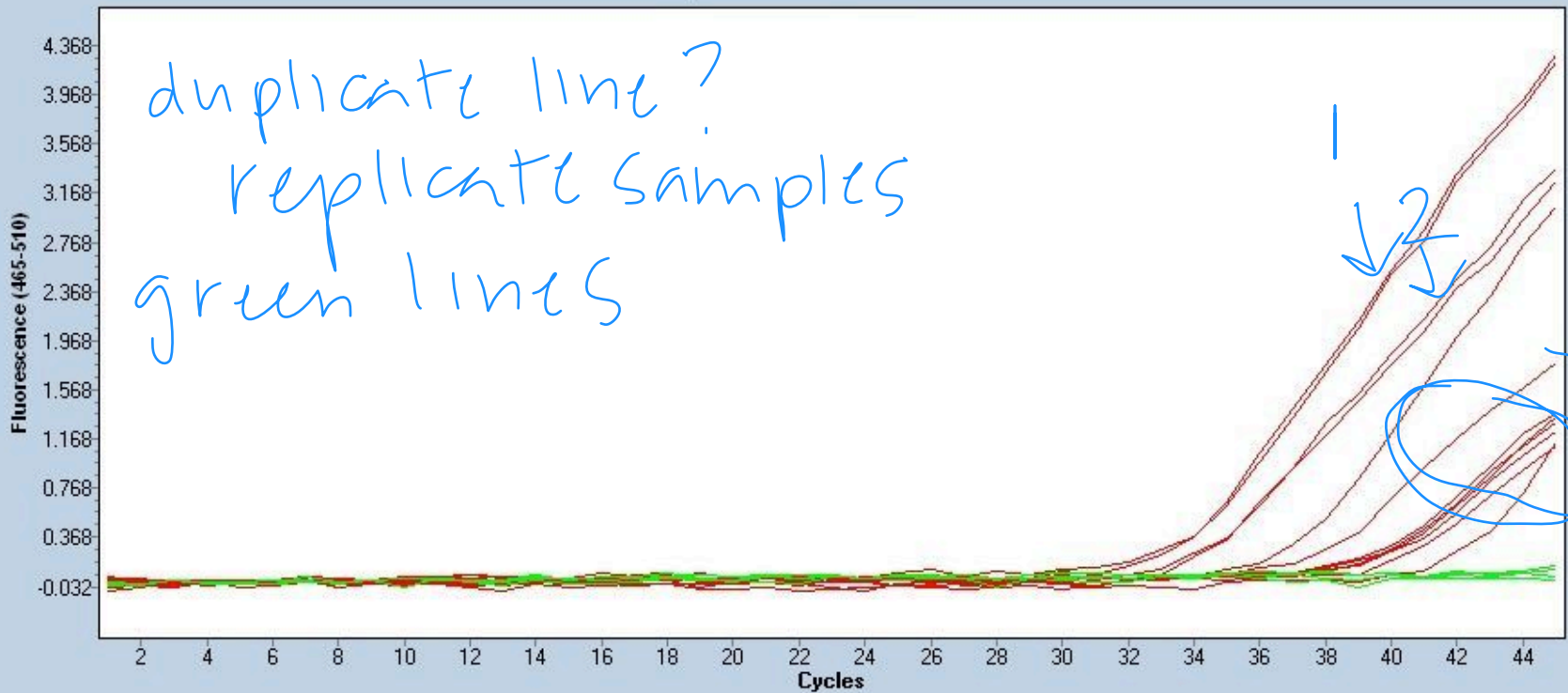
2. ID file (.txt)

   – Spaces and no spaces

3. Mapping file (.txt)

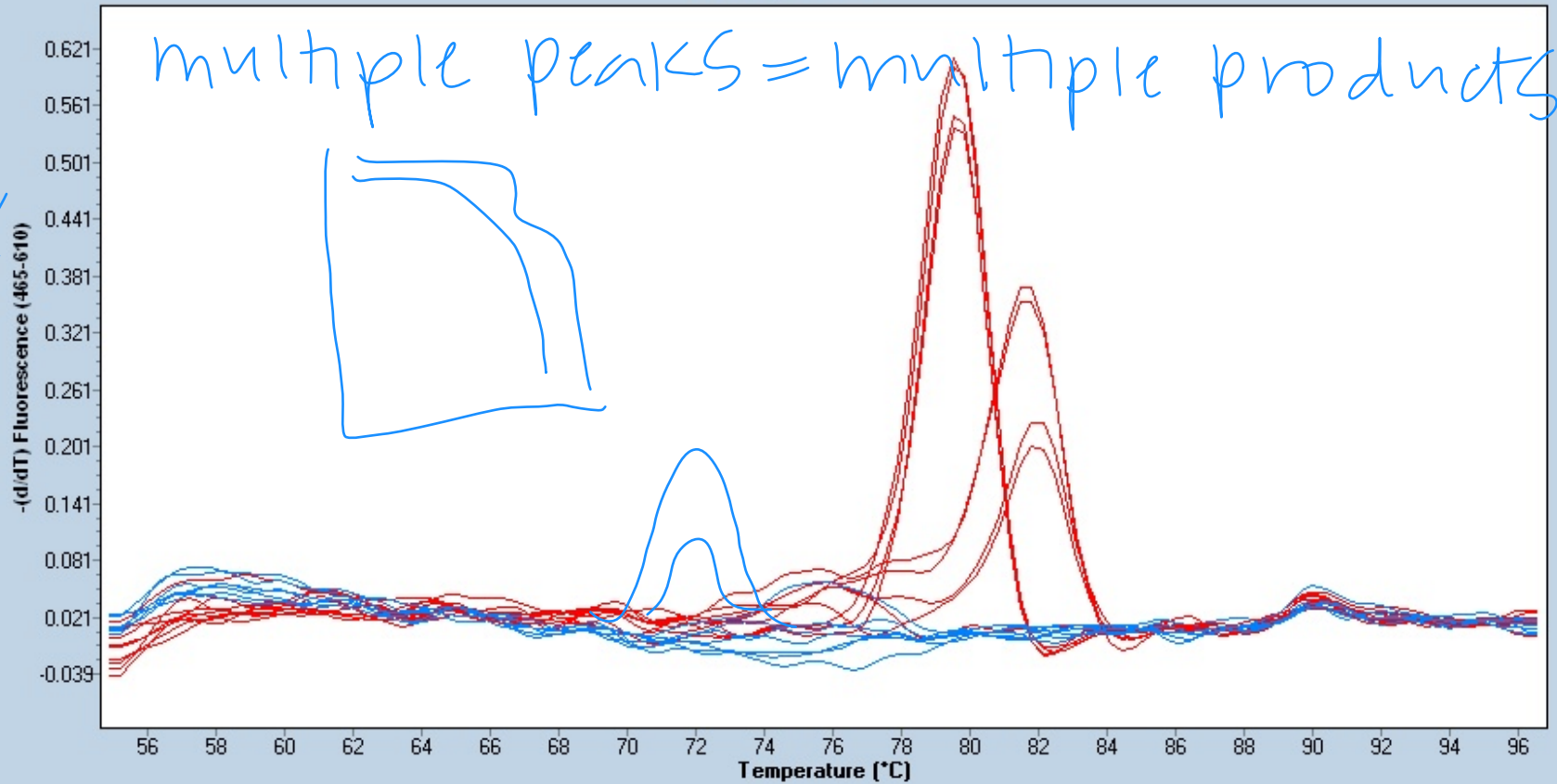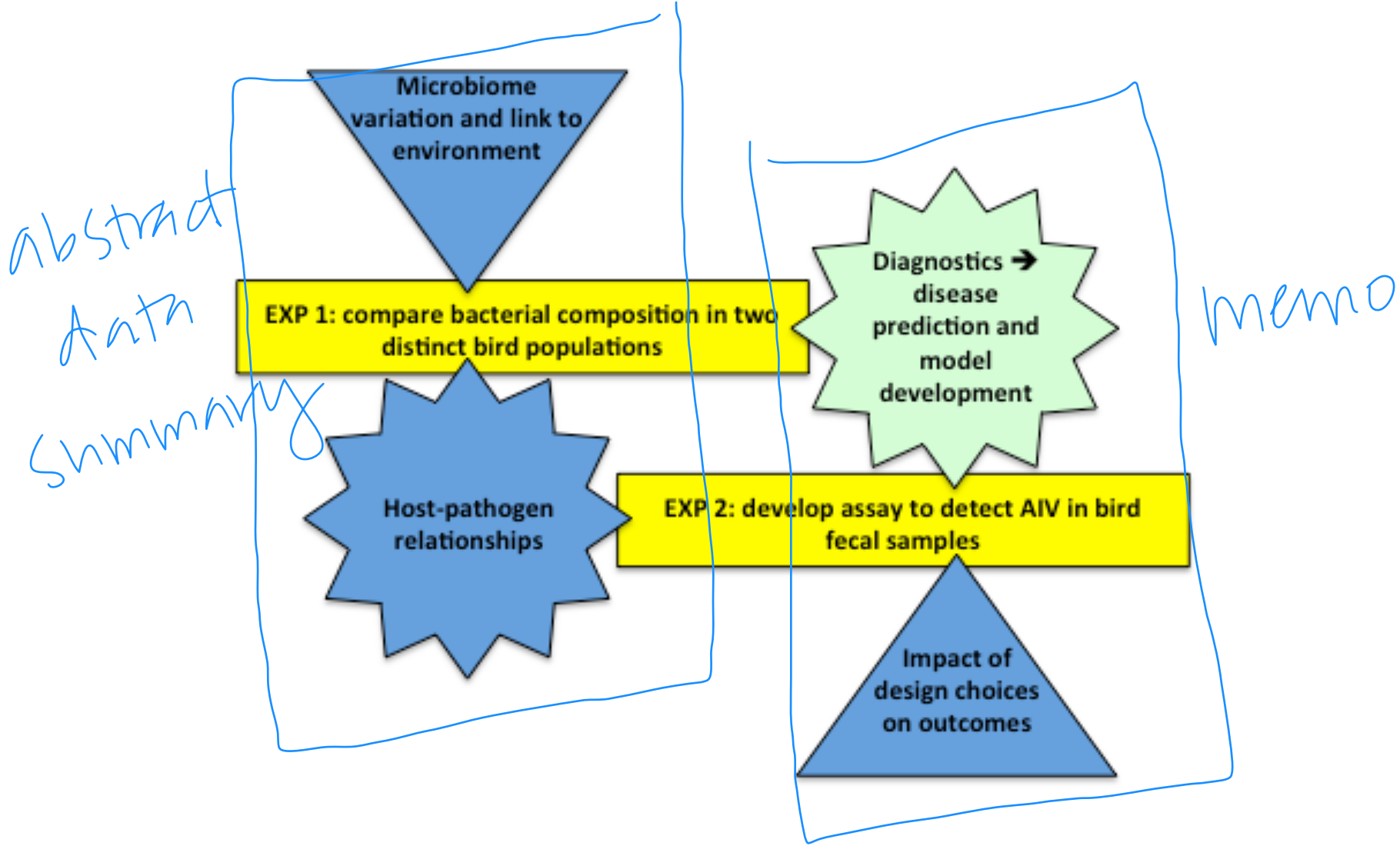   – Sample ID, sex, and location

Notes on AIV detection assay

More notes on AIV detection assay

Module 1 conceptual overview

Today

1. Align 'good' samples from Genewiz site
   – Refer to the spreadsheet

2. Analyze ~8 sequences
   – If all of your bird samples are complete, help your classmates

3. Post data on M1D8 Talk page
   – Full length sequence files (.txt)
   – MEGA alignment file (.mas)

4. All data must be posted by Saturday at 12:00p