Soojin Meeting Notes

July 11, 2023

Total = about 700 samples (many of known age)

Is aging accelerated in individuals that have experienced more infections? Demonstrate effect on lifespan. (just with this known age dataset, not future with PCR)

Options:

Clock Foundation = $160 per sample

All known age individuals, looking at $100,000

Certainly juvenile won’t need as many samples, could send far fewer

Subset but not entirely

More likely 50k

Inclination is to go that route as a first pass

Give us the raw data, and we can target certain cpg sites representative of signatures that we want

Those sites are not selected by functional means, could just be that they show a signal from epigenetic drift

Bat genome size: 2 gB, optimizing only gets us to $300 per sample

If you had a few really old, could do full genome on those older ones to get even more information 10-20 samples

* Can use to identify sites that are not included in the array
* Can also infer information about function then.

Clock foundation can help identify individuals of interest (young and old) and then follow up with the deep bisulfite sequencing

800.000 cpgs in the bat genome, so we’re actually not getting super all of them.

We aren’t a genomics lab, our goal is diagnostic tool

These 700 are ¼ of our total dataset. So we want to get all of them.

Soojin thinks it would be logical to make those steps.

Definitely be able to answer question of infection having impact on aging

DNA extracted from wing tissue. Common sample in bat biology world, and the prior clock was built on that.

**Sites that appear most representative from microarray, is there a reason to think that they might not be the top sites in an Illumina protocol?**

**YES, reason is that they are well correlated, but they aren’t perfect correlations. R2 is 0.8. there is variation. That’s because some sites that come up very well with RNAseq that aren’t coming up with microarray. And the methods have some stochastic differences. They should correspond, but we shouldn’t expect perfect correlation. But this is a general issue, so people will understand.**

We’ll go forward with sending a subset to the Clock Foundation

Ask: tradeoffs question: thoughts on proportion of samples needed to identify sites. What samples to send in?

Adult: up to 40 years in the wild

Oldest: 16 years.

Pteropus: more truncated, because of human hunting pressure. 8-10.

Juvenile: 0-1. Can be estimated pretty precisely based on capture date.

Have juvenile infection and maternally

More samples that you have around a certain age, can

?maybe put off adding many juveniles until someone is here to focus on those questions?

\*\*reach out to the clock foundation to see how many individuals they recommend

**Even distribution between the infection variables is important to answer the downstream analysis (infection and non infected)**

Look at data, see where the skews are. Are all old samples infected or not?

\*\*\*Connect Soojin and Manny – SPRI bead protocol

Power analysis: can do that for this!

**Grants and future funding**

Cara is submitting EEID grant, NSF/NIH join funded

Equivalent to an RO1

Interested in age-structured serology (money to get ages of bats, but not to dive into genomics)

Soojin is going to look into it. Soojin can help us out with building the datasets. Blood is possibly more important for epigenetic stuff.

What blood details? – Soojin is going to look into how people do field collection for molecular biology

Fresh blood in heparin tubes then process immediately

* We have a pretty good cold chain method bc we have study RNA viruses
* We take whole blood in Qiagen RNA protect tubes
* Our goal with that is transcriptomics
* If going forward there is a better preparation method, let us know and we can change.
* We take aliquot of blood, separate to serum and red blood cells

Soojin is just federal grants so far. Foundation grants possible for infectious disease. Epigenetics too complex.

**TLDR;**

* Recaptures
* Known ages from a range of ages
* Juveniles
* Infection Y/N

Do a power analysis to figure out how many samples we need! Step 1: figure out our exact questions.

Look into recaps

-catch-plots

1. Histogram of age by sex by species including juveniles
   1. Subset aged and send you sample ID
2. Merging with infection data
3. Include recaptures (Eidolon and Rousettus information)
   1. Know if it’s working in the same way
4. Recaps
5. Use histogram by
   1. Which of these are already extracted (subsample from those that are already done we can prefer those)

Sample ID, age, sex, species, whether they’ve been extractedCloc

MAKE RECAP LIST