Epigenetics Meeting

Who: Bobby Brook (Clock Foundation), Cara Brook, Sophia Horigan

When: May 17, 2024

Purpose: Discuss data back from clock foundation

Background:

Hi Bobby, great to meet you. Won’t take up too much of your time, just had some clarifying questions.

Questions:

* What are the data we were given?

A screenshot of a computer

Description automatically generated

* Illumina data
  + Summary intensities for each probe set on an array

A number on a white background

Description automatically generated

* Three arrays?
* Mammalianarray\_default\_predictions
  + IDATsFound
  + ExtBasename
  + Predictions from clocks
  + How do you get demographic estimates?
  + Fraction failed – high!
  + Outliers
    - All museum samples – does this mean that we don’t trust clock predictions?
* \*\* how do we evaluate our confidence in the prediction made from any specific clock? Does that come from our confidence in the accuracy of the data?
  + i.e. prediction of species, sex, tissue?
  + Which clock has the highest pearson’s correlation with our tooth aged bats?
  + Core sample vs gold standard
* - age
* Large difference in predicted age between clocks – which clock? All bat – Wilkinson
* Normalized Data
* Regression with our tooth predicted ages

Go through csv column by column, link to QCwebsite

Understand raw data

Questions:

* If data wasn’t an outlier for any column, good to go?
* How to determine which clock is best to use?
  + Wilkinson paper
* Understand the raw data
* Epigenetic vs dentition algorithm

Similar work with dolphins and polar bears

* Methylation arrays
  + 12 sample bead chip
    - Chip ID and position
    - Centrix ID and centrix position
* Demographic based on predicted species – which is wrong
  + anAge database
  + all pulled from there
* outlier qualifictions
  + IAC outlier – based on clustering approach
* FFPE restoration kits
  + Museum samples
* Include tooth age and send to them
  + Tooth
  + Recap
    - Anchored by juvenile age
* Built some clocks using recap ages
  + Custom measure for this group
* Universal clock vs species specific vs tissue specific
  + Presumably more accurate the more specific you get
  + But multi tissue clocks work well
* Separate out sites
  + Different seasonality – overpredicts in one versus other, might give some confidence