

Contents

Week 5

| | |
|--------------------------------|---|
| Heritability Exercise. | 2 |
|--------------------------------|---|

Week 4

| | |
|---------------------------|---|
| Molecular Clock | 4 |
| Zika Evolution. | 5 |
| Background | 5 |
| Methods | 5 |
| Results | 6 |

Week 3

| | |
|--------------------------------------|---|
| <i>Populus</i> Stimulations. | 7 |
|--------------------------------------|---|

Week 2

| | |
|--|----|
| Dog Domestication | 8 |
| How Farming Reshaped Our Smiles and Our Speech | 10 |

Week 5

Heritability Exercise

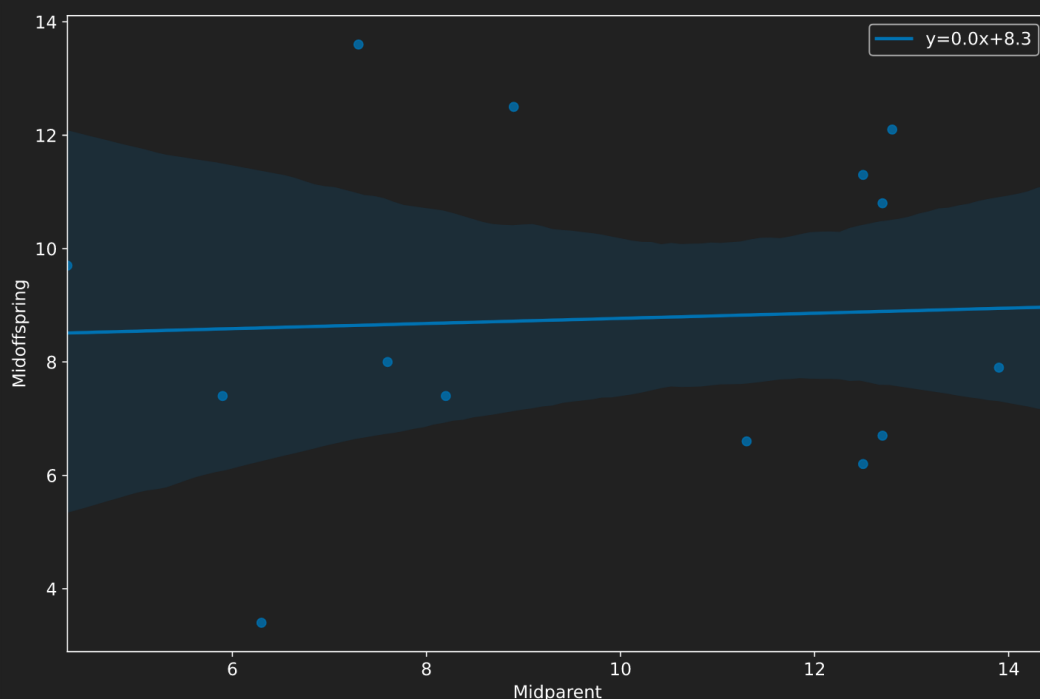
1. Approximately, what is the heritability of running speed in the breeder's dog population?

```
# data given used to create dataframe
d = {
  'Midoffspring':[10.8, 8, 8, 9.7, 6.6, 6.2, 12.5, 7.4,
3.4, 6.7, 7.9, 13.6, 7.4, 12.1, 11.3],
  'Midparent':[12.7, 7.6, 14.4, 4.3, 11.3, 12.5, 8.9,
8.2, 6.3, 12.7, 13.9, 7.3, 5.9, 12.8, 12.5]
}
df = pd.DataFrame(data=d)

# get coefficients of linear fit
slope, intercept, r_value, p_value, std_err =
stats.linregress(df['Midparent'], df['Midoffspring'])

# plot graph, make label for
ax = sns.regplot(df['Midparent'], df['Midoffspring'],
line_kws={'label':
"y={0:.1f}x+{1:.1f}".format(slope, intercept)})

# plot legend
ax.legend()
plt.show()
```



2. If she selectively breeds her dogs, will the next generation run substantially faster than the dogs she has now?
 - ▷ No, very low heritability, nearly 0. ($y=0.0x+8.3$)
3. What else would you suggest the breeder try if she wants to win more races?
 - ▷ Continue to track the midoffspring and midparent speeds, selectively breeding those who tend have a higher heritability ratio in regards to speed.

Week 4

Molecular Clock

CHEETA EXAMPLE:

Divergence: 5.2% per 1mya, Diversity for this locus: 0.182% per [x=time]

$$\frac{0.182}{x} = \frac{5.2\%}{1 \times 10^6}$$

$$1.82 \times 10^5 = 5.2x$$

$$x = \frac{1.82 \times 10^5}{5.2} = 35\,000 \text{ yrs}$$

Divergence: 6.5% per 1mya, Diversity for this locus: 0.182% per [x=time]

$$\frac{0.182}{x} = \frac{6.5\%}{1 \times 10^6}$$

$$1.82 \times 10^5 = 6.5x$$

$$x = \frac{1.82 \times 10^5}{6.5} = 28\,000 \text{ yrs}$$

DOGS EXAMPLE:

Divergence: 7.5% per 1mya, Diversity for this locus: 1% per [x=time]

Note: 0.075, and 0.01 converted to %

$$\frac{1}{x} = \frac{7.5\%}{1 \times 10^6}$$

$$1 \times 10^6 = 7.5x$$

$$x = \frac{1 \times 10^6}{7.5} = 133\,333 \text{ yrs}$$

Zika Evolution

Metsky HC et al. 2017. Zika virus evolution and spread in the Americas. *Nature* 546: 411-415. ⇨

Background

- ▷ What was known about Zika virus evolution before this paper was written?
 - Fewer than 100 ZIKV genomes were previously sequenced directly from clinical samples, providing little information about its epidemiology and evolution.
 - The virus had a link to birth defects and other neurological complications.
- ▷ What were the goals of the authors of this paper?
 - "to gain a deeper understanding of the viral populations underpinning the ZIKV epidemic by extensive genome sequencing."
 - Understand how sequences are changing and how selection might be acting on the virus.
 - Where the virus originated.

Methods

- ▷ What taxa were sampled for this study and where specifically were they from?
 - Samples were collected from United States, Guatemala/ El Salvador, Haiti, Dominican Republic, Honduras, Puerto Rico, Martinique, Jamaica, Brazil, and Columbia.
 - Puerto Rico, Honduras, Columbia, and the Caribbean (includes USA) were the four well supported clades.
- ▷ Why was this sampling strategy important to address their goals?
 - Unbiased metagenomic sequencing of 38 samples proved to have insufficient ZIKV RNA for genome assembly, so PCR amplification (110/229 genomes) and hybrid capture (37/66 genomes) had to be used for enrichment.
 - They had a short sampling period, so deleterious mutations were hard to catch, as well as other changes. They needed a large and diverse set of samples to establish a molecular clock and a substitution rate.
- ▷ How was evidence of selection on the Zika virus genome determined?
 - Total of 110 genomes, plus 64 previous published genomes were used in a reconstructed phylogenetic tree based on a molecular clock.

- Each lineage was then traced back to root and used to estimate the substitution rate, which would provide evidence of selection.
- Aimed to compare ratio between synonomys and nonsynonymous mutations.

Results

- ▷ Looking at the results shown in Figure 2, explain how the authors arrived at the conclusion that the Zika virus outbreak originated in Brazil.
 - The Brazil ZIKV genomes appear on all deep branches of the tree, and the most recent common ancestor is the root of the tree.
- ▷ When did the authors determine the Zika virus outbreak(s) occurred?
 - Early 2014, 95% confidence interval from August 2013 to July 2014.
- ▷ What was the substitution rate for the Zika virus since the outbreak in the Americas?
 - Within-outbreak substitution rate of 1.15×10^{-3} , 95% confidence interval of 9.78×10^{-4} to 1.33×10^{-3}
 - A 1.3–5 times higher rate than other flaviviruses.
 - May be higher than expected as samples did not have time for purifying selection produce an effect.
- ▷ Why did the authors search for whether there was an excess of nonsynonymous mutations in the Zika virus envelope glycoprotein (nicknamed E)?
 - "Viral surface glycoproteins are known targets for positive selection; mutations in these proteins can confer adaption to new vectors or aid immune escape."
 - If there was an increased substitution rate it would indicate for selection.
- ▷ Was evidence of selection found anywhere in the Zika virus genome? Explain.
 - Adaptive mutations are more likely to be found at high frequency among the nonsynonymous mutations.
 - Strong evidence for unrestrained diversifying selection was not found due to similar nonsynonymous substitution rates compared to other coding regions.
 - So no, no strong evidence for selection was found.

Week 3

Populus Stimulations

Exercise 1: Stimulating Strong Selection

1. Describe how the frequency of the A allele changes due to strong selection.
 - ▷ The frequency of the A allele **increases rapidly** due to strong selection.
2. How many generations passed before the A allele was fixed or lost?
 - ▷ Roughly **60–70** generations.
3. How many generations passed before the a allele was fixed or lost?
 - ▷ Roughly **60–70** generations.

Exercise 2: Simulating Modereate Selection

4. Describe how the frequency of the A allele changes due to slightly weaker selection.
 - ▷ Moderate sigmodal increases in frequency.
5. How many generations passed before the A allele was fixed or lost?
 - ▷ Roughly **120–130** generations.
6. How many generations passed before the a allele was fixed or lost?
 - ▷ Roughly **120–130** generations.

Exercise 3: Simulating Weak Selection

7. Describe how the frequency of the A allele changes due to weak selection.
 - ▷ **Slightly increasing** exponential curve.
8. How many generations passed before the A allele was fixed or lost?
 - ▷ Was not fixed or lost after 1000 generations.
9. How many generations passed before the a allele was fixed or lost?
 - ▷ Same, neither allele was not fixed or lost.
10. If a generation was equal to 10 years, how many years would it take for the frequency of the A allele to double in frequency (from 10% to 20%)
 - ▷ About 550 generations, so approximately **5,500** years.

Week 2

Dog Domestication

▷ Background and Goals

- What hypotheses about dog domestication are examined in this paper?
 - The origin of domestic dogs was the main focus; genetic data suggests East Asia, but other data suggests Europe and Siberia.
- What predictions do these hypotheses make? What would the expected phylogenetic trees look like under each of these hypotheses (i.e. what are the expected topologies)?
 - The most recent common ancestor would indicate origin and genetic data would help confirm time frame.

▷ Methods

- What taxa were sampled for this study?
 - Hypothesis was tested using DNA extracted to trace mitochondrial genomes from 18 prehistoric canids and 20 modern wolves from Eurasian and American origin.
- What gene sequences were used for this study?
 - Wolves, dogs, including divergent breeds, recently published Chinese indigenous dogs, and coyotes. (148 mitochondrial genomes)
- What phylogenetic methods were used to reconstruct the phylogenetic tree in Figure 1?
 - Maximum likelihood, coalescence, and Bayesian.

▷ Results and Discussion

- In Figure 1, which wolf sequence is most closely related to the extant dog clade A, and what geographic region is it from?
 - Wolves: Switzerland 214,500.
 - Dogs: Argentina, USA (8,500, 1,000)
- In Figure 1, which clade of modern dog sequences has the oldest origin? Approximately how old is this clade?
 - Dog clade A, Russia, 18,000 years ago.
- Based upon Figure 1, how many times does it appear that modern extant dogs could possibly have been domesticated?

- Figure 1 suggests 4, but the authors state, "the inferred recent divergence of clade B from wolves now found in Sweden and Ukraine implies that it might represent a mitochondrial genome introgressed from wolves rather than one established by domestication."
- What was the topology of the phylogenetic tree that was recovered, and therefore, which hypothesis was supported?
 - The findings support the conclusion that legacy of dogs derives from wolves of European origin.
 - Analysis of coalescence times support divergence time of >15,000 years ago.
 - Past mitochondrial and Y chromosome analysis suggested non-European, but had less supported data than these findings.
- ▷ Conclusions
 - What are the major conclusions of the paper?
 - Three of four modern dog clades are more closely related to sequences from ancient European rather than extant wolves with divergence times >15,000 years ago.
 - Was the taxonomic sampling sufficient to rule out alternative hypotheses?
 - Ancient panel did not contain specimens from Middle East or China. No, no ancient dog remains older than 13,000 years are known from those regions.
 - The mtDNA sequence tree is well supported, but represents a single genetic locus. Independent loci could offer more power to resolve phylogenetic relations.

How Farming Reshaped Our Smiles and Our Speech

▷ Gibbons A 2019; doi: 10.1126/science.363.6432.113

1. What is this news article about?

▷ How farming may have reshaped our jaws and thus indirectly influencing how we speak.

2. How could diet change the structure of the jaw?

▷ Less wear and tear changed how jaws and teeth developed, creating an overbite with the now less worn down teeth.

3. How could jaw structure change speech patterns of humans?

▷ Jaw and teeth placement changes how easy it is to make certain sounds, and overbites made "f" and "v" easier to make.

4. Are these changes in jaw structure permanent? Why or why not?

▷ Not likely, if you ate diets similar to early humans your teeth would wear down more and impact from wisdom teeth would minimize overbite. Though, there may have been some lasting selection pressure for a natural overbite.

5. Did natural selection shape these changes in jaw structure? Explain.

▷ Not likely, mostly likely from behavior change due to mostly diet change.

6. Can natural selection shape speech patterns? Explain.

▷ Yes, sexual selection could be a significant. Or if the change in speech pattern could confer a advantage in communication.

7. Can you think of any parallels with other animals where diet influenced feeding structures?

▷ Lots of animals, birds (finches!!), dogs, snakes. Huge impact.