

Exercise Set 05: Good and Bad Visualizations

BEE 4850/5850, Fall 2024

Name: Anthony Nicolaides

ID: ajn68

Due Date

Friday, 2/23/24, 9:00pm

Overview

Instructions

The goal of this exercise is for you to find and evaluate data visualizations which you think do a particularly good and bad job.

Load Environment

The following code loads the environment and makes sure all needed packages are installed. This should be at the start of most Julia scripts.

```
In [ ]: import Pkg
        Pkg.activate(@__DIR__)
        Pkg.instantiate()
```

Activating project at `~/Documents/BEE5850/exercises/ex_week05_BEE5850`

The following packages are included in the environment (to help you find other similar packages in other languages). The code below loads these packages for use in the subsequent notebook (the desired functionality for each package is commented next to the package).

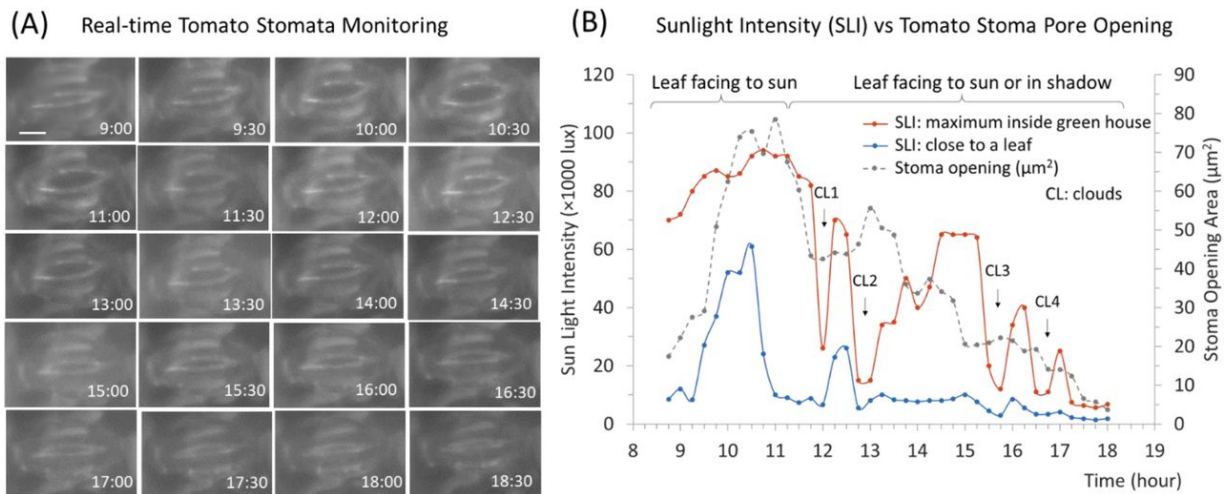
```
In [ ]: using DataFrames # tabular data structure
        using CSVFiles # reads/writes .csv files
        using Plots # plotting library
        using StatsBase # statistical quantities like mean, median, etc
        using StatsPlots # some additional statistical plotting tools
```

Problem

Find an example of a data visualization (could be from any reasonable source: journalism, a scientific paper, generated from data) that you think does a particularly good job of communicating something about the underlying data, and one which does a particularly bad job. Write a brief summary (one paragraph) for each about what the visualization is trying to communicate and what makes it (in)effective. Make sure to include where I can find a raw version of the figure (or if you generated it yourself) which is higher-resolution for use in the class discussion next week.

Effective Visualization

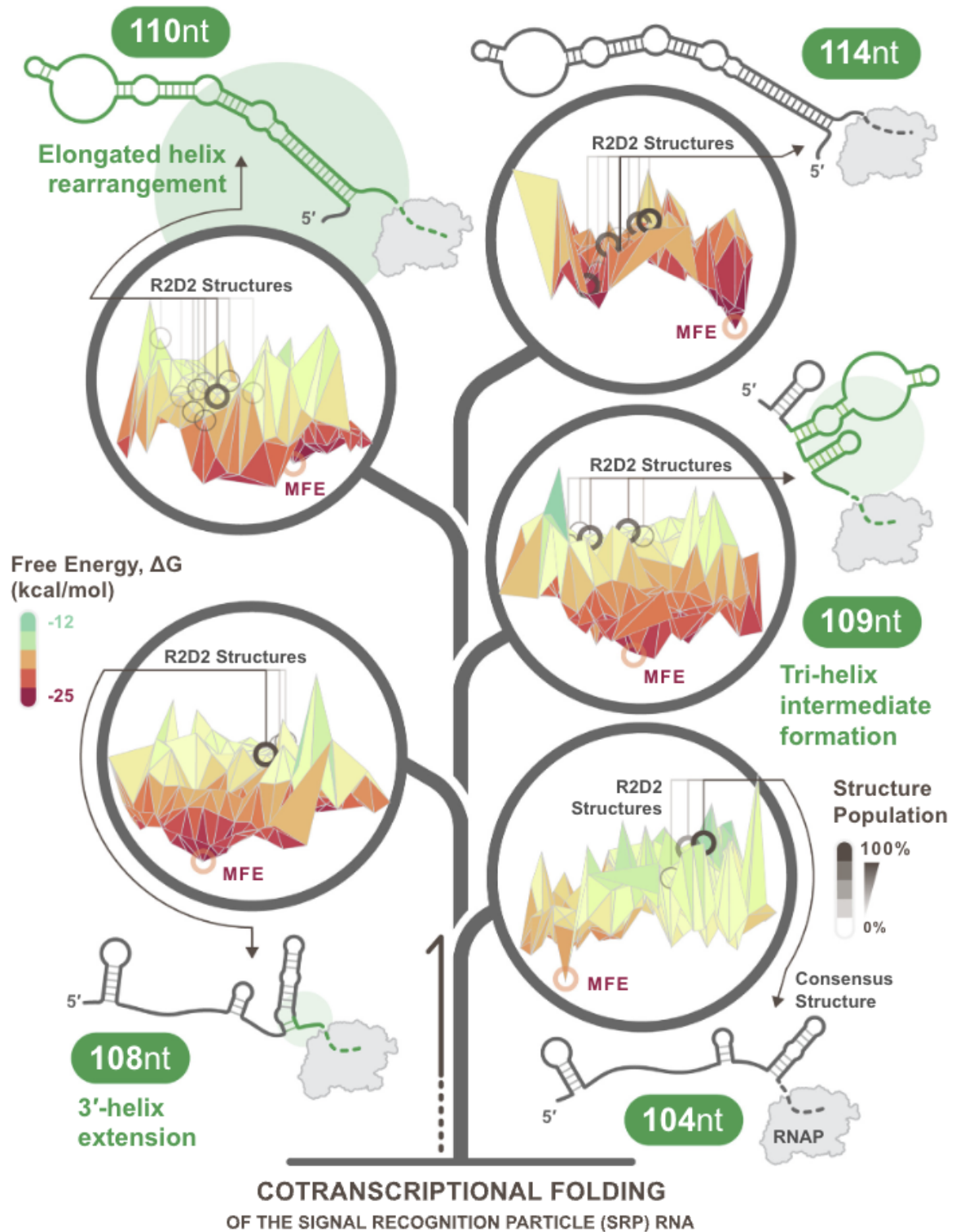
```
In [ ]: display("image/png", read("BEE5850_ex05_good.png"))
```



This paper is about leaf stomata activity and how to monitor it based on sunlight variability. Stomata are small leaf pores on the underside of leaves that help with gas and water exchange with atmosphere. This particular figure is meant to show stomata activity in a tomato plant through time. I think this is a good visualization of data because it is simply and to the point with what it is trying to say: stomata open when they are more exposed to sunlight. The figure (B) also shows when clouds are present without cluttering the graph too much. The color are distinct and distinguishable from each other. Figure (A) also shows what the stomata actually looked like at specific hours in the graph. Overall, I think this figure is concise and clearly conveys a message.

Ineffective Visualization

```
In [ ]: display("image/png", read("BEE5850_ex05_bad.png"))
```



This figure demonstrates the RNA folding pattern as it is transcribed from DNA. As RNA polymerase converts DNA to RNA, the RNA freely floats in the cell and undergoes folding. This paper looks at how RNA folds like it is still being transcribed, so it is still tethered to RNA polymerase and new nucleotides get introduced. The figure itself is showing that as new nucleotides get added, the RNA undergoes new folding conformations. Counterintuitively, the RNA does not fold into the minimum free energy structure (MFE), as shown by the free energy landscape in the R2D2 structures.

The reason I think this is not a great visualization of data, is that the message does is not clear when you look at it. The figure relies on color gradients (free energy), saturation gradients (how common a particular structure is), and 3D representation (3D free energy landscape). When description of the figure does help provide context, however without it the figure looks clustered. Furthermore, the free energy landscape does not provide as much information as the author might think, while the lowest free energy is marked with a red circle and MFE, there are multiple gray circles. Some of these gray circles are blocked by the 3D landscape itself, which is counterproductive to clear communication of the idea.

References

Good visualization figure (Figure 7):

[In-situ Real-time Field Imaging and Monitoring of Leaf Stomata by High-resolution Portable Microscope](#)

<https://doi.org/10.1101/677450>

Bad visualization figure (Figure 1):

[How does RNA fold dynamically?](#)

<https://doi.org/10.1016/j.jmb.2022.167665>