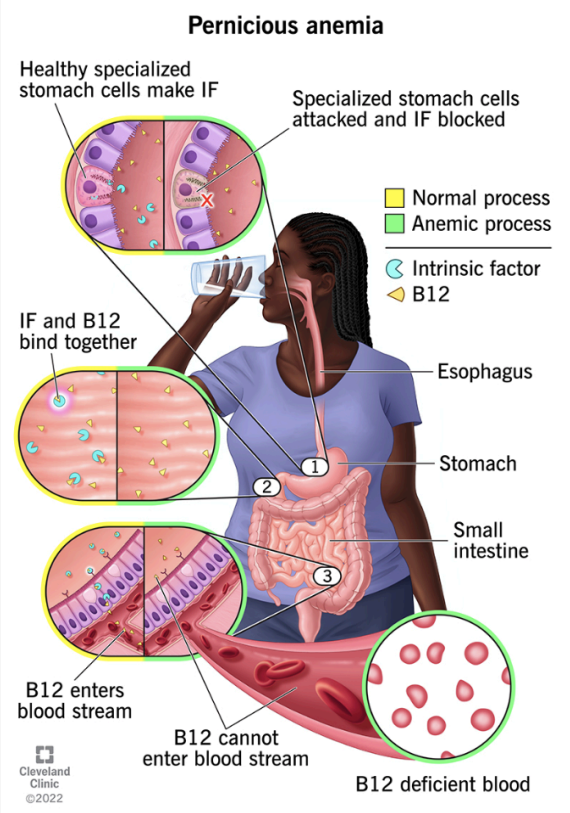
Project Report: Analyzing the Level of Corrinoid in Different Fermented Foods

1. Introduction



It has long been hypothesized amongst the scientific community that the fermentation process of plant products affects the level of corrinoid produced. Corrinoids are effectively vitamins, the most well known being B12. However, many companies test for B12 in their products to report with methods that are flawed and lead to inconclusive results. What this leads to is a misreport in the quantification of B12 in their foods which poses serious health threats to those who are B12 deficient, mainly in the condition of pernicious anemia.

In this project, we analyzed the level of corrinoid in two store bought products, Nancy's Kefir and Humm Kombucha in efforts to calculate the level of B12 in each and see if the advertising was false or true. Seen below are the reported concentrations of B12 in the two tested samples that will be used to compare to the final results, to see if their advertising is accurate.

Nancy's Organic Lowfat Peach Kefir	yes	<0.1ug/100g
Humm Kombucha (any flavor)	yes	212.9 ug/226.8g

2. Data Source

The data used was taken from experiments done by Clara Bardeen in the Taga lab. Effectively what was tested was the growth of the bacteria "MetE" over time in wells that contained the foods of interest. If MetE grows, then there is corrinoid or methionine present in the food and it can be quantified. The data itself was taken on an instrument called "Tecan" that measures the optical density of bacterial strains over time which then quantifies their relative level of growth.

3. Equation to Fit Data

The equation used to calculate the actual level of corrinoid present in each sample is used through the generation of a standard curve. This standard curve is created by having increasing dilutions of B12 inoculated with MetE and then measuring their optical density over time. The standard curve at the end of an 11 hour period should resemble a linear line with an R^2 value close to 1. This standard curve is separate from the growth curves that are used simply to see if corrinoid even is present. If it is, then the standard curve's equation is used to calculate the actual level of corrinoid present. The level of corrinoid present would be x, where y is the optical density of the strain (Met E) at time 11 hours.

Standard Curve Formula; y = ax + b
Coefficient of Determination; R^2 = 1 - RSS/TSS

RSS = sum of squares of residuals
TSS = total sum of squares

4. Data Filtering

The data was filtered through using the loc method, where only desired wells at the specific time points were used. This method was mainly used for the calculation of corrinoid concentration, since an exact time point is necessary to plug in for the y value. Other data filtering methods were not necessary as the graphs created relied on the growth of one well for the entire time period, with no filtering to be applied.

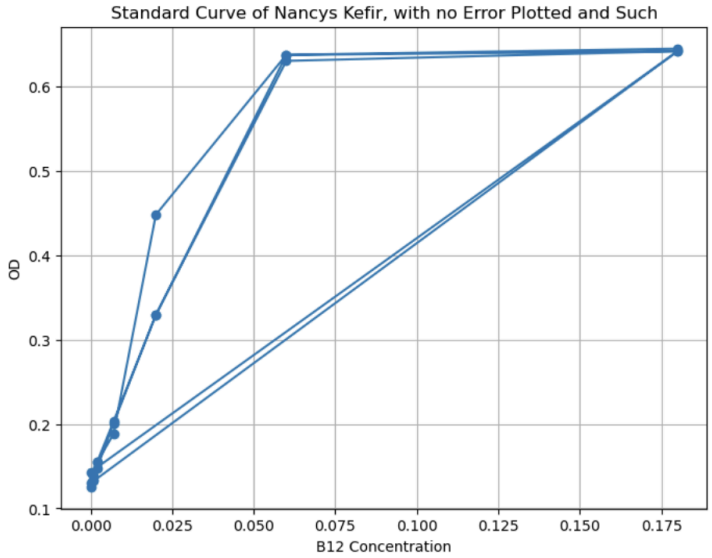
5. Data Fitting/Thought Process

This project created three types of data displays. Firstly, the growth curves of Nancy's kefir over time. Secondly, the standard curve for both Humm Kombucha and Nancy's Kefir. Thirdly, a data frame to display the final calculated concentrations of corrinoid for both tested samples. After fitting that data to generate a standard curve as well as individual growth curves pertaining to the foods of interest we were able to create:

- A logarithmic growth curve of each bacterial strain in increasing concentrations of B12 over time, with a coefficient of determination close to 1.
    a. The standard curve equation was also included.
- The growth of bacterial strains in Nancy's Kefir and its different dilutions over time.
- A quantification of how much corrinoid is present in each fermented food and at each concentration using the equation generated by the standard curve.

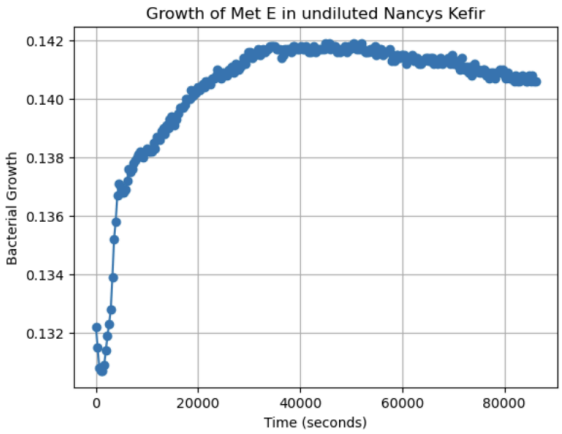
An overview of how we approached this project is detailed below:

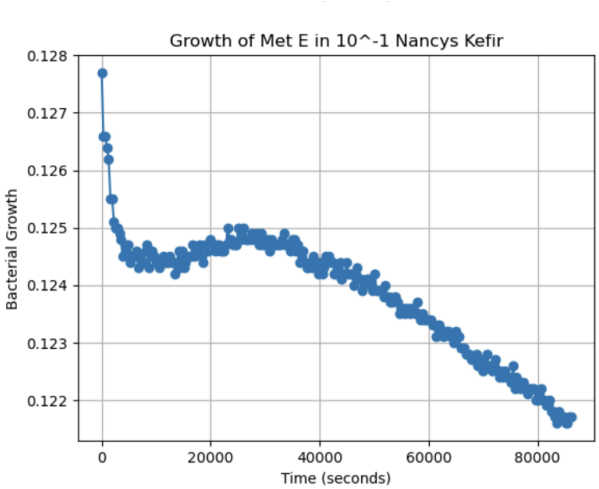
In the beginning we used the launchers numpy, pandas, matplotlib.pyplot, and also we attempted to use the launcher seaborn for our final, but failed and found a way around it. Matplotlib was used to generate the general structure of a standard curve using uploaded CSV from lab data.

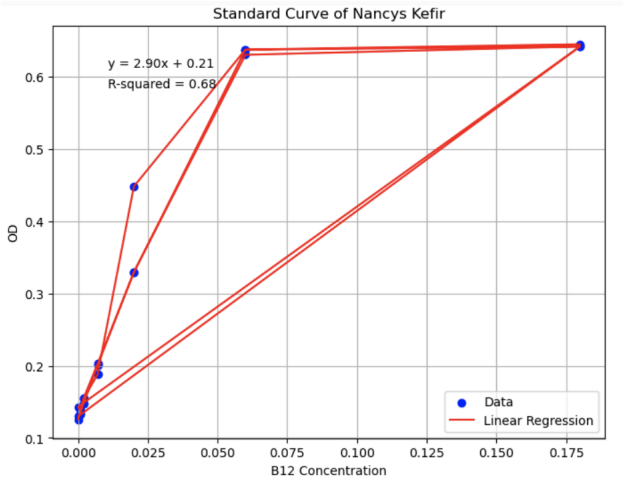


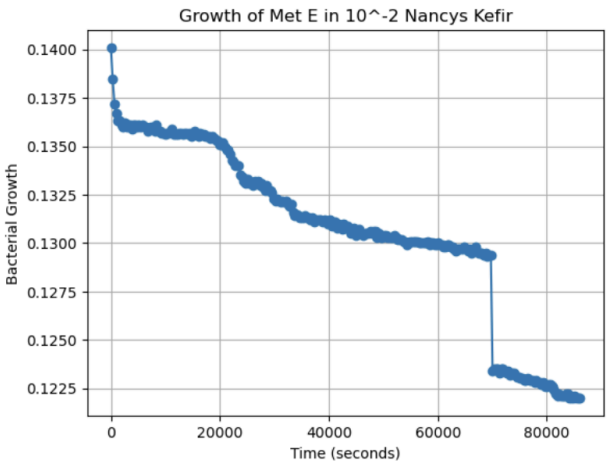
Next, we plotted the standard curve of both samples (Humm kombucha and Nancy's Kefir). Initially, we consulted reddit where we found advice on how to calculate the r squared value using sklearn. While this did work, we did not fully understand what it meant so we rewrote the code using pandas, numpy, and matplotlib.pyplot. First, we defined a function to calculate linear regression using the slope and intercept. Then we defined a function to calculate r squared by making predictions and finding the residual sum of squares. After this baseline of formula was written, we could input the data from the two different CSV's. The data used for the generation of these standard curves was the growth of the bacterial strain MetE in increasing concentrations of B12 (Vitamin B). By graphing this gradual growth of the bacteria in pure vitamin, the data for the growth of the foods themselves can be compared to quantify how much corrinoid is present in them. Thus, we graphed the first 3 columns of each CSV which pertained to the growth of the bacteria in varying B12 concentrations.

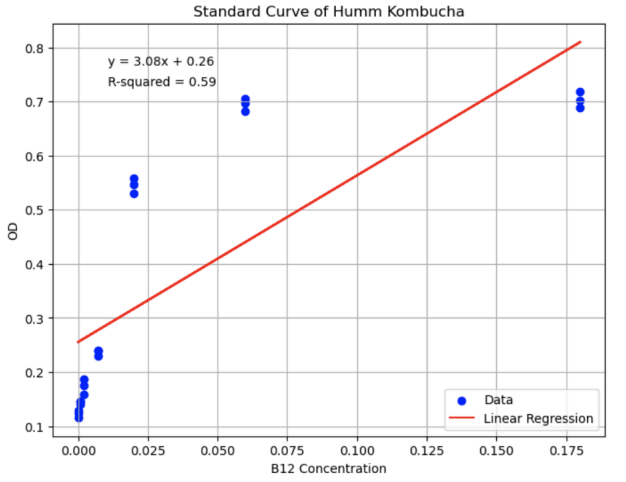
Then, we graphed the growth of MetE in varying dilutions of Nancy's Kefir. This is just to prove that there is or is not a corrinoid present. Since MetE grows in the presence of B12 and methionine, if it also grows in Nancy's Kefir then there is likely methionine or B12 present. The process of doing so was quite simple. We uploaded the CSV containing the growth of MetE in Nancy's Kefir over a 11 hour period and used matplotlib to plot the columns of interest. We can see those graphs along with the two standard curves below:


Growth of Met E in undiluted Nancys Kefir
Bacterial Growth
Time (seconds)
0.142
0.140
0.138
0.136
0.134
0.132
0
20000
40000
60000
80000


Growth of Met E in 10^-1 Nancys Kefir
Bacterial Growth
Time (seconds)
0.128
0.127
0.126
0.125
0.124
0.123
0.122
0
20000
40000
60000
80000


Standard Curve of Nancys Kefir
y = 2.90x + 0.21
R-squared = 0.68
OD
B12 Concentration
Data
Linear Regression
0.6
0.5
0.4
0.3
0.2
0.1
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175


Growth of Met E in 10^-2 Nancys Kefir
Bacterial Growth
Time (seconds)
0.1400
0.1375
0.1350
0.1325
0.1300
0.1275
0.1250
0.1225
0
20000
40000
60000
80000


Standard Curve of Humm Kombucha
y = 3.08x + 0.26
R-squared = 0.59
OD
B12 Concentration
Data
Linear Regression
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175

6. Conclusion

After generating the standard curves for each food sample, we were able to rearrange the equation to solve for x where y is the OD of each food sample at the 11 hour time interval. We ended up getting the table below. Where 0 represents the undiluted sample, 1 represents 10^-1 dilution and so on for the two food samples. If the values are negative, then there is no corrinoid present. The positive values must be multiplied by 10 to account for the dilution that went into processing the initial food sample. In short, the only corrinoid present was 2.17532 ug in Humm kombucha which is off by a factor of 100 in its advertising as included in the Introduction section of this report. While this might be due to environmental factors such as contamination, an inaccurate standard curve, etc. we can conclude, for the most part, that the heavily supplemented Humm kombucha does not have the appropriate level of B12 reported on its packaging and is likely contributing to high levels of B12 deficiency amongst the general population. Nancy's kefir does not have corrinoid detected which follows its reported packaging.

	Nancys Kefir	Humm Kombucha
0	-0.023552	0.217532
1	-0.029586	-0.016753
2	-0.027241	-0.042240