Theses/Capstones/Creative Projects   University Honors Program

Comparative Analysis of Metabolic Pathways of Bacteria Used in Fermented Food

Keanu Hoang
kkhoang@unomaha.edu

Kiran Bastola
University of Nebraska at Omaha, dkbastola@unomaha.edu

Comparative Analysis of Metabolic Pathways of Bacteria Used in Fermented Food

Keanu Hoang

School of Interdisciplinary Informatics
College of Infomation Science & Technology
University of Nebraska Omaha
kkhoang@unomaha.edu

Abstract—**This study presents a novel methodology for analyzing metabolic pathways. Utilizing KEGG REST API through a Biopython package and file parser, data about whether or not a bacteria has an enzyme or not was extracted. The results found that differences in metabolic pathway enrichment values follow along the lines of genera and pathway type. In particular, bacteria found in food spoilage and commercial nitrogen fixing products had high values of enrichment.**

Index Terms—metabolism, KEGG, enzyme, fermentation, zymology, milk, fish sauce

I. Introduction

Fermentation is an ancient food processing technique that people have employed primarily to increase the shelf-life of food. The science of fermentation is known as zymology or zymurgy but it is often referred to the chemical reactions that converts sugars to alcohol. However, this process also has secondary benefits including the supply of nutrients and supplemental microorganisms. In this project, we will first review relevant literature revolving around the health benefits of fermented food products, especially those that could impact the starting chemical compositions of fermented food products.

A. Milk

Milk is one of the many food products that are fermented. These fermented milk products have nutritional benefits that go beyond the ones that are obtained from its base ingredient. According to Korhonen and Pihlanto, there is a growing focus "on physiologically active peptides from milk proteins," [2]. One of the ways these peptides are liberated from the parent protein is through the use of proteolytic starter cultures which assists in the fermentation of the food product [2]. This phenomenon is well documented [2] but the mechanisms in which these fermented products "exert their effects are largely unknown," [6].

However, the functionalities of these peptide have been extensively studied [2], [6], [8]. It is well-known and documented that these products have nutritional benefit. Korhonen and Pihlanto outline that many of the bioactive peptides have ACE-inhibitory effects. This is corroborated by Parvez, they write "probiotic bacteria are important in the downregulating inflammation associated with hypersensitivity reactions in patients with atopic eczema and food allergy" [6]. There is variability among ACE-inhibitory activity among fermented milk products (i.e. cheese). Korhonen writes:

[Researchers] detected ACE-inhibitory activity in several cheese varieties and measured the highest activity in Gouda cheese aged 2 years. In feeding experiments on SHR, the decrease in systolic blood pressure was statistically significant with four cheese varieties (Gouda, Blue, Edam and Havarti) at 6 h after gastric intubation. Several peptides were isolated and identified from 8-month-old Gouda cheese, and two peptides derived from as1-casein f(1-9) and b-casein f(60-68), respectively, showed potent angiotensin converting enzyme (ACE)-inhibitory activity. For Manchego cheese, which is prepared from ovine milk, only cheese that was at least 15 days old showed ACE-inhibitory activity that was comparatively low. [2]

With the power of bioinformatics tools, novel information about the effects of the processes involved in fermentation can be found and later studied. This information can be stored and annotated in databases that researchers can conveniently access. Parvez describes known health benefits of probiotic bacteria present in fermented milk products including yogurt and cheese [6]. These effects include a role in blood pressure control and possible reduction of cancer risk via anti-carcinogenic metabolism and counteracting mutagenic and genotoxic effects [6]. However, there are multiple hypotheses on the methods in which this happens. One such hypothesis could be "the bacteria engaged in fermentation produce volatile organic compounds, which gets absorbed by the cells to control inflammation". Additionally, bioinformatics approaches have been used to discover and characterize the bioactive components of fermented food items. For example, Sanchez-Rivera et al. describes a plethora of approaches including in silico structure prediction and analyses of peptidesin the food matrix [7]. Nielsen et al. created a milk bioactive peptide database to map peptides to know bioactive peptides in milk and apply models to predict novel functional peptides in human milk [4].

B. Fish Sauce

Fermentation is a process that has been used world-wide. Many of the studies found that these cultural staples also had health benefit along with the characterization of complex flavor profiles and aromatic quality. Fermentation is especially

relevant in tropical countries because the food with high water content and nutrients cause an increase in perishing rate [5]. While this fact contributes to the importance of fermentation within these cultures and may contribute to their food habits, fermentation attributed to the probiotic quality of fermented products has been shown to correlate to antimutagenicity [5].

Fermentation of meat or fish is often associated with spoilage. However fish sauce is important to many Asian cultures and studies have shown that bioactive peptides produced in fermented marine food sauces by enzymatic hydrolysis by microorganisms have been involved in hypertension regulation, immune system modulation, and anticancer activities [8]. This prompts the question "How similar are the metabolism of these microbes which ferment meat compared to the lactic acid bacteria and others that are often to be involved in the fermentation of dairy and plant materials?" This question becomes especially important when it has been shown that the composition of different bacteria during fermentation may not result in the same sensory and nutritional values [1]. As of 2011, metagenomic and metabolic tools were used to characterize the microbial community and related metabolites of kimchi [Korean fermented cabbage] during spontaneous fermentation, [1].

II. MOTIVATION

In terms of food production and health, it is important that researchers know the root of the complexities of the taste and nutritional value of a given fermented product. Why do we eat what we eat? Is there a benefit that fermentation flora give that the microbiome of the gastrointestinal system does not have? With the information explored in this study, food scientists and bioengineers will have the tools to outline future studies and technologies that will be able to benefit the health and experience of consumers. The goal of this project is to illustrate the relationships of microorganisms involved in food fermentation. In particular, the fermentation of wines and cheeses will be studied.

There is a gap in the knowledge base about the fermentation process that involves the specific mechanisms in which microbiota liberate bioactive peptides and contribute to the other health benefits of consuming of fermented product. Also, it is not well documented as to how correlated the aromatic and flavor-contributing components are to the health benefits of fermented products. Research should concentration of the root of the fermentation process and determine how humans can not only improve upon advancing the fermentation methods to increase the amount of alcohol production through fermentation but also determine health benefits of consumed fermented food. The overarching goal of the present study is to determine potential dependencies of bacteria involved in fermentation processes under different fermentation conditions. In other words, how would different bacteria species known to ferment food products rich in carbohydrates differ or behave in the protein rich environment? To accomplish this goal, the following aims will be addressed:

1) Identify pathways involved in fermentation

2) Compare the important pathways of plant & animal product fermentation versus protein-rich fermentation
3) Comparative analysis of different bacteria

III. MATERIALS & METHODOLOGY

A. Data Collection

The following bacteria were found in cheese [6] and data about which enzymes that bacterium has in a pathway were pulled:

- Lactobacillus brevis ATCC 367
- Lactococcus lactis subsp. lactis Il1403
- Leuconostoc sp. C2
- Leuconostoc lactis
- Leuconostoc mesenteroides subsp. mesenteroides J18
- Clostridium tyrobutyricum

The following bacteria were found in fermented fish sauce [8] and data about which enzymes that bacterium has in a pathway were pulled:

- Weissella cibaria
- Lactobacillus sakei
- Bacillus licheniformis ATCC 14580
- Halobacterium salinarum
- Tetragenococcus halophilus

Escherichia coli and Staphylococcus aureus were also analyzed as they are a part of food spoilage and found in the gut biome [8].

A selection of metabolic pathways were looked at, with the main focus on carbohydrate metabolisms; in particular,

- Glycolysis
- Fructose/mannose metabolism,
- The citric acid cycle (TCA)
- Pyruvate metabolism
- Carbon fixing
- Nitrogen fixing

A Python script was created to pull primary IDs of data that are cross-referenced among all the relavent databases (Ortholog, Gene, Enzyme) within KEGG using their API REST. Specifically, the script utilizes REST's linking function that connects the Enzyme Commission (EC) number to KEGG orthologs to all the genes that an ortholog has to verify whether an organism has a specific enzyme or not. The script iterates through each ortholog involved in a pathway and links that ortholog to all the genes in KEGG that are categorized under that ortholog as well as the EC number assigned to it. The script then parses through each line of the link between ortholog to gene and checks whether or not it belongs to an organism in the set organism list. Fig. 1 illustrates the table format that the code outputs. The first column is the three or four letter code assigned by KEGG to a species of bacteria. The second column and third column is a pathway and an enzyme's EC number respectively. The fourth column denotes whether or not a bacteria has that enzyme or not (0 is DOES NOT HAVE and 1 is HAS). Code can be found here: https://github.com/kkhoang/FermentKEGGRest

TABLE I
ENRICHMENT VALUES CALCULATED

Organism	Glycolysis	TCA	Fructose/Mannose Metabolism	Pyruvate Metabolism	Carbon Metabolism	Nitrogen Metabolism
Spoilage Bacteria						
Escherichia coli O157:H7 EDL933 (EHEC)	0.74	0.70	0.80	0.70	0.61	0.63
Staphylococcus aureus RF122	0.74	0.74	0.31	0.62	0.55	0.38
Cheese Bacteria						
Lactobacillus brevis ATCC 367	0.47	0.17	0.11	0.36	0.26	0.19
Lactococcus lactis subsp. lactis Il1403	0.59	0.39	0.34	0.43	0.42	0.19
Leuconostoc sp. C2	0.50	0.13	0.20	0.34	0.26	0.13
Leuconostoc lactis	0.50	0.13	0.23	0.36	0.23	0.13
Leuconostoc mesenteroides subsp. mesenteroides J18	0.47	0.17	0.14	0.32	0.29	0.19
Clostridium tyrobutyricum	0.56	0.35	0.37	0.49	0.45	0.44
Fish Sauce Bacteria						
Tetragenococcus halophilus	0.62	0.39	0.40	0.45	0.35	0.31
Weissella cibaria	0.38	0.13	0.20	0.30	0.29	0.19
Bacillus licheniformis ATCC 14580	0.79	0.78	0.63	0.70	0.65	0.69
Lactobacillus sakei	0.56	0.13	0.29	0.36	0.26	0.13
Halobacterium salinarum	0.59	0.65	0.17	0.40	0.68	0.19

```
BACTERIA   PATHWAY EC NUMBER  HAS? (0 | 1)

  B1        P1       E1           1
  B1        .        .            .
  B1        .        .            .
  B1        .        .            .
  B1        PN       EN           0
  B2        P1       E1           1
  B2        P2       E2           1
  .         .        .            .
  .         .        .            .
  .         .        .            .
  BN        PN       EN           1
```

Fig. 1. The format of the output from the Python script. Does not denote actual values. BACTERIA column contains 3-4 letter code for a given bacterial species, PATHWAY column contains the reference code for a given pathway, EC number column contains all the EC numbers of an enzyme in a pathway, and HAS? column contains the number 1 if a bacteria has that enzyme and a 0 if not.

B. Calculations

Based on the nature of the tables created from the script, networks could be made to visualize how bacteria connect to enzymes and subsequently to pathways, but the networks were too large to make qualitative analysis viable. Rather, data was quantified to make interpretation more streamline. The number of edges between a bacterium species and enzymes for a given pathway (i.e. the total number of 1's a bacterium has for a pathway) was divided by the total possible number of enzymes in that pathway.

$$EnrichmentValue = \frac{N_{Edges}}{N_{Pathway}} \qquad (1)$$

Where N_{Edges} equals the number of edges found for a given bacteria and $N_{Pathway}$ is the total number of enzymes

in a pathway. For example, if a bacteria has 35 enzymes in a pathway out of the possible 50 enzymes in that pathway, that enrichment value would be calculated to 0.70.

C. Statistical Calculations

Qualitative analysis of the distributions of enrichment values (from Eq. 1) between bacteria found in fish sauce and that of cheese prove useful in distinguishing the differences of metabolisms. Using R, histograms were made of the enrichment values per metabolic pathway. Distributions of enrichment values of bacteria found in cheese fermentation vs. bacteria found fish sauce fermentation for a pathway were created. Code for the generation of histograms can be found here: https://repl.it/@KeanuHoang/bacenz2

IV. RESULTS

Using a spreadsheet, enrichment values were calculated for each organism per metabolic pathway using Eq. 1. Table 1 describes the values for the enrichment values. What was found was that all the glycolytic pathways had high values relative to other pathways regardless of species. Lactic acid bacteria (LAB) had lower values for TCA. This is expected as LAB do not rely on aerobic metabolism for energy production. LAB also had typically lower values for each metabolic pathway analyzed. The bacteria found in meat spoilage (E. coli & S. aureus) had partiarly high enrichment values across the board for all metabolic pathways analyzed. Species in the Leuconostoc genera typically have lower values for all metabolic pathways.

Figs. 2-7 describe the distribution of enrichment values for each particular pathway that were pulled from Table 1. The small sample size and non-Normal distribution pushed towards the lack of use of hypothesis testing. Qualitatively, these histograms showed a slight difference in metabolic processes of fish sauce fermentation and milk fermentation.
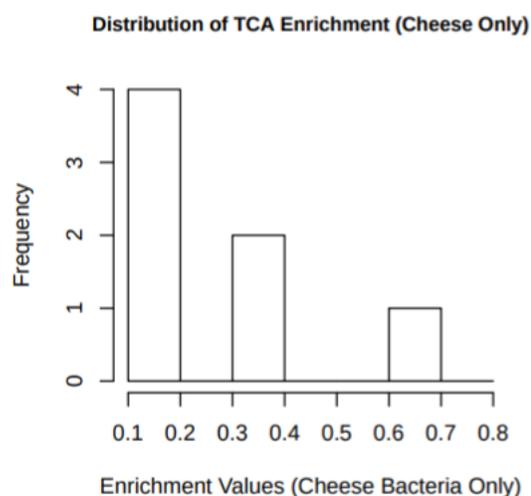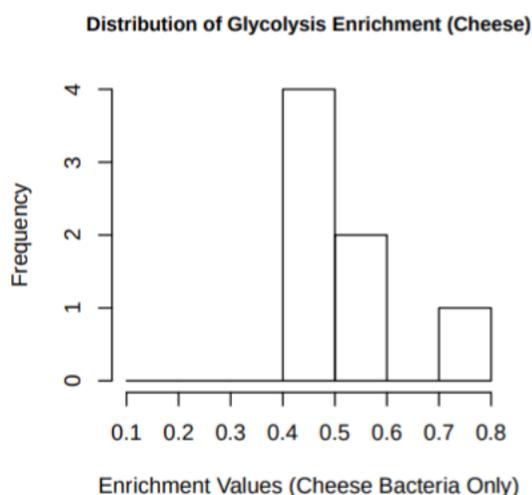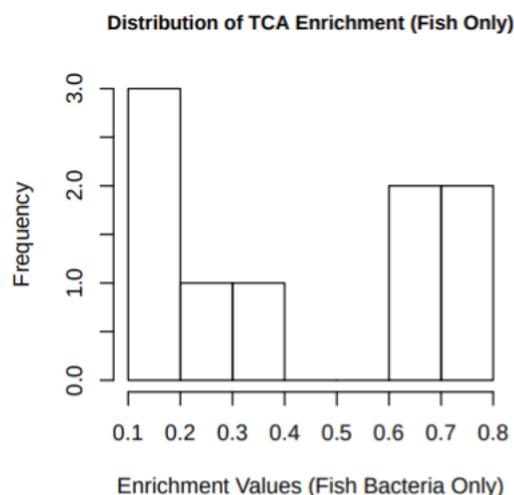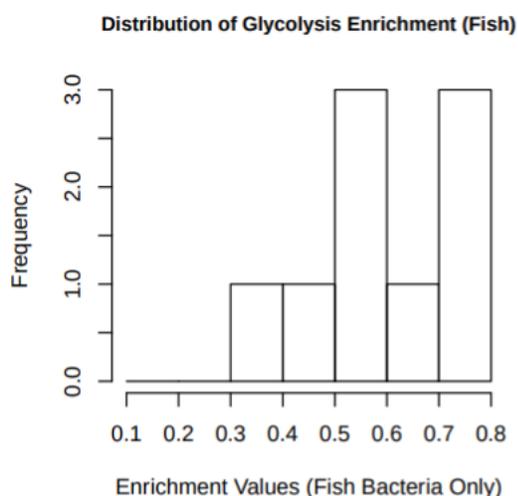
Fig. 2. Distribution of enrichment values of bacteria found in fish sauce versus that of milk fermentation in glycolisis.

Fig. 3. Distribution of enrichment values of bacteria found in fish sauce versus that of milk fermentation in the citric acid cycle (TCA).

The distibutions typically have bimodality for bacteria found in fish fermentation as well as higher median/mode values. Due to the small number of bacteria and metabolic pathways involved, these distributions are sparse in frequencies. Fig. 2 illustrates the differences between bacteria in fish sauce and cheese in glycolysis. Although the range of enrichment values were nearly the same, the frequency of higher values are the main difference in glycolysis. This is mostly due to the high values calculated for B. licheniformis and H. salinarum. Figs. 5 and 6 show a similar result for pyruvate metabolism and carbon fixing metabolism respectively. There exists more bacteria that have high enrichment values for pyruvate metabolism and carbon fixing metabolism. Fig. 3 illustrates within the citric acid cycle, that although both fish sauce bacteria and cheese bacteria have low frequencies in TCA enrichment values, fish sauce bacteria have more

than one species that have enrichment values higher than 0.6 while cheese bacteria do not. Fig. 7 describing nitrogen fixing metabolism mirrors this result. While there are cheese there exists a bacterium found in cheese that have an enrichment value over 0.6, there are more bacteria that are found in fish sauce that have high enrichment values. Fig. 4 shows that Fructose/Mannose metabolism among fish sauce bacteria and cheese bacteria similar. The distribution of enrichment values have high frequency at low enrichment values while a few outliers have an enrichment value higher than 0.5.

V. DISCUSSION

A. Interpretation

For the glycolytic pathway, it appears that there is little discrepency in whether or not fish sauce bacteria and cheese bacteria can perform glycolysis well as the enrichment values typically are higher than 0.4 as described by Fig. 2. This makes

Distribution of Fructose/Mannose Enrichment (Fish)
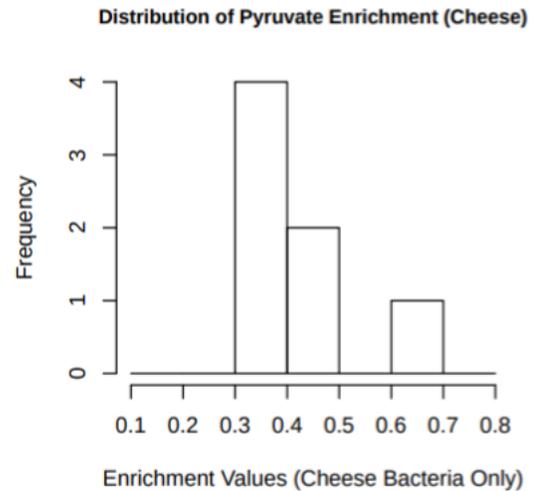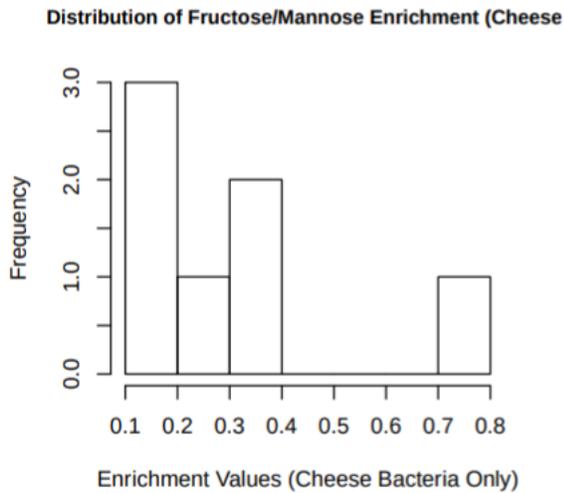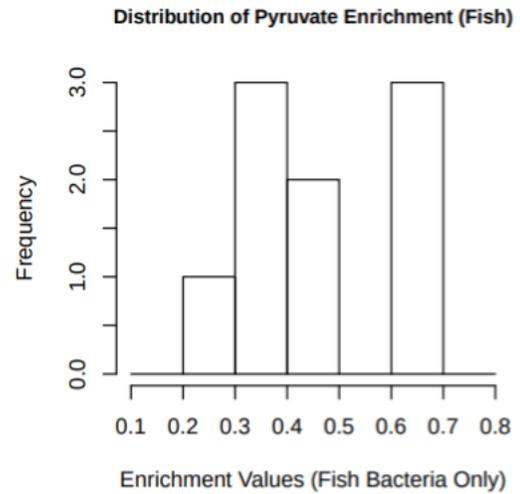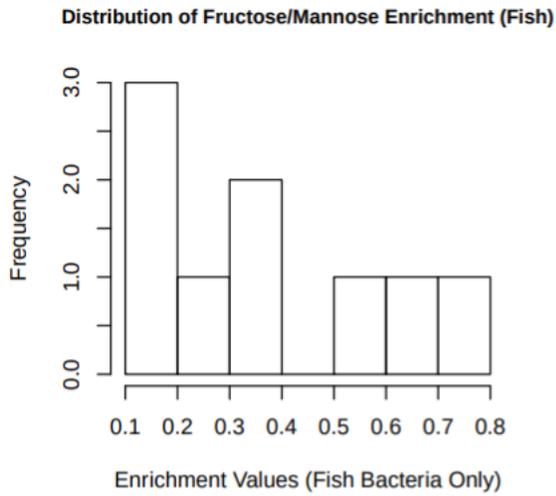
Distribution of Pyruvate Enrichment (Fish)

Fig. 4. Distribution of enrichment values of bacteria found in fish sauce versus that of milk fermentation in fructose/mannose metabolism.

Distribution of Fructose/Mannose Enrichment (Cheese

Distribution of Pyruvate Enrichment (Cheese)

Fig. 5. Distribution of enrichment values of bacteria found in fish sauce versus that of milk fermentation in pyruvate metabolism.

sense as glycolysis is the major pathway that leads to other energy-producing glycolytic pathways. Pyruvate metabolism and carbon fixing metabolism, showing a similar result in the distribution of enrichment values, as illustrated by Figs. 5 and 6, among cheese and fish sauce bacteria populations. This means that both sets of bacteria have a set of enzymes that can tranform pyruvate and other carbon molecules to other metabolites at fairly similar ability as the values of enrichment belong in the same range. However, since there is a greater number of bacteria found in fermented fish products that have higher enrichment values, there are more fish sauce bacteria that have higher ability in transforming pyruvate and energy storing carbon to different metabolites.

In fructose/mannose metabolism, the values are low for enrichment among both populations, the distributions are similar. This suggests that the enzymes that these bacteria do have for fructose/mannose metabolism most likely are the enzymes that

overlap in glycolysis (as fructose/mannose are intermediates in glycolysis). Cheese and fish sauce bacteria have similar ability levels in metabolizing fructose/mannose, most likely to glycolytic intermediates.

Major differences appear for the citric acid cycle and nitrogen metabolism's enrichment values. The nitrogen metabolism differences correspond to the cheese vs. fish fermentation difference. Fish is higher in protein content and the bacteria involved are known for its nitrogen fixing capabilities (as B. licheniformis is used commercially to fix nitrogen). For TCA, the differences in enrichment values can be attributed to whether or not a bacteria prefers to exist in aerobic conditions. As most of the bacteria in cheese are lactic acid bacteria, the process to produce energy carriers is anaerobic, utilizing the Cori cycle.

The differences not only appear for bacteria used in cheese vs. fish sauce; but with the genera in each category as

Distribution of Carbon Fixing Enrichment (Fish)

Distribution of Nitrogen Fixing Enrichment (Fish)

Distribution of Carbon Fixing Enrichment (Cheese)

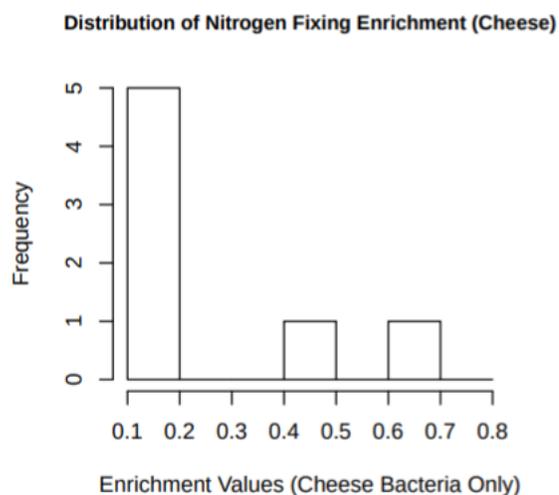
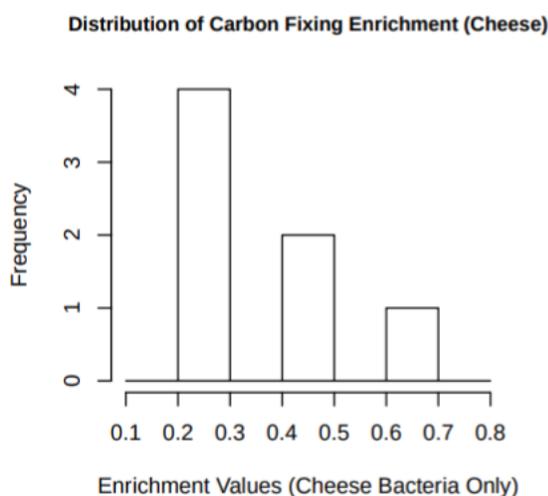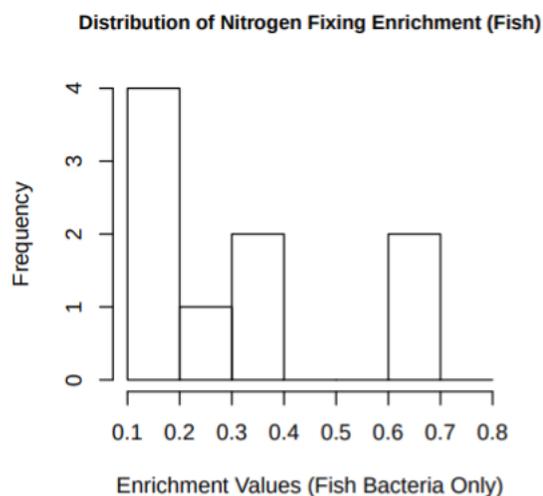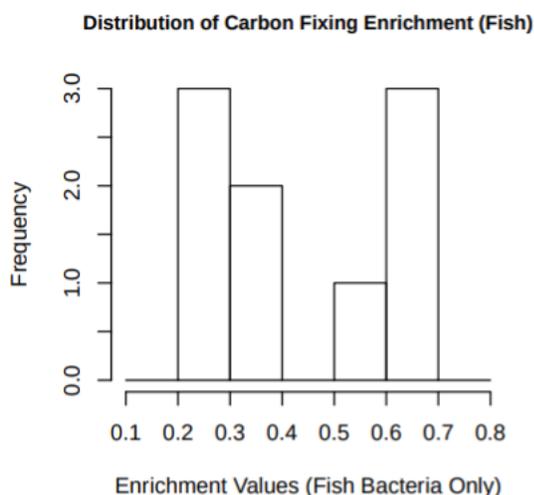**Distribution of Nitrogen Fixing Enrichment (Cheese)**

Fig. 6. Distribution of enrichment values of bacteria found in fish sauce versus that of milk fermentation in carbon metabolism.

Fig. 7. Distribution of enrichment values of bacteria found in fish sauce versus that of milk fermentation in nitrogen metabolism.

well. Leuconostoc and lactobacillus overall tended to hold lower number enzymes for a pathway while Lactococcus and Bacillus tended to hold the higher values (Table 1). The low values in the citric acid cycle was expected for the cheese fermenting bacteria as those species rely on the anaerobic Cori cycle to produce energy more than the citric acid cycle.

E. coli and Staph. aureus having high values for all metabolic pathways corresponds to the fact that they are bacteria often involved in spoilage. The high enrichment values for each pathway indicate the high flexibility of these bacteria, as these species are found everywhere versus that of other lactic acid bacteria which are found in anaerobic conditions like the microflora of cheese fermentation or that of bacteria found in fish sauce which require higher salinity. The rate at which the food degrades increases because of the high ability to degrade food product due to these bacteria.

What was not expected to have high values for all metabolic

pathways was Bacillus licheniformis as described by Table 1. But after reviewing the KEGG description of this species, the result makes sense given that the bacteria has high commercial use to fix nitrogen and other metabolites.

B. Limitations

KEGG REST API was difficult to use especially when trying to work with the BRITE hierarchy. Reflecting upon the literature, the aim is to mirror the proteolytic and biopeptidolytic discovery in the literature and suggest other possible aspects of metabolisms that contribute to the health effects of fermented food items.

The result could be the result of research/curation bias, as high values were found for E. coli, S. aureus, and B, licheniformis. These bacteria are highly studied due to the fact that they are found in many places and utilized as models in research or as commercial products. Conversely, Leuconostoc

species have lower values because of the lack of bias towards them.

C. Future Research

Further research requires utilizing other databases to validate the conclusion drawn. Utilizing sequencing, one can find homologous sequences to the enzymes to verify whether or not a bacteria has specific metabolite for metabolisms. A graphical database of fermenting microorgansims could be developed in order to make novel discovery of metabolic differences in the zymological processing of food.

Further research can go into analyzing the metabolic processes that breakdown food in the microbiome of the human gut and make comparisons there. The framework of this methodology has been built so that this research can be expanded in that direction. Further research requires looking into amino acid metabolisms specifically protein degradation. A focus on protein degradation that results in bioactive compounds or volatile compounds that are produced could be explored. Connecting the end products of a process would show how a specific organism operates.

VI. Conclusion

This study explored the metabolic differences in the fermentation of food products utilizing a framework to pull curated data entries to verify whether or not a species of bacteria had an enzyme for a particular pathway. Understanding how the processes works will help future researchers discover the root at the health benefits of such food items and utilize this information to manufacture more desired food products. This preliminary study can be expanded to a greater number of pathways and species to give statistically relevant results that correspond to biological relevance.

References

[1] Di Cagno, R., Coda, R., De Angelis, M., & Gobbetti, M. "Exploitation of vegetables and fruits through lactic acid fermentation." Food Microbiology, vol. 33 iss. 1, pp. 1–10. 2003.

[2] Korhonen, H., & Pihlanto, A. "Bioactive peptides: Production and functionality." International Dairy Journal / Published in Association with the International Dairy Federation, 16(9), 945–960. 2006.

[3] McSweeny, P. "Biochemistry of cheese ripening." Society of Dairy Technology. 2004.

[4] Nielsen, S. D., Beverly, R. L., Qu, Y., & Dallas, D. C. "Milk bioactive peptide database: A comprehensive database of milk protein-derived bioactive peptides and novel visualization." Food Chemistry, 232, 673–682. 2017.

[5] Nuraida, L. "A review: Health promoting lactic acid bacteria in traditional Indonesian fermented foods." Food Science and Human Wellness, 4(2), 47–55. 2015.

[6] Parvez, S., Malik, K. A., Ah Kang, S., & Kim, H.Y. "Probiotics and their fermented food products are beneficial for health." Journal of Applied Microbiology, 100(6), 1171–1185. 2006.

[7] Sánchez-Rivera, L., Martínez-Maqueda, D., Cruz-Huerta, E., Miralles, B., & Recio, I. "Peptidomics for discovery, bioavailability and monitoring of dairy bioactive peptides." Food Research International , 63, 170–181. 2014.

[8] Shivanne Gowda, S. G., Narayan, B., & Gopal, S. "Bacteriological properties and health-related biochemical components of fermented fish sauce: An overview." Food Reviews International, 32(2), 203–229. 2016.