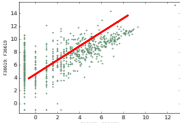


MOTIVATION and BACKGROUND

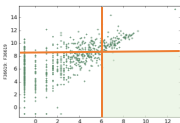
- Microbes have been shown to influence various diseases, such as cancer, diabetes, and obesity
- Communities of microbes and other microorganisms that live in a particular habitats called **microbiome**
- No studies have focused on finding fundamental rules (invariants) between microbes using Boolean Analysis
- Determination of fundamental rules can lead to clinical utility such as cancer diagnosis

PREVIOUS ANALYSIS:



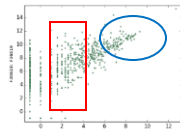
Weaker correlation using Pearson's Correlation Coefficient

OUR ANALYSIS:



Boolean Relationship:
A high \rightarrow B high

FUTURE DISEASE ANALYSIS:



Using Boolean Analysis:
RED rectangle represents disease
BLUE oval represents healthy

SOLUTION

We propose using Boolean analysis to comprehensively study large amounts of microbiome data and formulate universal Boolean rules that each microbe community will follow.

To reach the goal, the following steps are taken:

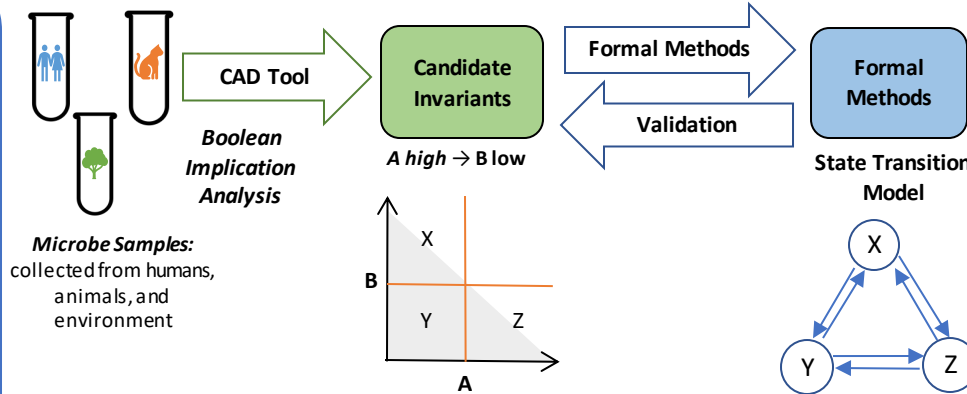
Phase 1 – Collect Data:

1. Collect microbiome data from multiple databases, such as HMP and QITA
2. Convert the data into OTU/ BIOM tables using QIIME2
3. Using frequencies from the tables, plot pairs of microbes on graphs using Hegemon*

Phase 2 – Analyze Data:

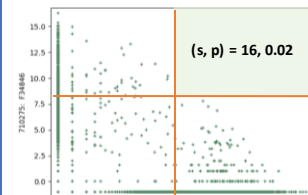
4. Perform Boolean analysis on the generated graphs

CAD APPROACH



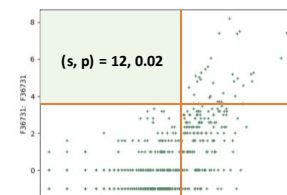
Data Analysis

- Boolean relationships are shown through implications (\rightarrow)
- Boolean relationships are determined by any sparse quadrant(s) – shown by the light green quadrants
- S-numbers denote the randomness and P-numbers denote the error rate
- Let A and B represent arbitrary microbes in the figures below



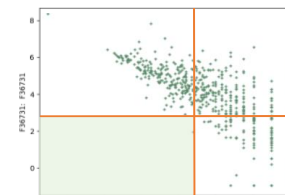
A high \rightarrow B low

When microbe A is highly present, microbe B is low



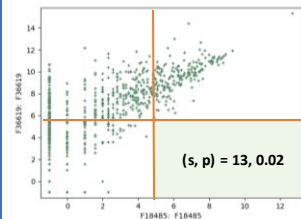
A low \rightarrow B low

When the frequency of microbe A is low, the frequency of microbe B is also low



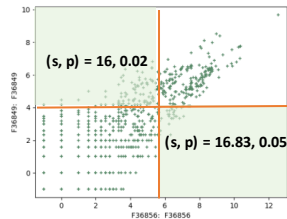
A low \rightarrow B high

When there is a low frequency of microbe A, microbe B is highly present



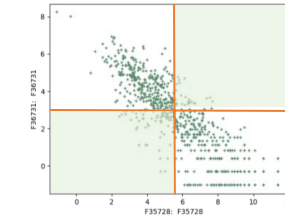
A high \rightarrow B high

Both microbe A and B are highly present at the same time



A Equivalent B

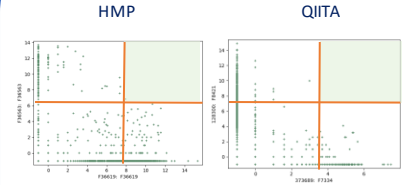
As microbe A increases, so does microbe B



A Opposite B

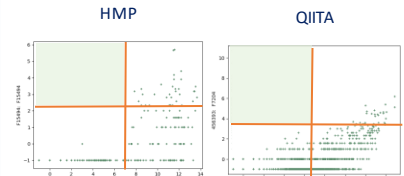
As microbe A decreases, so does microbe B

RESULTS



RULE:

Staphylococcus high \rightarrow Lactobacillus low



RULE:

Lactobacillus y high \rightarrow Lactobacillus x high

DISCUSSION

- Staphylococcus is the most common cause of food poisoning
- Lactobacillus species are known for their Probiotic benefits.
- Lactobacillus inhibits Staphylococcus

FUTURE DIRECTIONS

- Continue searching for more relationships
- Validate the relationship in multiple independent data sets
- Validate the strong relationships in the wet lab
- Explore the clinical utility of these relationships by determining relationships between diseases and microbes

ACKNOWLEDGEMENTS

We would like to thank FMCAD for this opportunity, as well as Dr. Debashis Sahoo and the rest of the Boolean lab for guidance throughout this project.

CITATIONS

- Debashis Sahoo, Dawid L. Dill, Andrew J. Gentles, Rob Tibshirani, Sylvia K. Plehritis. Boolean implication networks derived from large-scale, whole genome microarray datasets. *Genome Biology*, 9:R253, Oct 30 2008
- Piero Dalerba*, Debashis Sahoo*, Soomyung Park, Xiangjian Guo, Greg Youthers, Nan Song, Nate Wilcox-Fogel, Erna Forgo, Pradeep S. Rajendran, Stephen P. Miranda, Shigeo Hisamori, Jacqueline Hutchison, Torger Kalsky, Daling Qian, No-riem Wolmark, George A. Fisher, Matt van der Wijk, and Michael F. Clarke. CD22 is a Prognostic Biomarker in Stage II and Stage III Colon Cancer. *N Engl J Med*. 2016 Jan 21;374(3):212-22. doi: 10.1056/NEJMoa1506597.