

# Chaos in small microbial communities

Behzad D. Karkaria<sup>1</sup> Kiran R. Patil<sup>1</sup>

<sup>1</sup>MRC Tox

\*To whom correspondence should be addressed;

E-mail:

## Abstract

Genome scale metabolic model reconstructions are relatively poor at predicting species viability in specified media. Methods such as gap-filling, require experimental data to identify which metabolic reactions should be added to achieve species growth. However, this therefore requires *in vitro* screening of each species in different media environments. We have developed a machine learning pipeline that can predict the viability of a microbial species in a given defined media. We propose that this pipeline can be used to perform automatic gap-filling, to improve the output of existing automated metabolic reconstruction frameworks.

## 1 Introduction

...

## 2 Results

GSMM reconstruction pipelines such as metaGEM, take the genome sequence of a species and use databases of annotated genes to construct a network of reactions that describe the organisms metabolism. Incomplete genome sequences and incomplete annotation databases result in missing reactions (gaps). Gapfilling methods take a defined media that the species is known to grow in, and adds reactions to the model that would enable such growth.

When performing a large number of genome reconstructions, gapfilling is generally performed using a "complete media". T

Figure x shows a dendrogram of species, coupled with a heatmap showing their growth rates in different defined media. The dendrogram is created by hierarchical clustering, using reactions in the GSMMs as features. The heatmap shows patterns of growth between related organisms. This figure provides a good indication that GSMMs include features that can be used in machine learning techniques to predict growth.

### **3 Methods**

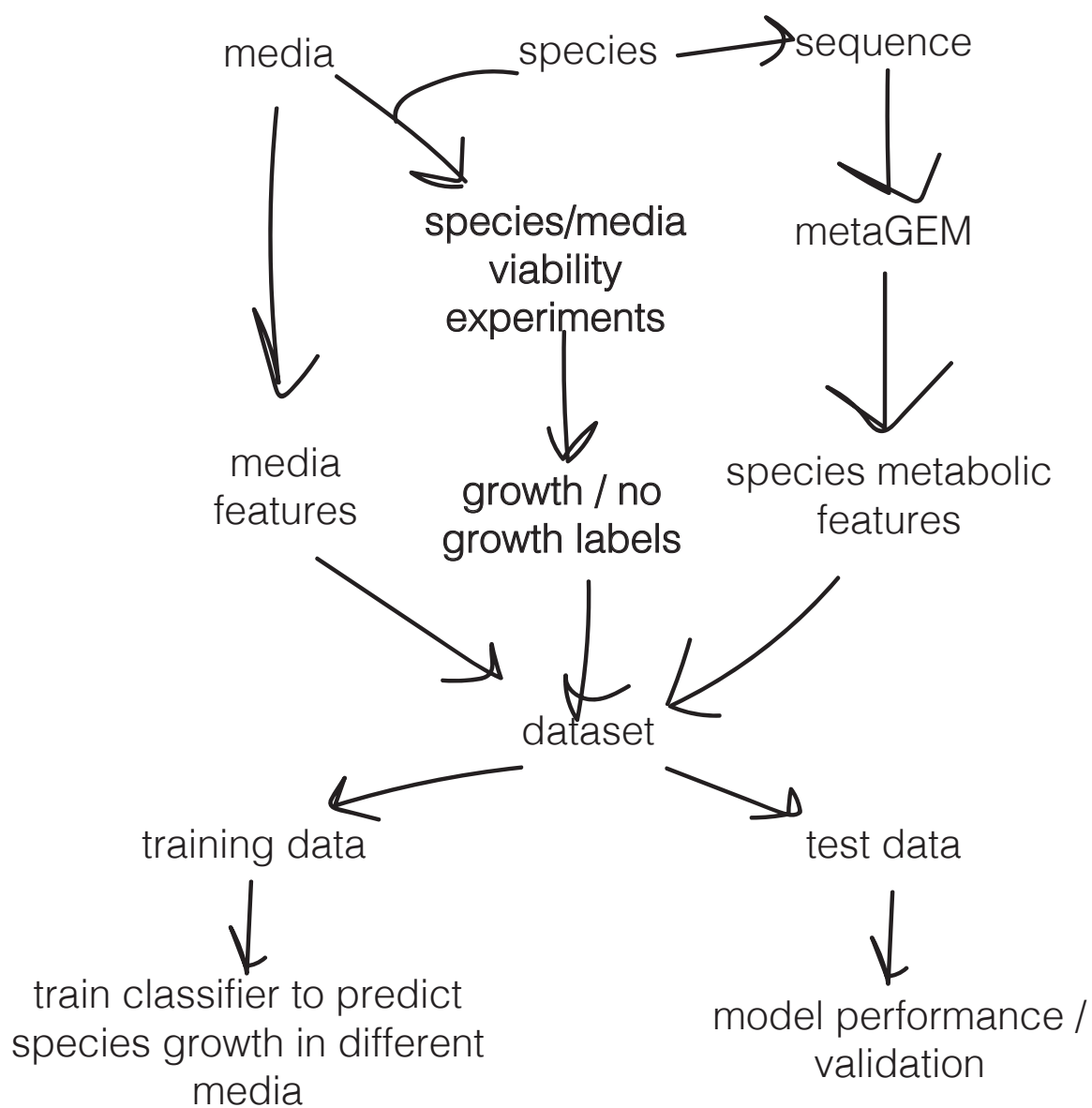
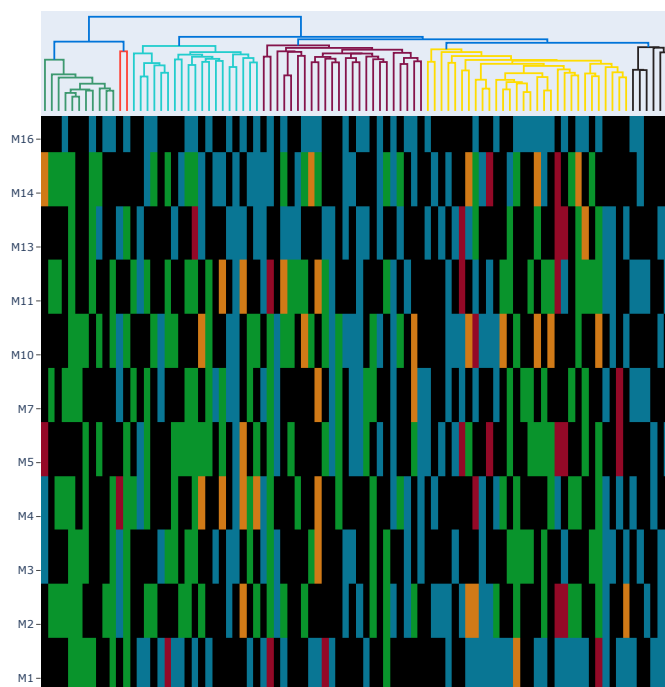
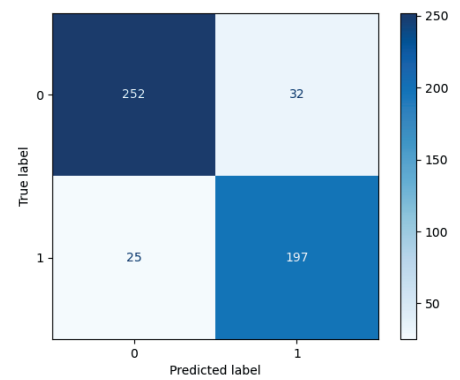
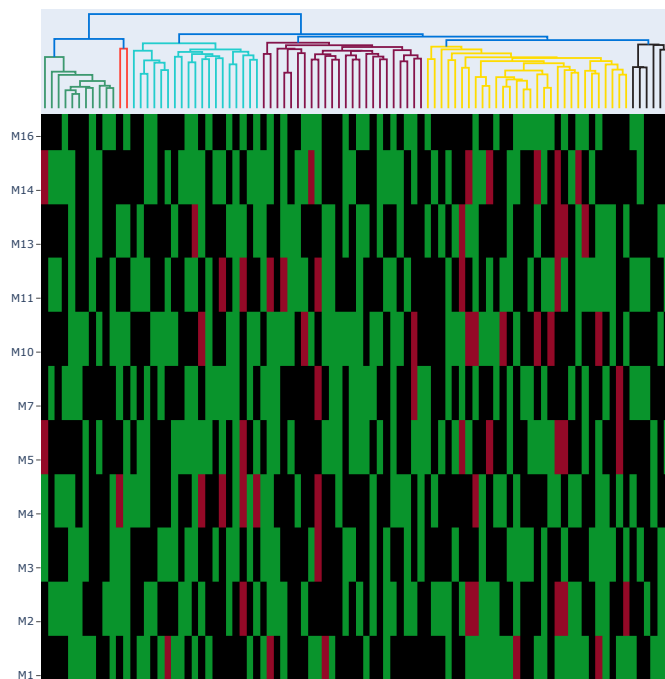


Figure 1:





■ Training data

Correct classifications:

■ True positive

■ True negative

Incorrect classifications:

■ False positive

■ False negative