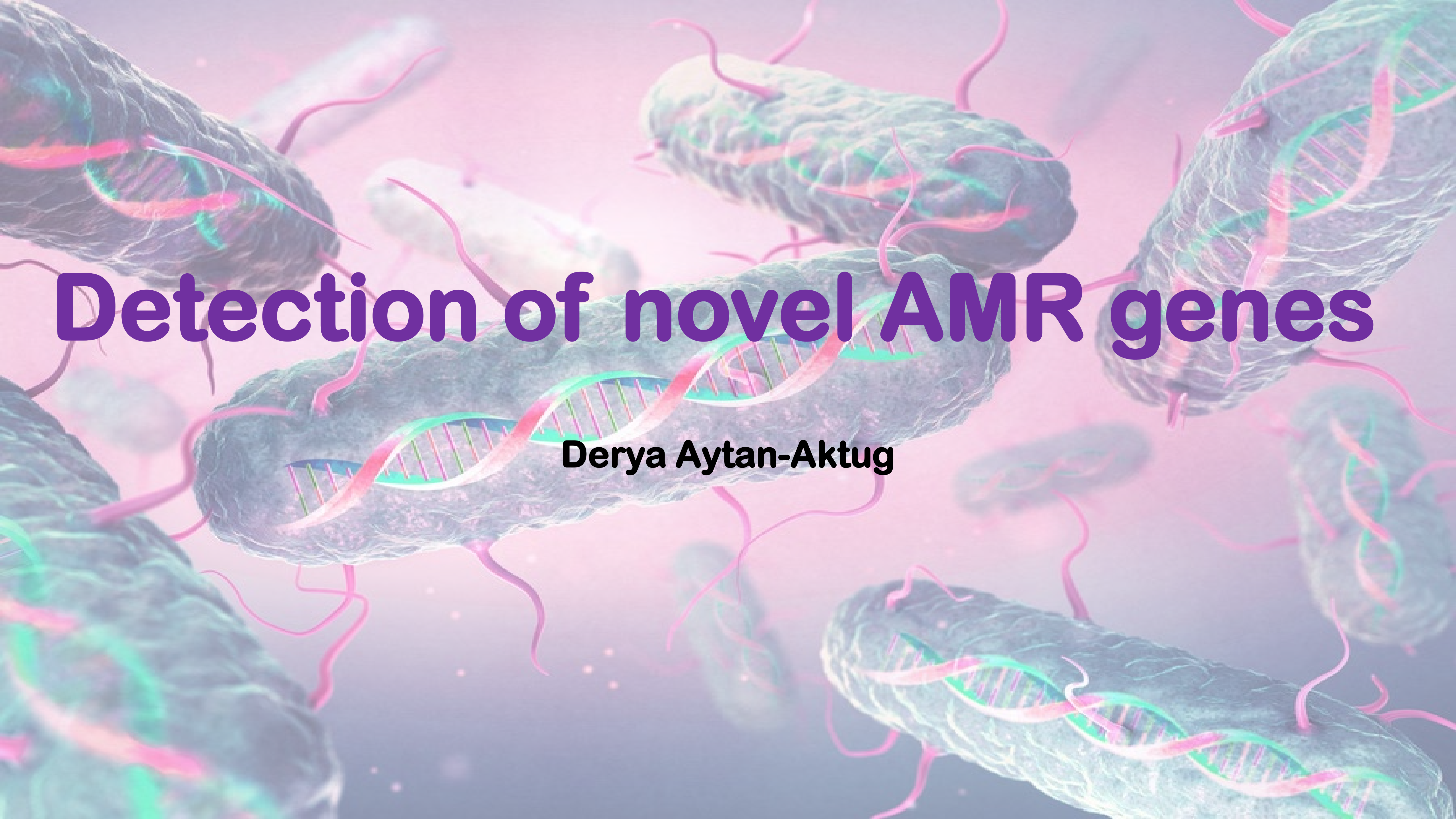
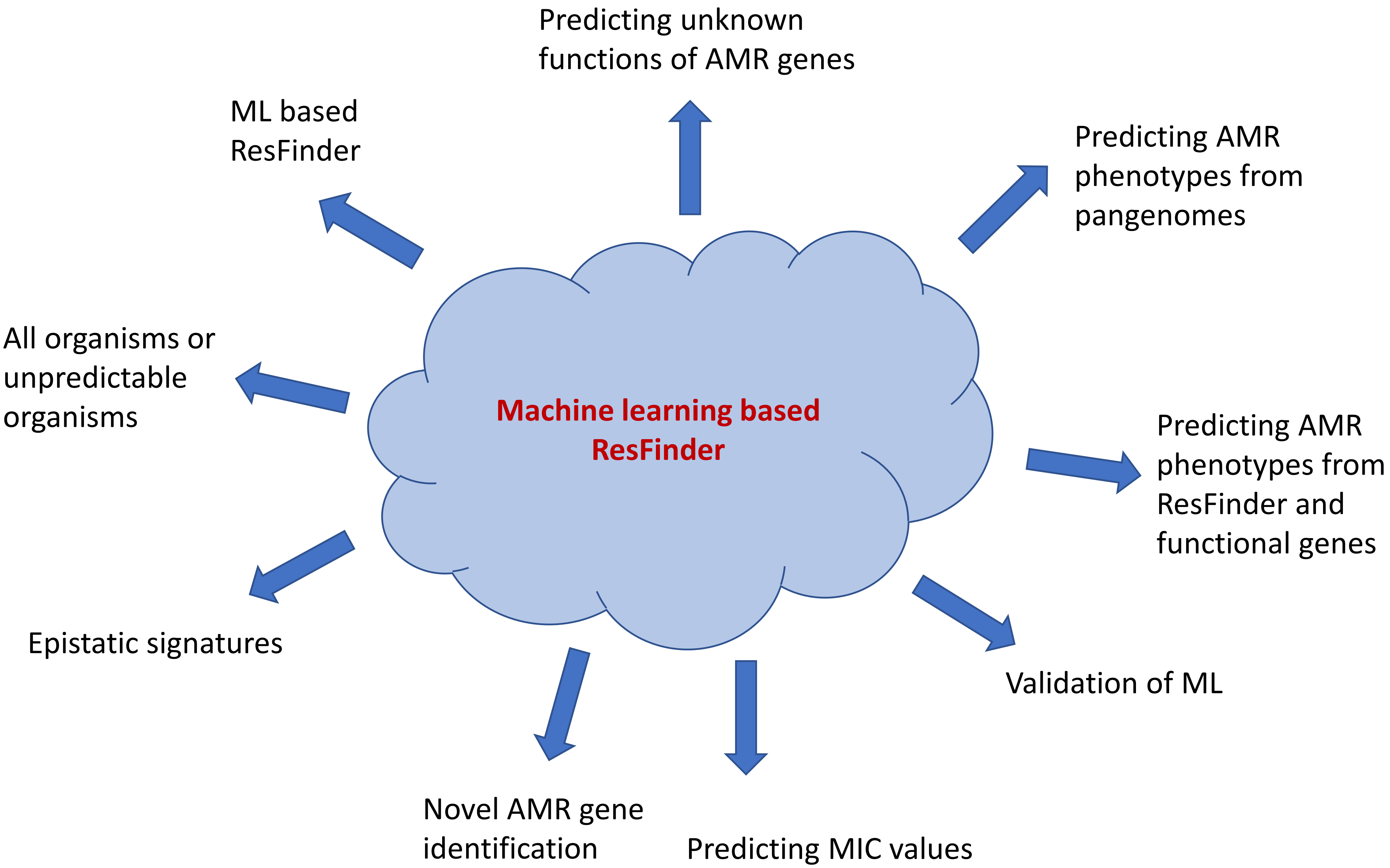


Detection of novel AMR genes

Derya Aytan-Aktug



Machine learning based ResFinder

Predicting unknown functions of AMR genes

Predicting AMR phenotypes from pangenomes

Predicting AMR phenotypes from ResFinder and functional genes

Validation of ML

Predicting MIC values

Novel AMR gene identification

Epistatic signatures

All organisms or unpredictable organisms

ML based ResFinder

Project-1: Detection of novel AMR genes

Targeted organism(s): Acinetobacter baumannii and Pseudomonas aeruginosa

Features: ResFinder genes, CARD genes, functional genes, pangenomes, variations

How good can we predict AMR with known resistance genes?

Is there a room for novel genes?

Validation?

Project-1: Detection of novel AMR genes

Targeted organism(s): Acinetobacter baumannii and Pseudomonas aeruginosa

Features: ResFinder genes, CARD genes, functional genes, pangenomes, variations

	Gene-1	Gene-2	Gene-3
Isolate-1	coverage	coverage	coverage
Isolate-2	coverage	coverage	coverage
…	…	…	…

How good can we predict AMR with known resistan

Is there a room for novel genes?

Validation?

Project-1: Detection of novel AMR genes

Targeted organism(s): Acinetobacter baumannii and Pseudomonas aeruginosa

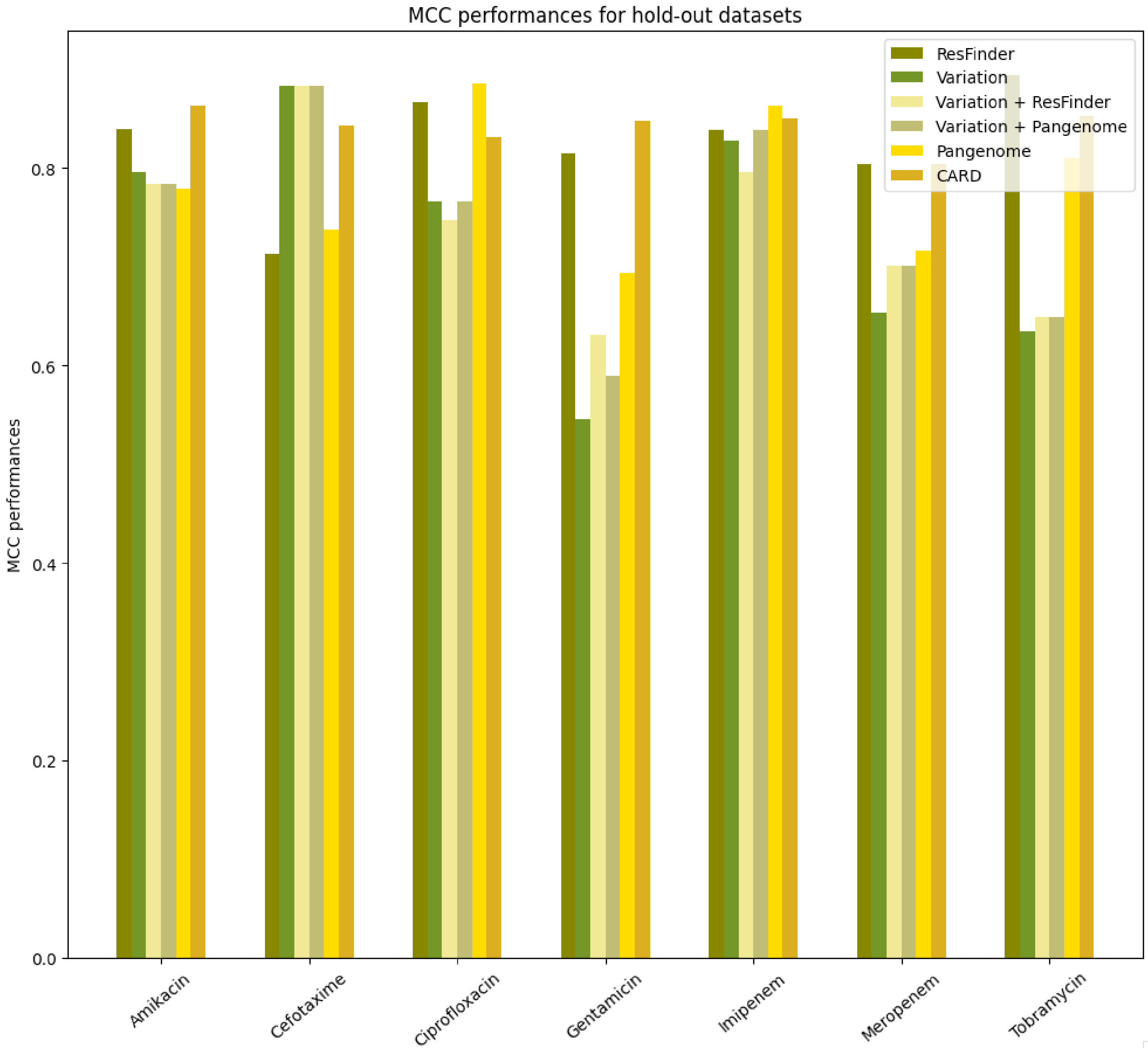
Features: ResFinder genes, CARD genes, functional genes, pangenomes, variations

How good can we predict AMR with known resistance genes?

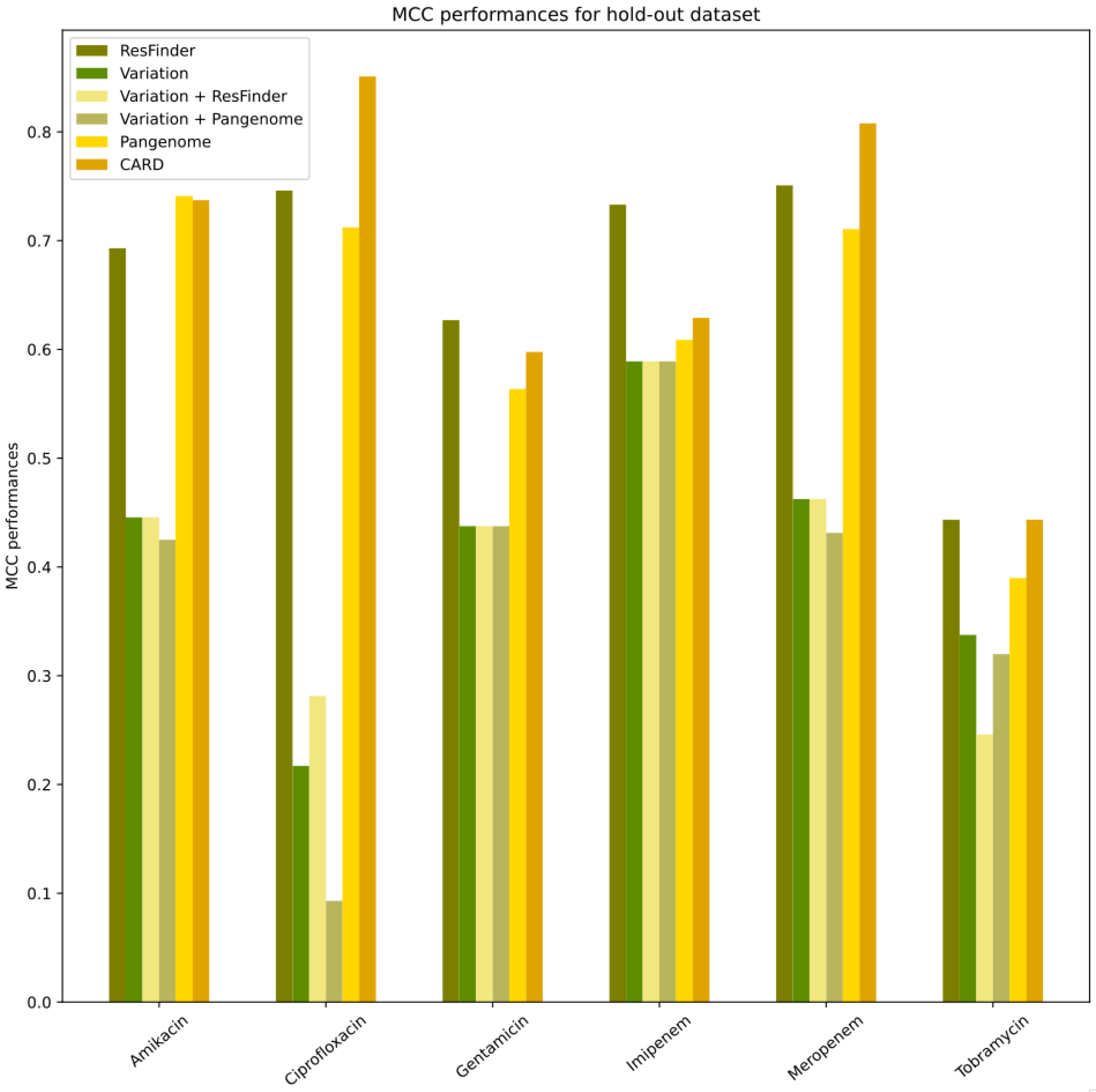
Is there a room for novel genes?

Validation?

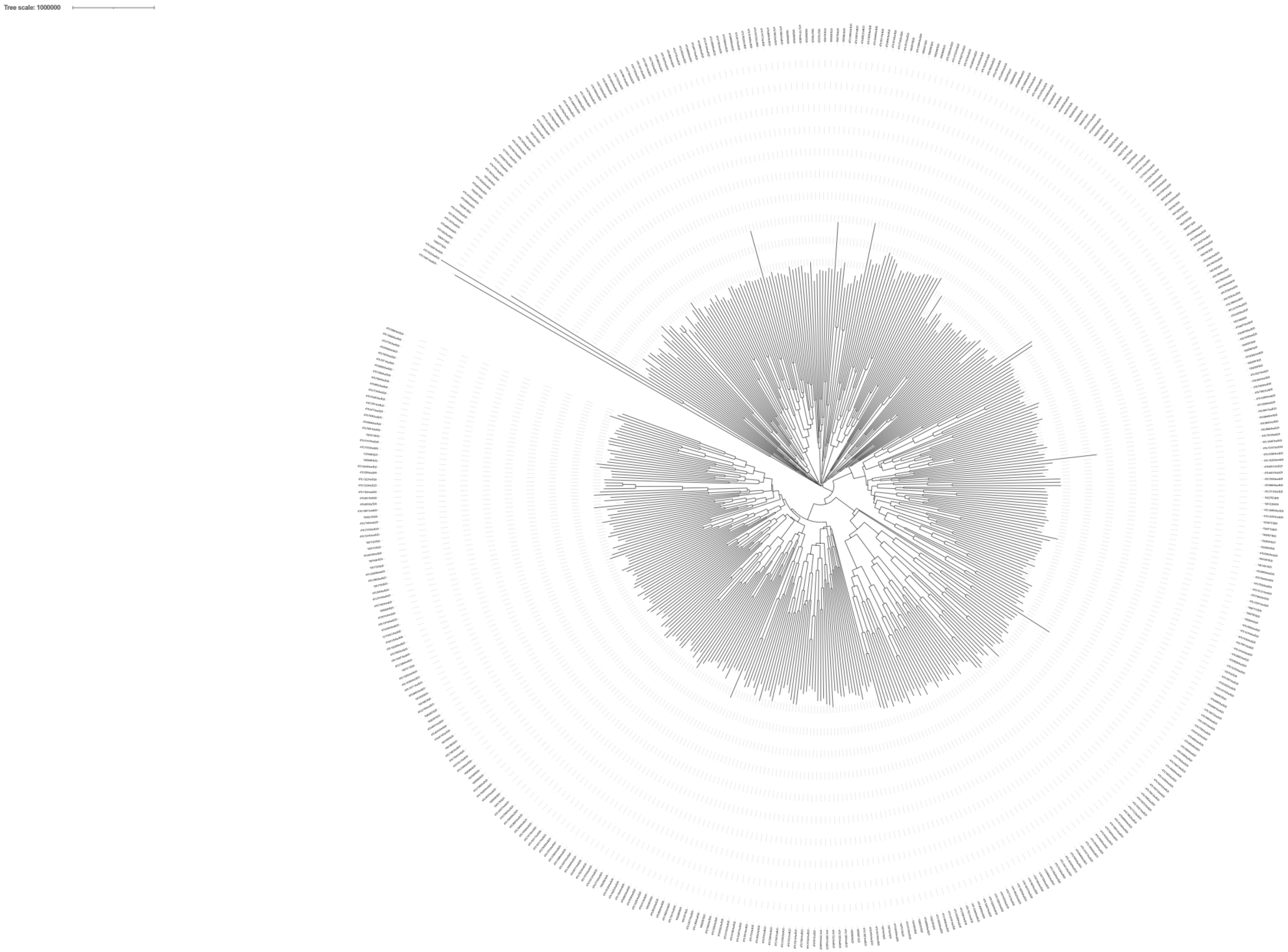
How much can we explain with known resistance genes?



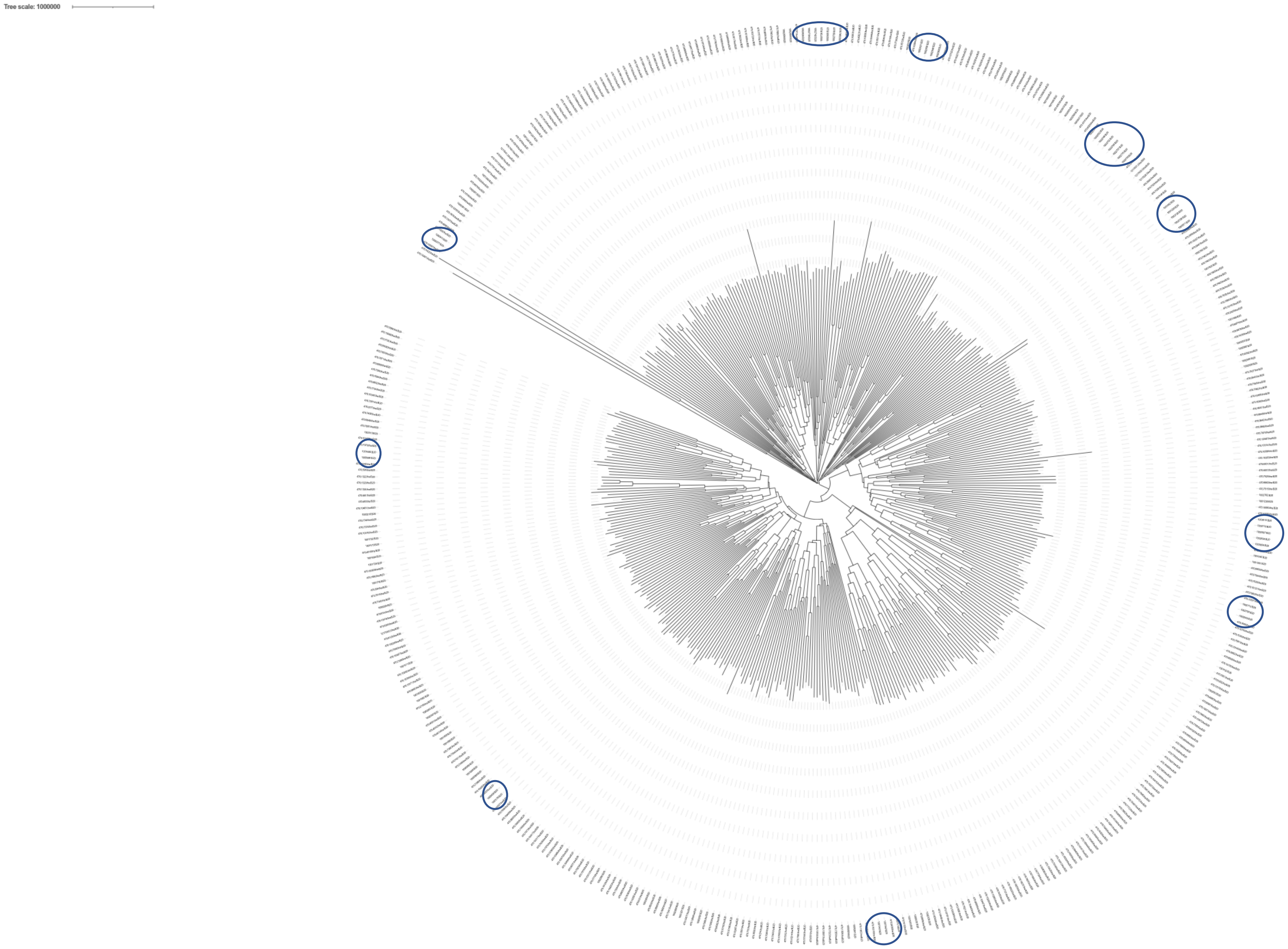
Validation with TWIW isolates



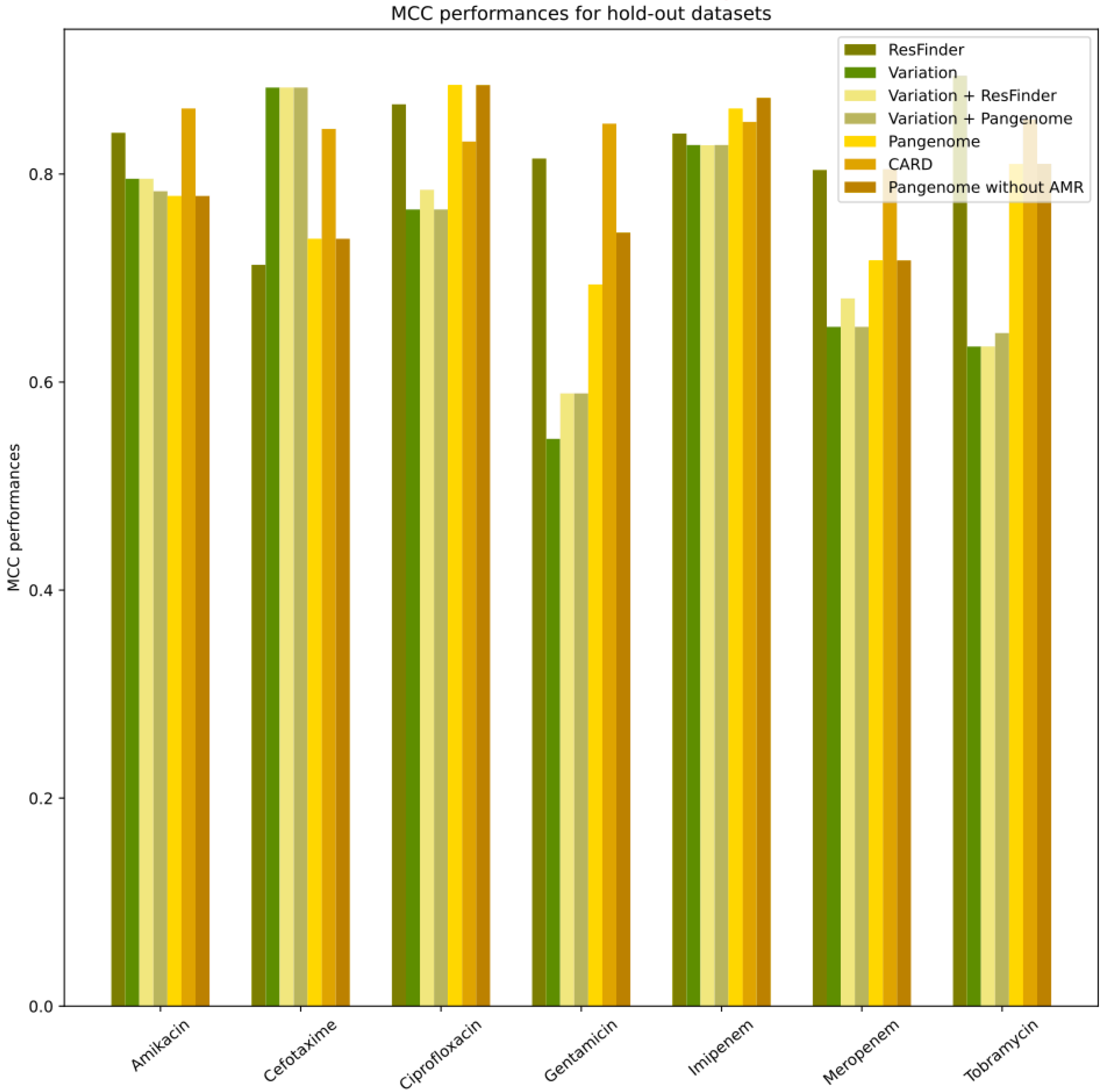
Distances between hold-out and validation data



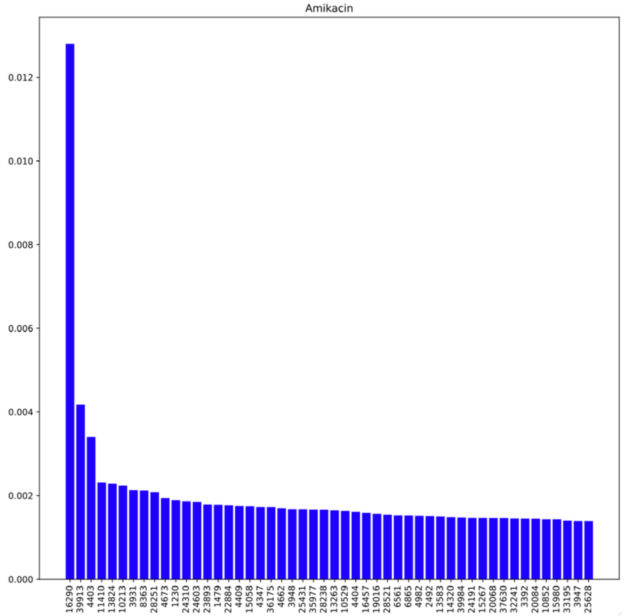
Distances between hold-out and validation data

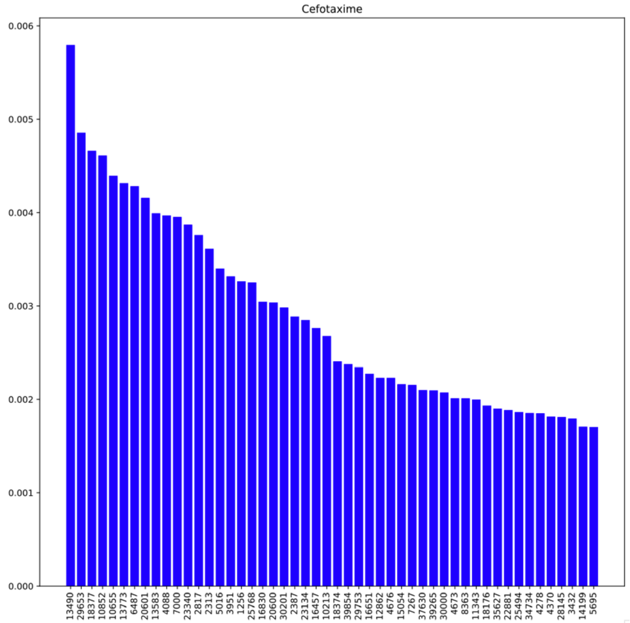


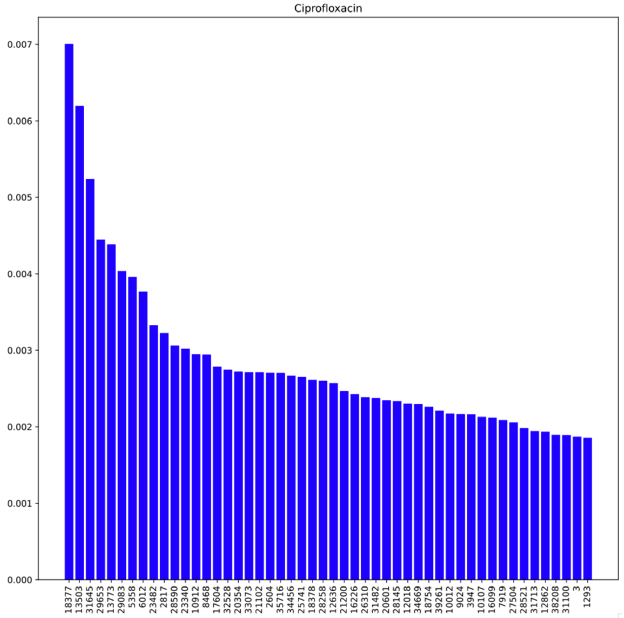
If we remove known resistance genes from pangenome?

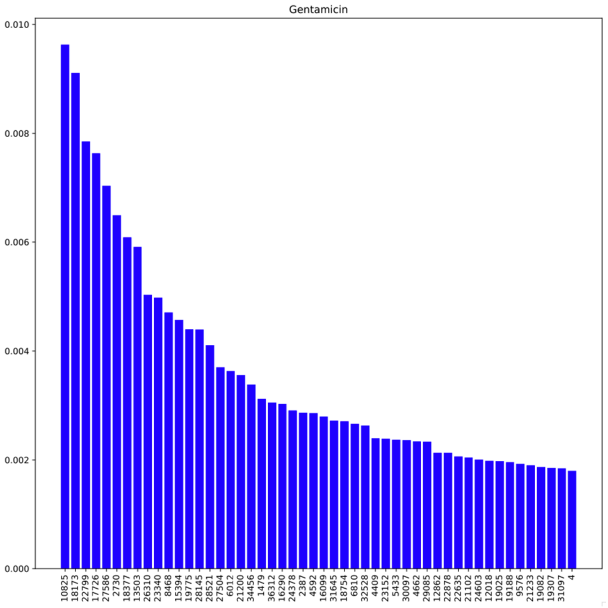


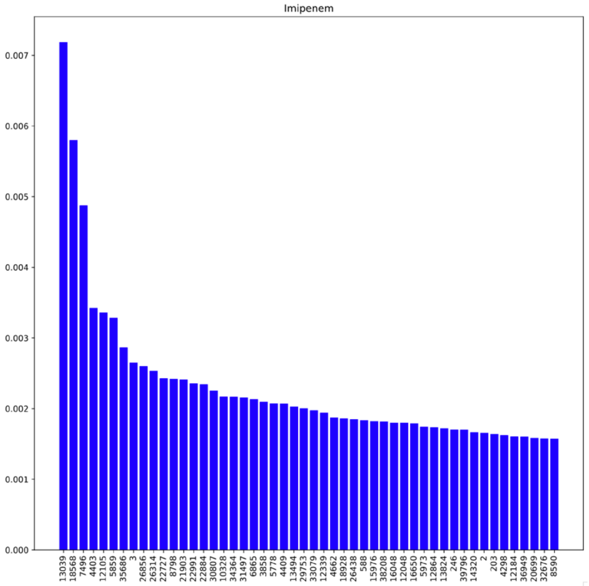
Which features were found important including AMR genes?

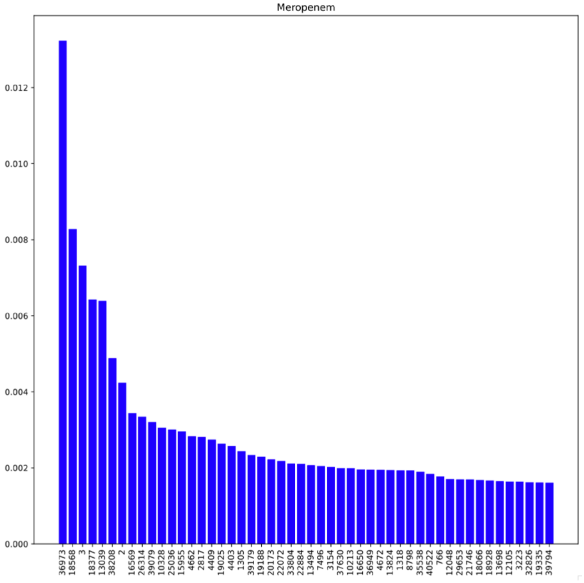


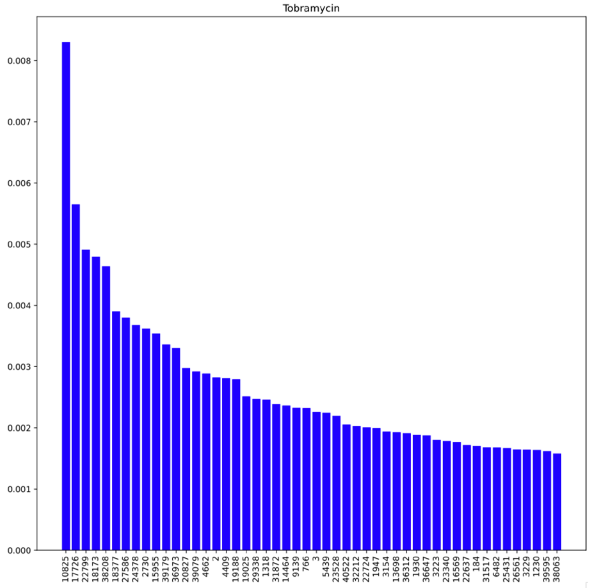




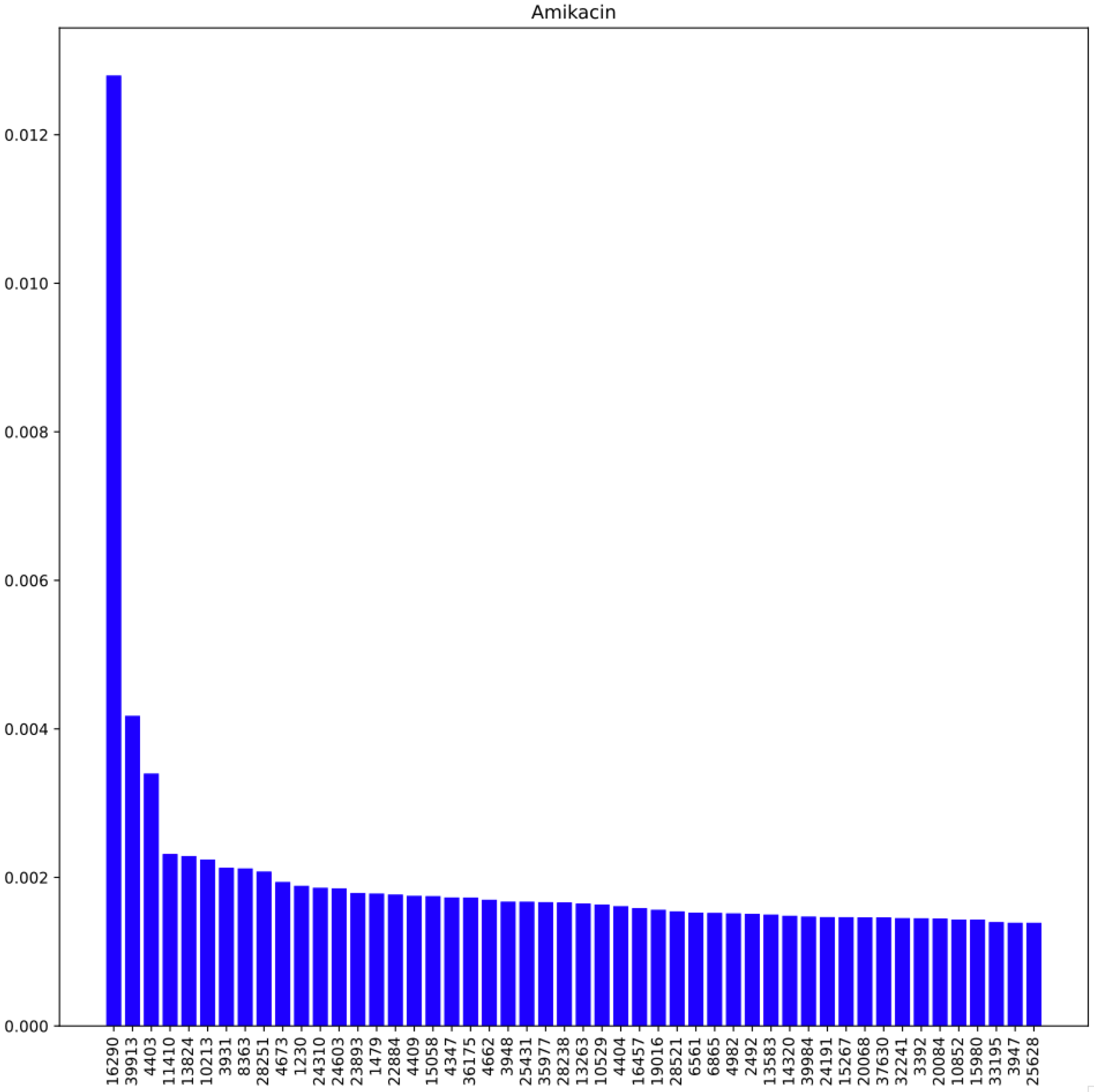


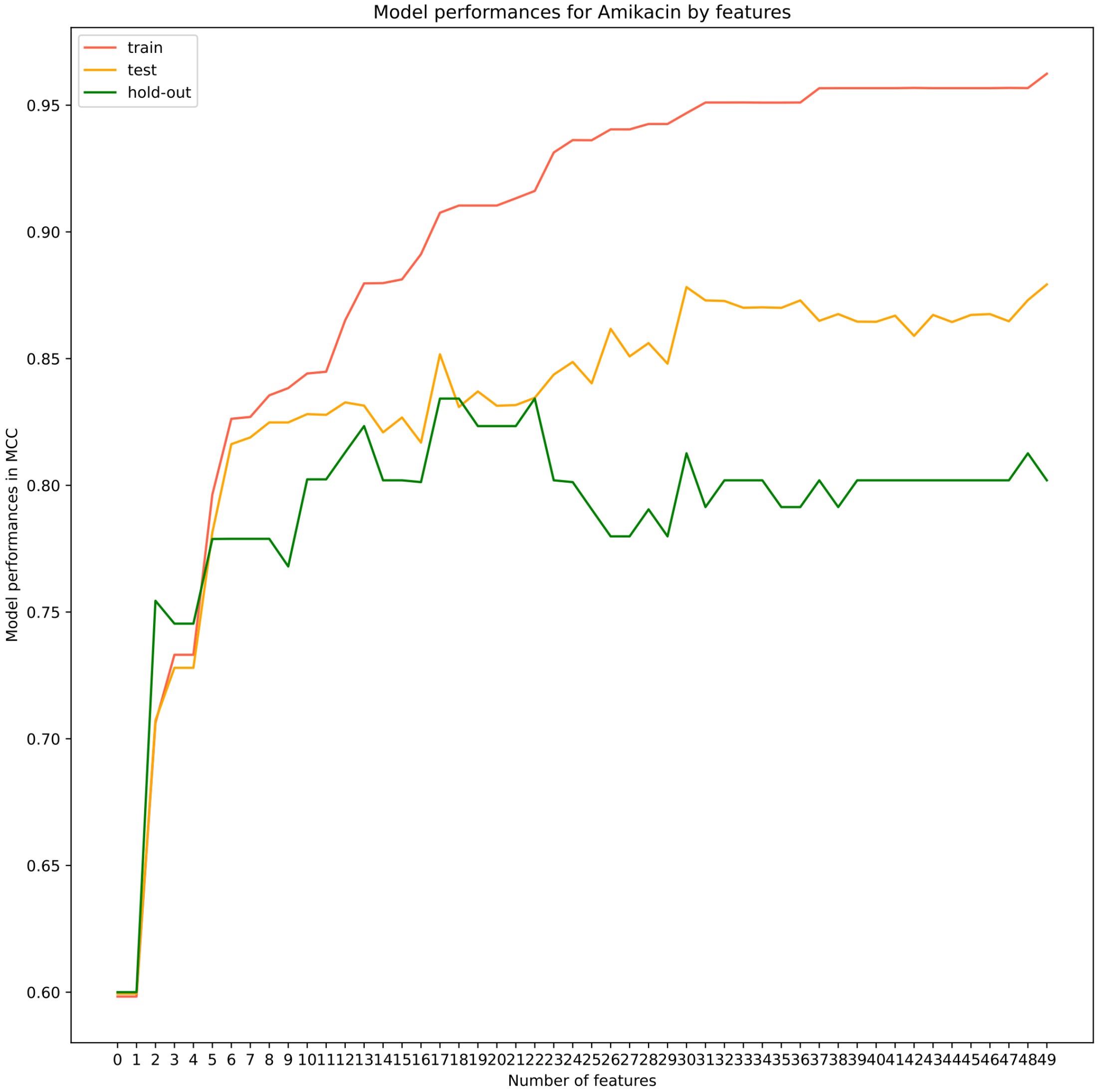




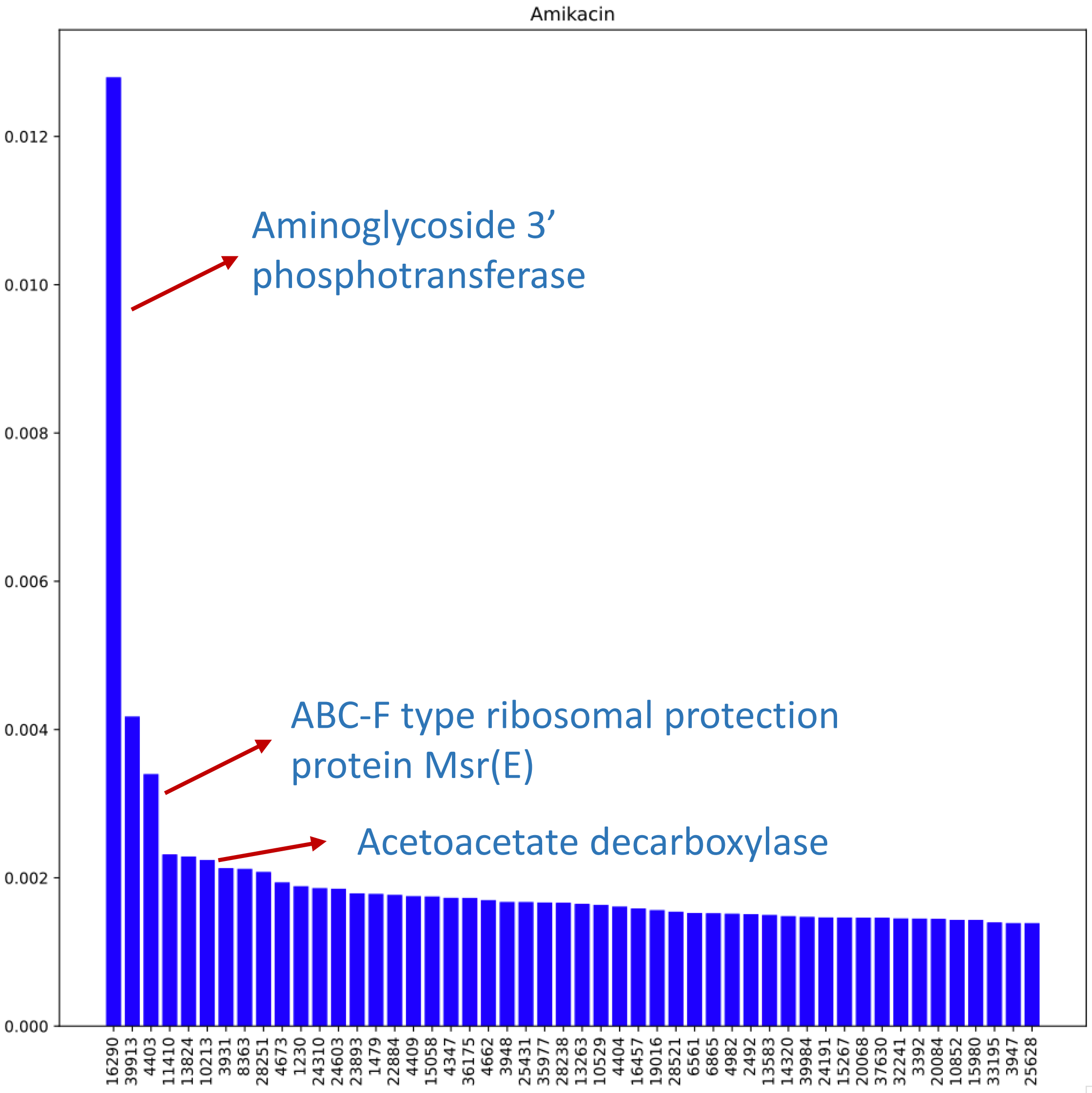


Model performances feature by feature





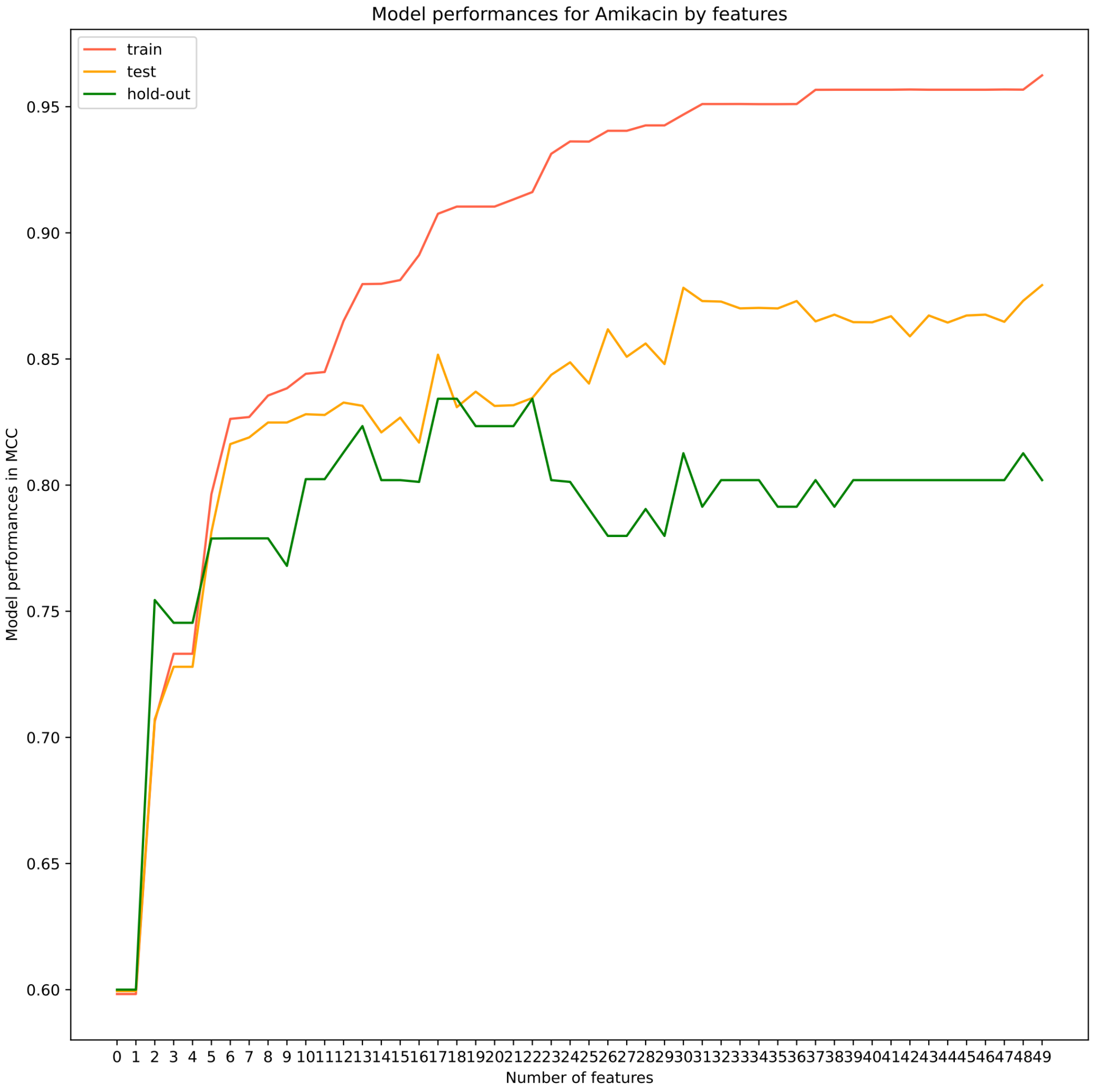
Model performances feature by feature



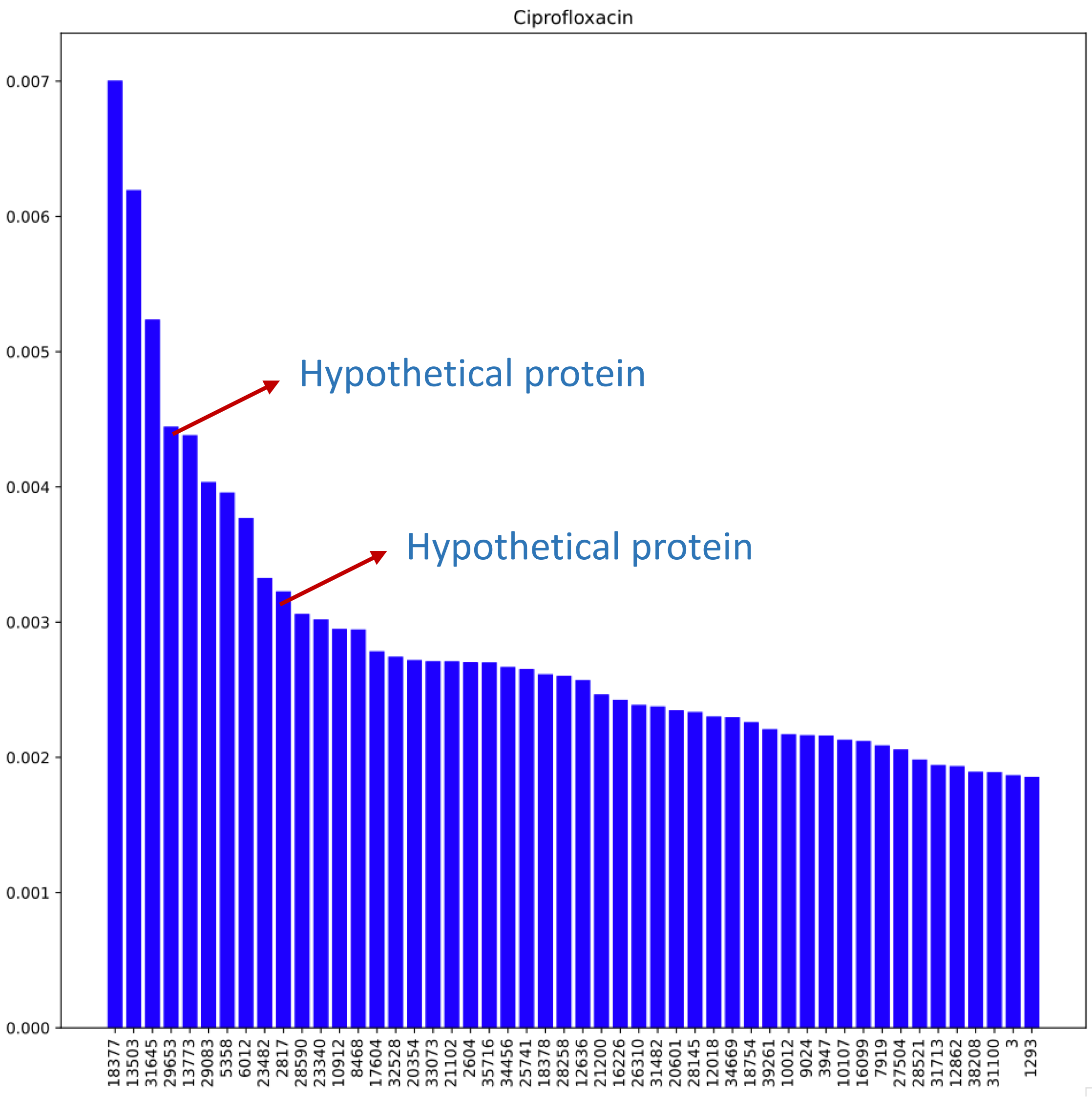
Aminoglycoside 3’ phosphotransferase

ABC-F type ribosomal protection protein Msr(E)

Acetoacetate decarboxylase

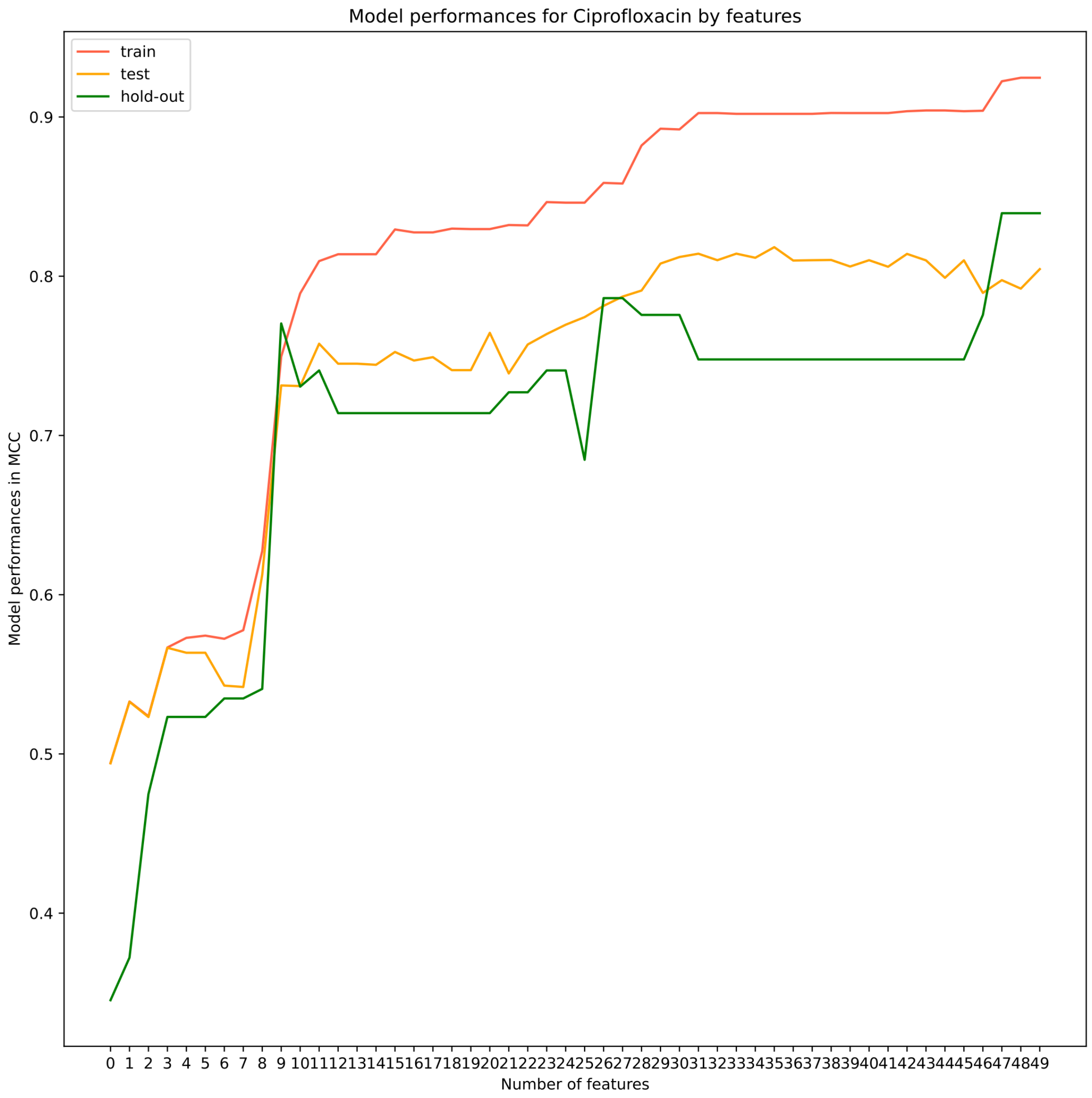


Model performances feature by feature



Hypothetical protein

Hypothetical protein



How prevalent are they in resistance and susceptible isolates of hold-out set?

Antibiotic	Gene product	Nr. Of resistance isolates	Nr. Of resistance isolates with the gene	Nr. Of susceptible isolates	Nr. Of susceptible isolates with gene
Amikacin	ABC-F type ribosomal protection protein Msr(E)	482	124	408	10
Cefotaxime	IS6 family transposase IS15	81	77	111	45
Ciprofloxacin	Hypothetical protein	1435	807	180	11
Ciprofloxacin	Hypothetical protein	1435	1029	180	30
Gentamicin	IS6 family transposase IS15	765	651	84	14
Imipenem	ydaU family protein	436	207	517	37
Meropenem	ATP binding protein	324	172	197	1
Tobramycin	Hypothetical protein	545	103	280	190

How prevalent are they in resistance and susceptible isolates of hold-out set?

Antibiotic	Gene product	Nr. Of resistance isolates	Nr. Of resistance isolates with the gene	Nr. Of susceptible isolates	Nr. Of susceptible isolates with gene
Amikacin	ABC-F type ribosomal protection protein Msr(E)	482	124	408	10
Cefotaxime	IS6 family transposase IS15	81	77	111	45
Ciprofloxacin	Hypothetical protein	1435	807	180	11
Ciprofloxacin	Hypothetical protein	1435	1029	180	30
Gentamicin	IS6 family transposase IS15	765	651	84	14
Imipenem	ydaU family protein	436	207	517	37
Meropenem	ATP binding protein	324	172	197	1
Tobramycin	Hypothetical protein	545	103	280	190

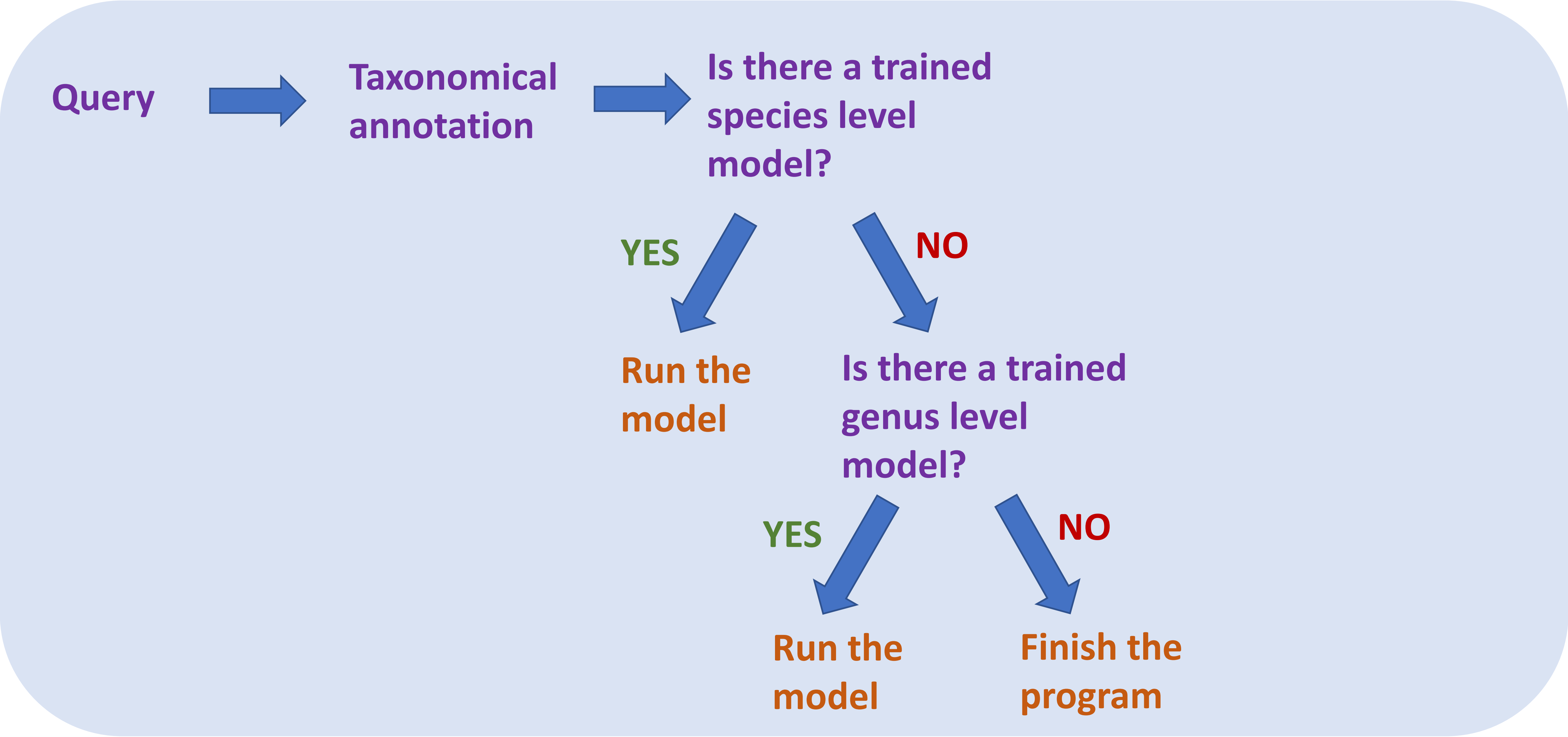
Next steps?

- In vitro validation of the candidate AMR genes
- Writing the manuscript
  - Showcase of detection and validation of AMR genes
- Any other suggestions?

Project-2:ML-based AMR predictors

- ML-based web-servers for predicting AMR (ML-based ResFinder)
  - There is no state of the art
    - Training per species
      - Collecting data/metadata

Pipeline



Query

Taxonomical annotation

Is there a trained species level model?

YES

Run the model

NO

Is there a trained genus level model?

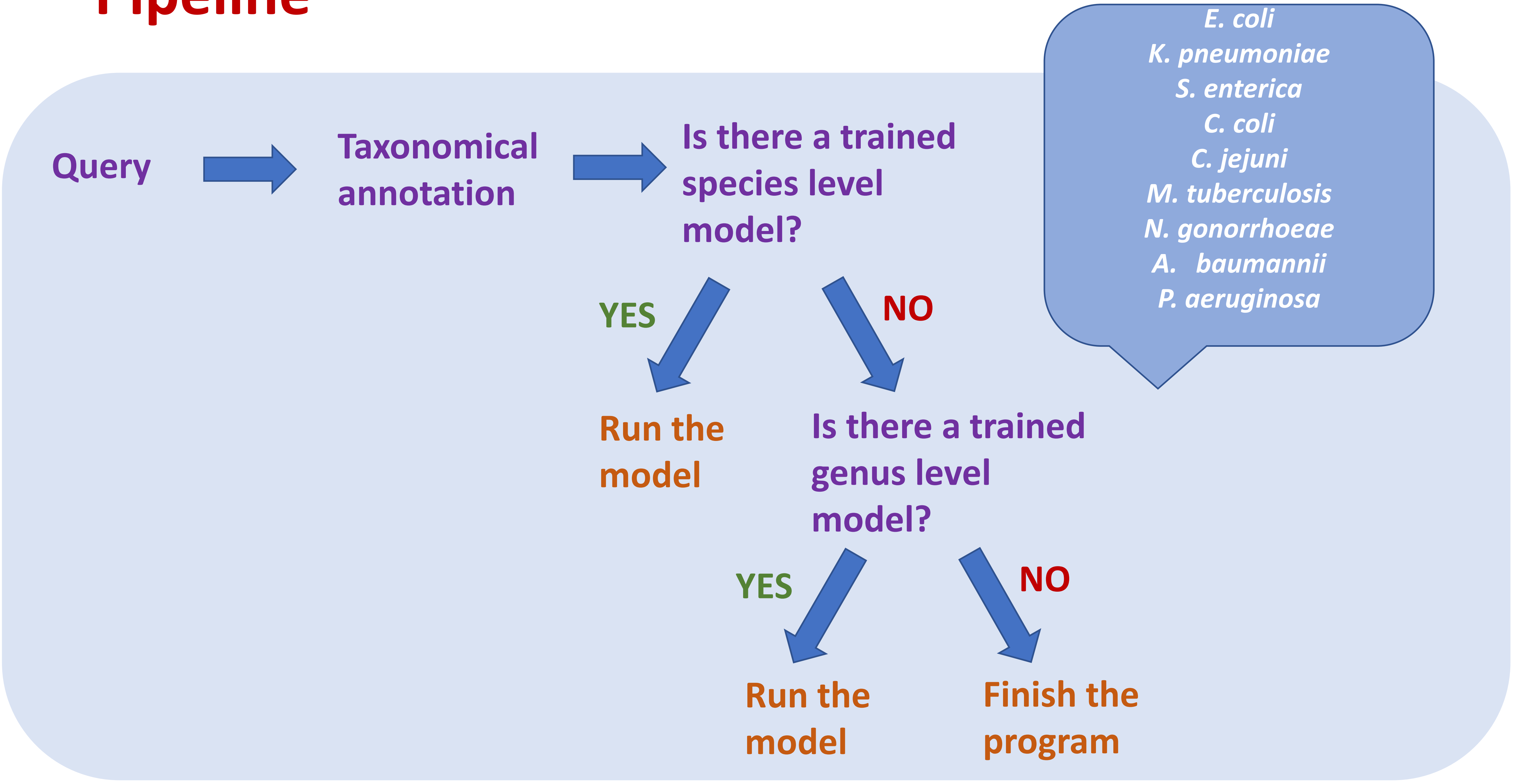
YES

Run the model

NO

Finish the program

Pipeline



Query

Taxonomical annotation

Is there a trained species level model?

E. coli
K. pneumoniae
S. enterica
C. coli
C. jejuni
M. tuberculosis
N. gonorrhoeae
A. baumannii
P. aeruginosa

YES

Run the model

NO

Is there a trained genus level model?

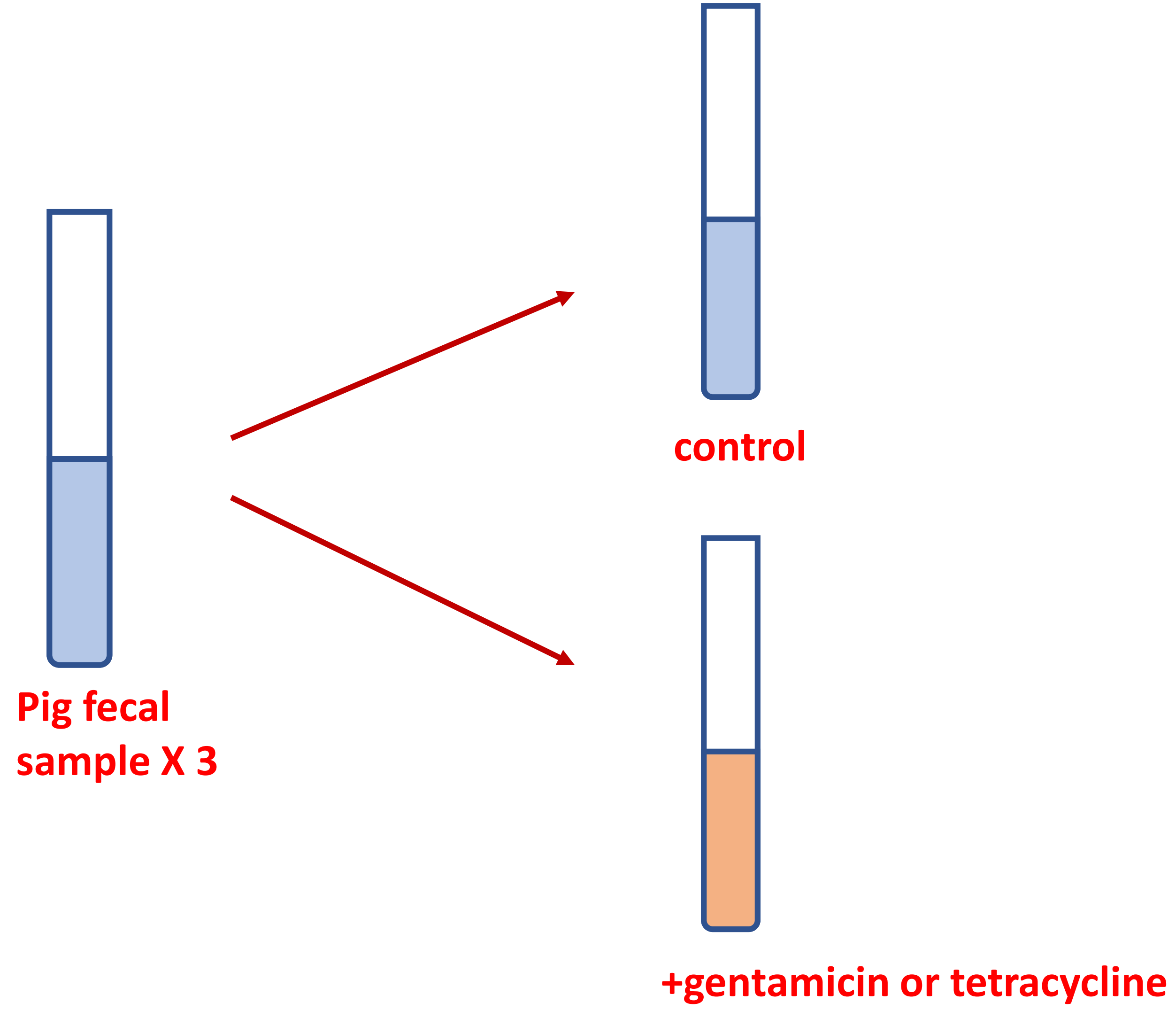
YES

Run the model

NO

Finish the program

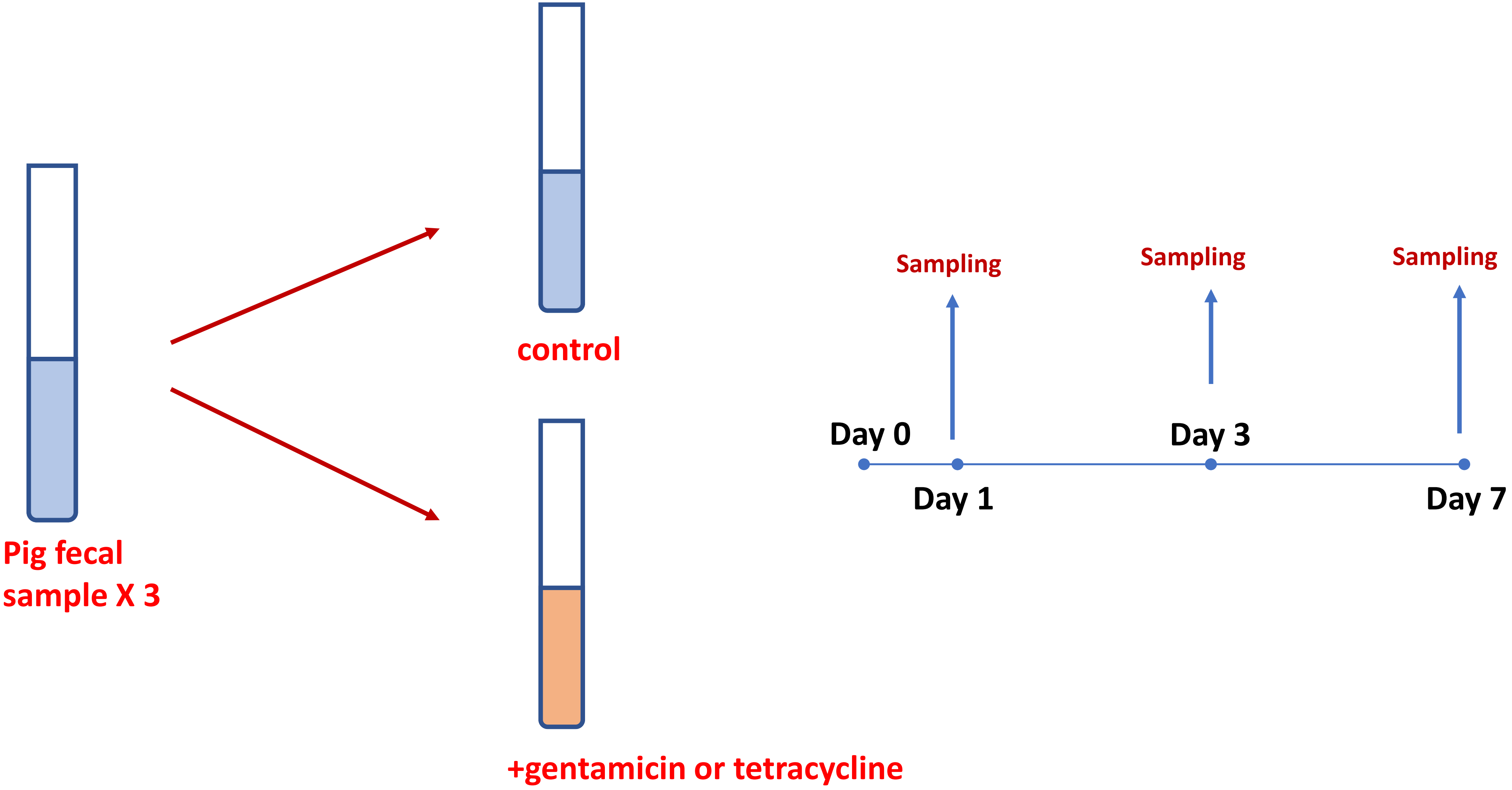
Project-3:Tracking emergence of AMR with metagenomics


Pig fecal
sample X 3

control

+gentamicin or tetracycline

Project-3:Tracking emergence of AMR with metagenomics



Pig fecal sample X 3

control

+gentamicin or tetracycline

Day 0

Day 1

Day 3

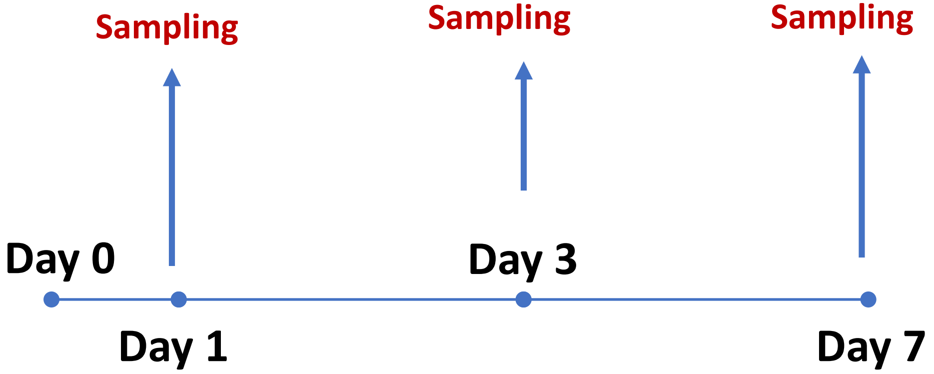
Day 7

Sampling

Sampling

Sampling

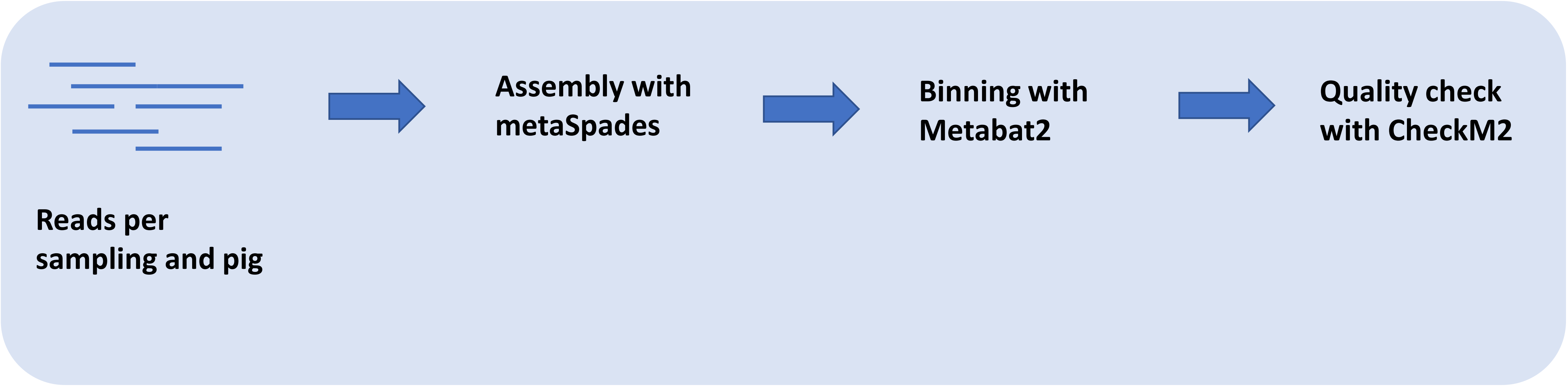
What I want from this project?



- Tracking organisms
  - Which species will take over over time?
- Tracking resistance genes
  - Abundance of resistance genes with and without stress conditions
- Tracking of spontaneous mutations
  - Randomly or logically?
- **Any other suggestions?**

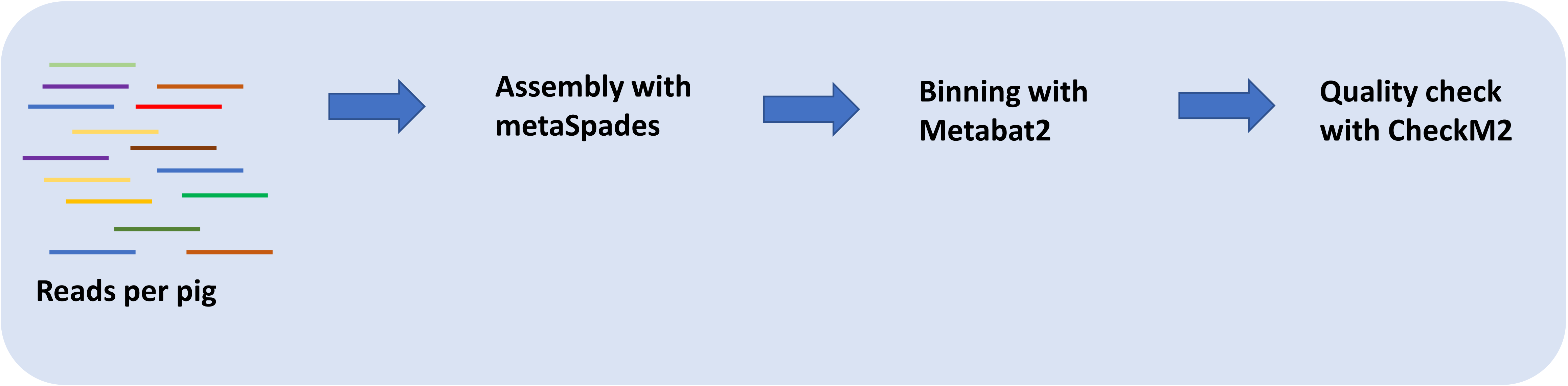
So far…

Individual assemblies



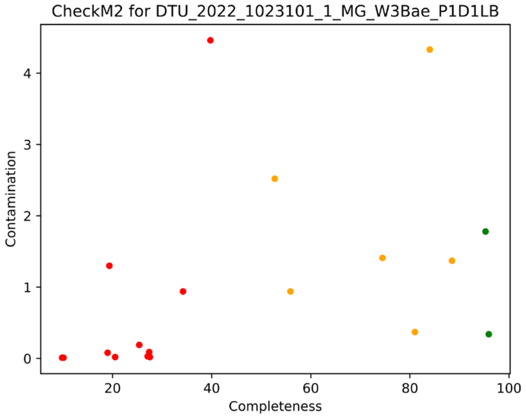
Reads per sampling and pig → Assembly with metaSpades → Binning with Metabat2 → Quality check with CheckM2

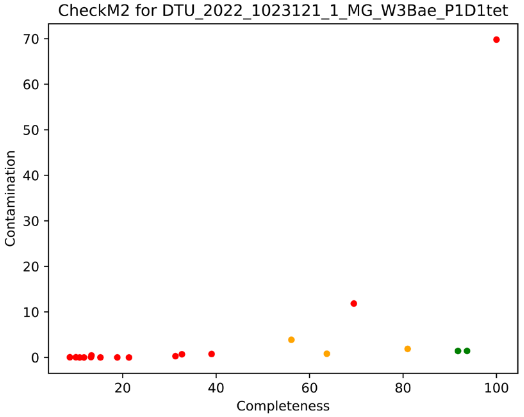
Co-assemblies

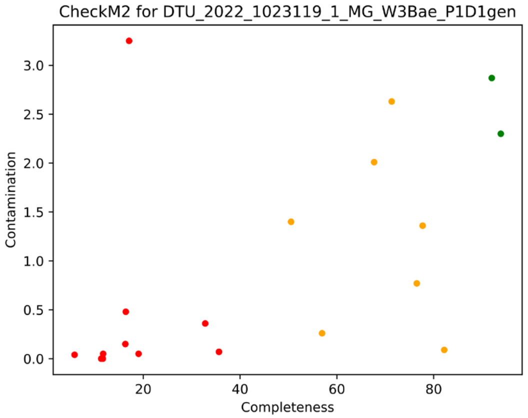


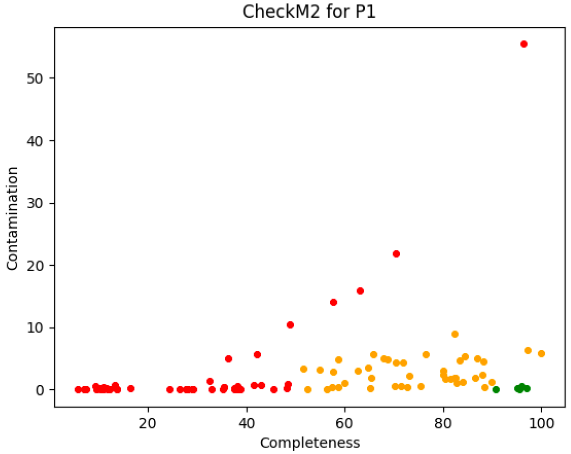
Reads per pig → Assembly with metaSpades → Binning with Metabat2 → Quality check with CheckM2

Quality of bins – Pig1

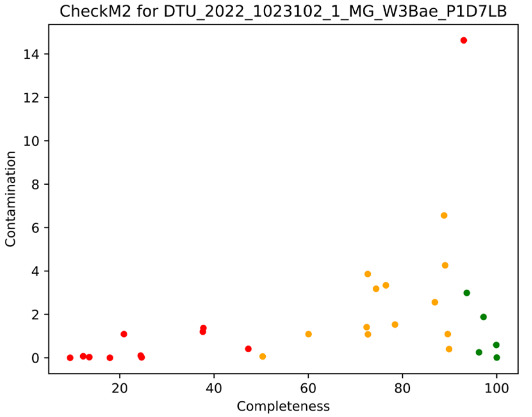

CheckM2 for DTU_2022_1023101_1_MG_W3Bae_P1D1LB

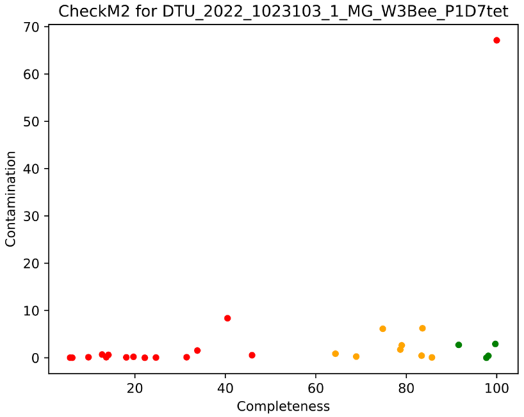

CheckM2 for DTU_2022_1023121_1_MG_W3Bae_P1D1tet

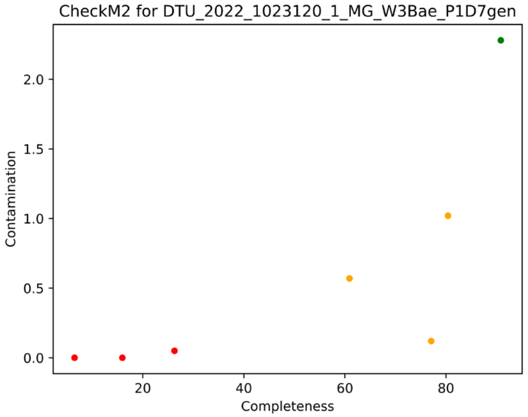

CheckM2 for DTU_2022_1023119_1_MG_W3Bae_P1D1gen


CheckM2 for P1

P1


CheckM2 for DTU_2022_1023102_1_MG_W3Bae_P1D7LB


CheckM2 for DTU_2022_1023103_1_MG_W3Bee_P1D7tet


CheckM2 for DTU_2022_1023120_1_MG_W3Bae_P1D7gen

Control

Tetracycline

Gentamycin

High quality
Medium quality
Low quality

Cake at the kitchen from 13:30 on Thursday☺