Antibiotic potential of soil bacteria targeting ESKAPE pathogens

Laura Marie Hagen Fur

MSc Student in Bioinformatics

Bioinformatics Research Center, Aarhus University

Spring 2025

Supervisors

Thomas Bataillon
Bioinformatics Research Center

Thomas Tørring
Biological and Chemical Engineering

Aarhus University

Abstract

Antibiotic resistance is a growing global concern, with the emergence and spread of antibiotic-resistant bacteria continuing to increase. Antibiotics are small metabolites naturally produced by microorganisms to ensure their survival in competitive environments by inhibiting or killing other competitive microorganisms. Conventional treatment options are getting exhausted, calling for an urgent need to discover novel antibiotics.

Here, I apply an *in silico* approach to discover natural products with potential antibacterial effects against ESKAPE pathogens by utilizing the bioinformatics pipeline, antiSMASH, to mine genomes of bacterial soil isolates for biosynthetic gene clusters (BGCs). In the study, unsupervised and supervised learning methods along with phylogenetic analysis, were integrated to find associations between observed bioactivity and the presence of specific BGCs. A total of 558 BGCs were predicted by antiSMASH, of which several BGCs were flagged in the analyses after correlating these with the results from the bioactivity assays. Some of these BGCs have reported antimicrobial activities with proposed broad- and narrow-spectrum activity, such as epilancin 15x, paenilamicin, and tridecaptin, which highlights the promising potential of combining supervised learning methods with laboratory work to the discovery of novel antibiotics.

Contents

Acronyms

A. baumannii	*Acinetobacter baumannii*	6
A. baylyi	*Acinetobacter baylyi*	6
ALS-PDC	amyotrophic lateral sclerosis–parkinsonism and dementia complex	1
AMR	antimicrobial resistance	1
antiSMASH	antibiotics and Secondary Metabolites Analysis SHell	4
BGC	biosynthetic gene cluster	4
BUSCO	Benchmarking Universal Single-Copy Orthologue	8
CDC	Center for Disease Control and Prevention	2
E. cloacae	*Enterobacter cloacae*	6
E. coli	*Escherichia coli*	6
EDA	exploratory data analysis	11

E. faecalis	*Enterococcus faecalis*	2
E. faecium	*Enterococcus faecium*	6
E. mori	*Enterobacter mori*	6
E. mundtii	*Enterococcus mundtii*	6
GTDB-Tk	the Genome Taxonomy Database Toolkit	8
K. oxytoca	*Klebsiella oxytoca*	6
MAG	metagenomics-assembled genome	7
MIBiG	Minimum Information about a Biosynthetic Gene cluster	6
P. aeruginosa	*Pseudomonas aeruginosa*	6
P. putida	*Pseudomonas putida*	6
pHMM	profile hidden Markov model	5

Prodigal procaryotic gene recognition and translation initiation site identification 5

S. aureus *Staphylococcus aureus* 6

S. epidermidis *Staphylococcus epidermidis* 6