

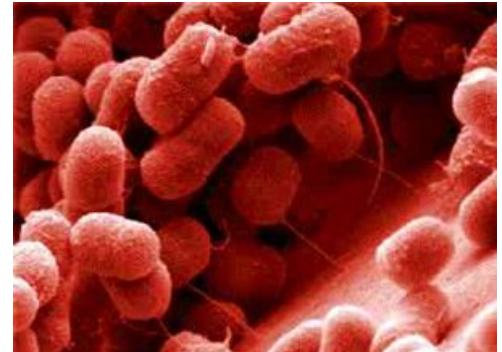
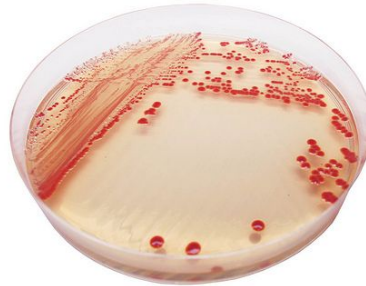
ACINETOBACTER BAUMANNII

FUNCTIONAL GENOME ANALYSIS

Expected Significance Team:
Troy Sincomb, Teva Bracha, Gayathri Kuppa

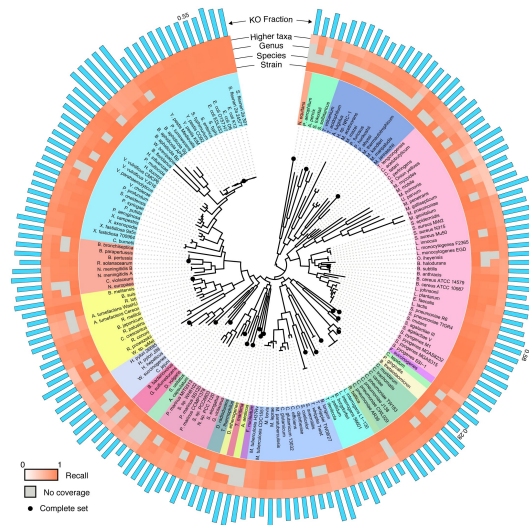
BACKGROUND

- Pleomorphic aerobic gram-negative bacillus, commonly found in hospital environments
- Associated with antibiotic resistance
- Some strains survive in presence of antibiotics
- Observe the mutated genes that develop resistance



FUNCTIONAL GENOME ANALYSIS

- Annotate sequences 301-400 to assign functional analysis to the genome.
- After we query sequences using different databases, we extract the functions assignments into a table.
- Each entry consists of keywords that are automatically extracted from the known annotations.

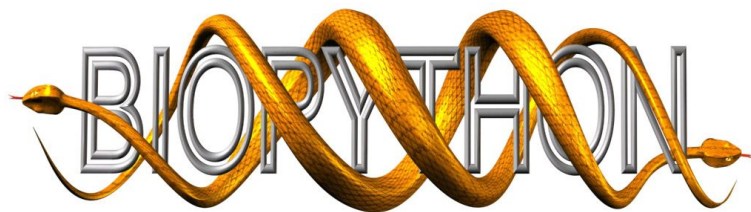


FUNCTIONAL GENOME ANALYSIS

Seq #	Function Names	Comments
300	putative extracellular nuclease; lamin tail domain	Likely to be a nuclease that cleaves nucleotides outside the cell or an intermediate filament in the nuclear domain. Not likely to be related to antibacterial resistance.
301	nuclease; lamin tail domain	Likely to be a nuclease or an intermediate filament in the nuclear domain. Not likely to be related to antibacterial resistance.
302	glutathione S-transferase	Enzyme that converts glutathione to other substances with detoxifying properties. Not likely to be related to antibacterial resistance.
303	alkyl hydroperoxide reductase subunit F	Reduces alkyl hydroperoxides to protect DNA from damage. Not likely to be related to antibacterial resistance.
304	alkyl hydroperoxide reductase subunit F	Reduces alkyl hydroperoxides to protect DNA from damage. Not likely to be related to antibacterial resistance.
305	Hypothetical Protein	In the DUF1653 pfam family.
306	lysine tRNA synthetase; KTSC domain protein	A ligase that uses ATP to bond lysine with tRNA that holds lysine for translation. Not likely to be related to antibacterial resistance.
307	ATP-dependent (RNA) helicase	Uses ATP to remodel RNA for regulatory purposes. Not likely to be related to antibacterial resistance.

QUERY TOOLS

- Biopython Blast
- Prody module
- Biopython ScanProsite
- Bioservices Kegg
- pfam2go & prosite2go geneontology



ProDy

Protein Dynamics & Sequence Analysis



OUTPUT FORMAT

- Only outputs changed for final format were Prosite and Blast
 - Prosite gave multiple hits, so cap was adjusted to top 10
 - Blast expected value was set to $1e-20$ and a cap of top 5
- Final format was collected file was made into a csv file and could be presenting in the pandas module.
- Using the pfam and prosite ids, the go-slim geneontologies for Pfam and Prosite were parsed.

SAMPLE OF DATA IN PANDAS DATAFRAME

	id	blast	pfam	prosite	kegg	go	comments
0	ABO11628	[['gi 1196949756 ref WP_086377352.1 ', 'nuclea...	noHomolog	noHomolog	noHomolog	noHomolog	301 -> nuclease; lamin tail domain -> Likely t...
1	ABO11629	[['gi 446703992 ref WP_000781338.1 ', 'MULTISP...	{ 'PF02798': { 'accession': 'PF02798.19', 'class...	[['sequence_ac': 'USERSEQ1', 'start': 1, 'stop...	{ 'acb00480': 'Glutathione metabolism'}	['GO:protein binding ; GO:0005515']	302 -> glutathione S-transferase -> Enzyme tha...
2	ABO11630	[['gi 1098612794 ref WP_071243622.1 ', 'alkyl ...	{ 'PF07992': { 'accession': 'PF07992.13', 'class...	[['sequence_ac': 'USERSEQ1', 'start': 97, 'sto...	noHomolog	['GO:oxidoreductase activity ; GO:0016491', 'G...	303 -> alkyl hydroperoxide reductase subunit F...
3	ABO11631	[['gi 126387133 gb ABO11631.1 ', 'alkyl hydrop...	noHomolog	[['sequence_ac': 'USERSEQ1', 'start': 115, 'st...	noHomolog	noHomolog	304 -> alkyl hydroperoxide reductase subunit F...
4	ABO11632	[['gi 445980391 ref WP_000058246.1 ', 'MULTISP...	{ 'PF07866': { 'accession': 'PF07866.10', 'class...	noHomolog	noHomolog	noHomolog	305 -> Hypothetical Protein -> In the DUF1653 ...
5	ABO11634	[['gi 691014800 ref WP_031976756.1 ', 'lysine ...	{ 'PF13619': { 'accession': 'PF13619.5', 'class'...	noHomolog	noHomolog	noHomolog	306 -> lysine tRNA synthetase; KTSC domain pro...
6	ABO11636	[['gi 126387138 gb ABO11636.1 ', 'ATP- dependen...	{ 'PF11898': { 'accession': 'PF11898.7',	[['sequence_ac': 'USERSEQ1', 'start': 1, 'stop...	noHomolog	noHomolog	307 -> ATP- dependent (RNA) helicase -> Uses

EXTRA CREDIT SEQUENCES 801-900

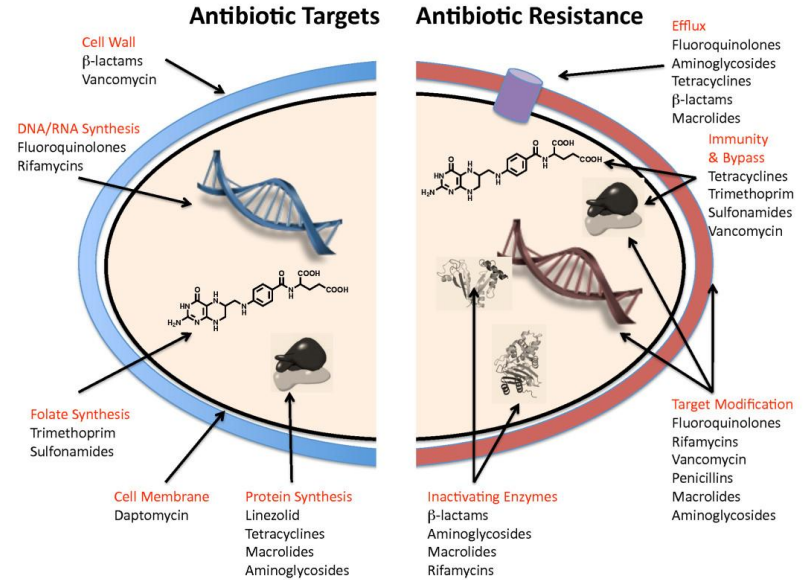
- ALL hits were either no Homolog or hypothetical proteins

	id	blast	pfam	prosite	kegg	go
0	ABS90243	[['gi 155030039 gb ABS90243.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
1	ABS90244	[['gi 155030040 gb ABS90244.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
2	ABS90314	[['gi 155030110 gb ABS90314.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
3	ABS90315	[['gi 155030111 gb ABS90315.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
4	ABS89919	[['gi 155029715 gb ABS89919.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
5	ABS89951	[['gi 446046034 ref WP_000123889.1 ', 'MULTISP...	noHomology	noHomology	noHomology	noHomolog
6	ABS90008	[['gi 155029804 gb ABS90008.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
7	ABS90009	[['gi 155029805 gb ABS90009.1 ', 'hypothetical...	noHomology	[['sequence_ac': 'USERSEQ1', 'start': 1, 'stop...	noHomology	noHomolog
8	ABS90072	[['gi 155029868 gb ABS90072.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
9	ABS90134	[['gi 446902484 ref WP_000979740.1 ', 'MULTISP...	noHomology	noHomology	noHomology	noHomolog
10	ABS90135	[['gi 588103822 gb EXD82750.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
11	ABS90171	[['gi 155029967 gb ABS90171.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
12	ABS90173	[['gi 446447182 ref WP_000525037.1 ', 'hypothe...	noHomology	noHomology	noHomology	noHomolog
13	ABS90271	[['gi 447043836 ref WP_001121092.1 ', 'hypothe...	noHomology	noHomology	noHomology	noHomolog
14	ABS90272	[['gi 155030068 gb ABS90272.1 ', 'hypothetical...	noHomology	noHomology	noHomology	noHomolog
15	ABS90316	[['gi 155030112 gb ABS90316.1 ', 'hypothetical...	noHomology	[['sequence_ac': 'USERSEQ1', 'start': 122, 'st...	noHomology	noHomolog

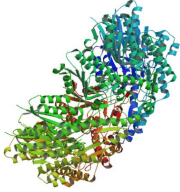
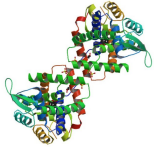
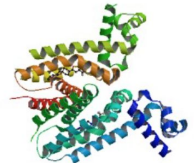
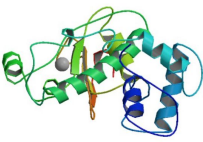
ANTIBIOTIC RESISTANCE

Three main mechanisms to antibacterial resistance:

- enzymatic degradation of antibacterial drugs
- alteration of bacterial proteins that are antimicrobial targets
- changes in membrane permeability to antibiotics



OUR GENES

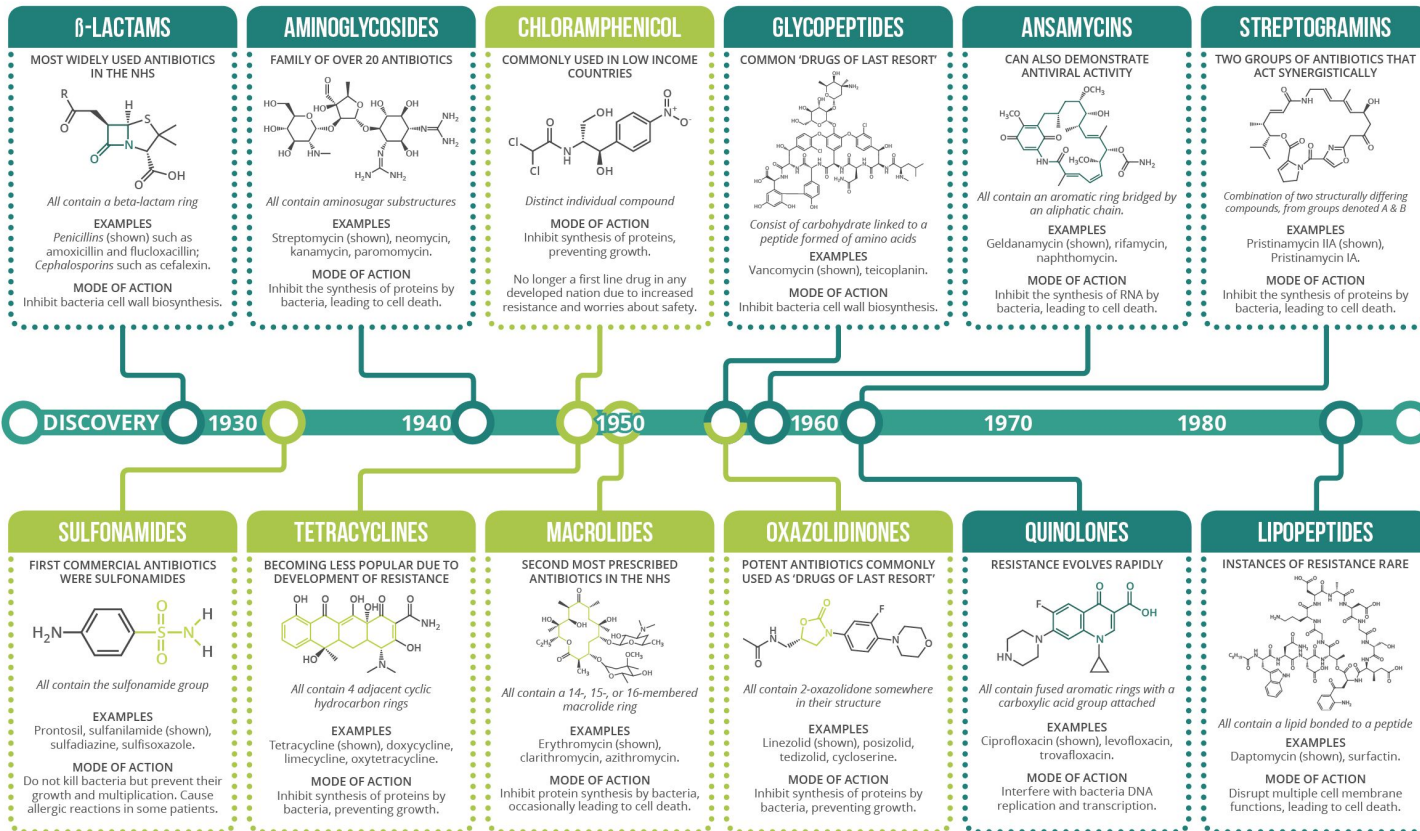
324		MBL fold metallo-hydrolase	Metallo-beta-lactamase protein that catalyzes the hydrolysis of a variety of β -lactam antibiotics. Could be responsible for antibacterial resistance
352		MarR family transcriptional regulator	Regulates the transcription of genes responsible for multiple antibiotic resistance.
359		TetR family transcriptional regulator	Transcriptional repressor that may be responsible for the efflux of tetracycline.
360		TetR family transcriptional regulator	Transcriptional repressor that may be responsible for the efflux of tetracycline.
382		Peptidase M15	Metallopeptidase family that can provide glycopeptide resistance to both vancomycin and teicoplanin. Potentially related to antibacterial resistance.

DIFFERENT CLASSES OF ANTIBIOTICS - AN OVERVIEW

Key:

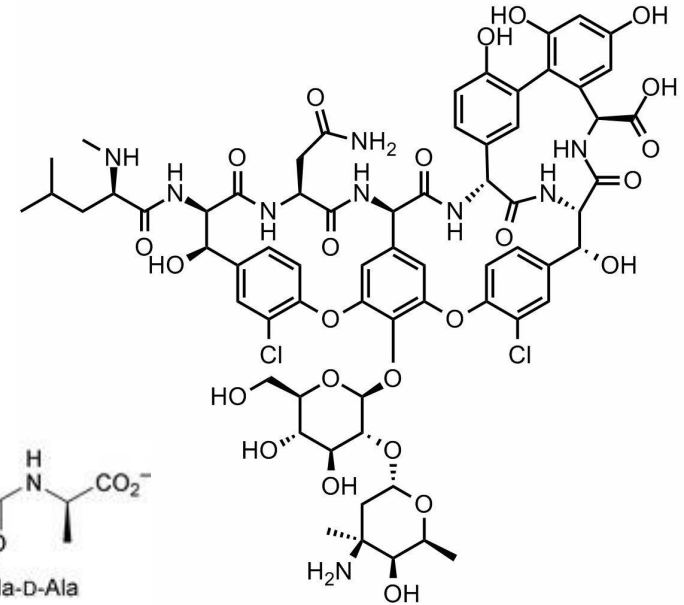
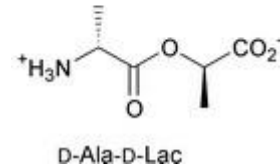
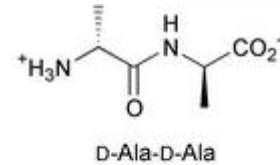
COMMONLY ACT AS BACTERIOSTATIC AGENTS, RESTRICTING GROWTH & REPRODUCTION

COMMONLY ACT AS BACTERICIDAL AGENTS, CAUSING BACTERIAL CELL DEATH



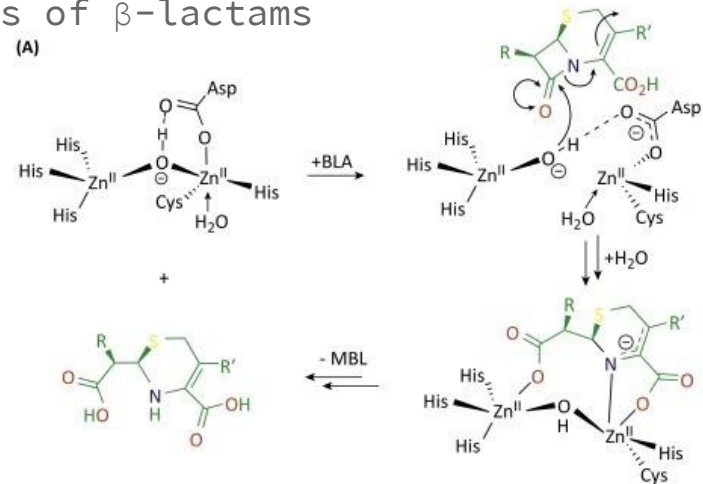
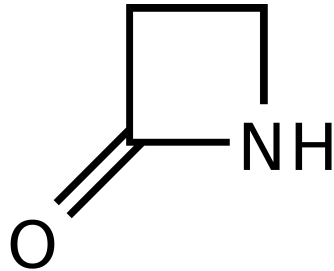
GLYCOPEPTIDES

- Stops cell wall synthesis by inhibiting peptidoglycan synthesis
 - Binds to precursors of which prevent their incorporation into the cell wall
- Vancomycin and Teicoplanin
- D-Ala D-Ala vs. D-Ala-D-lactate
- Our protein: Peptidase M15
 - Zinc dependent
 - VanX breaks down D-Ala-D-Ala
 - Potentially antibiotic resistant



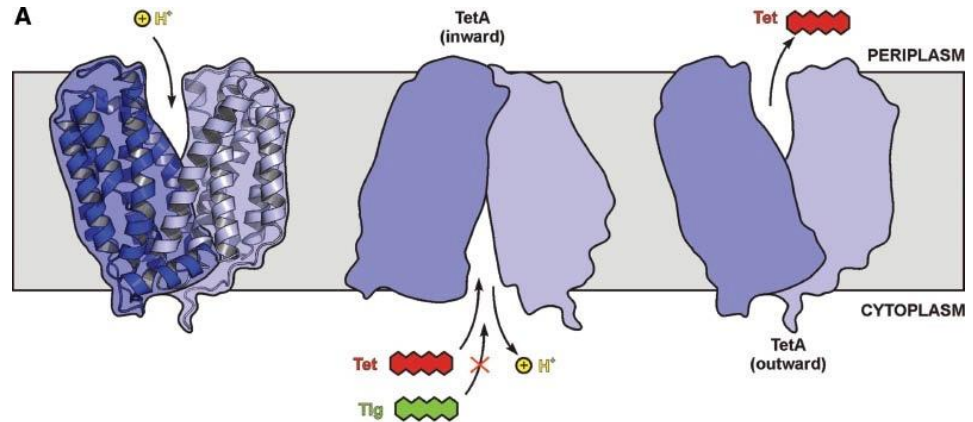
BETA-LACTAMS

- Also work by stopping bacterial cell wall synthesis
- Inhibits transpeptidases by mimicking D-Ala-D-Ala
- Our protein: Metallo-beta-lactamase fold metallo-hydrolase
 - Dependent on Zinc or Iron
 - Hydrolyze β -lactams into inactive β -amino acids
 - Neutralizes the antibiotic properties of β -lactams



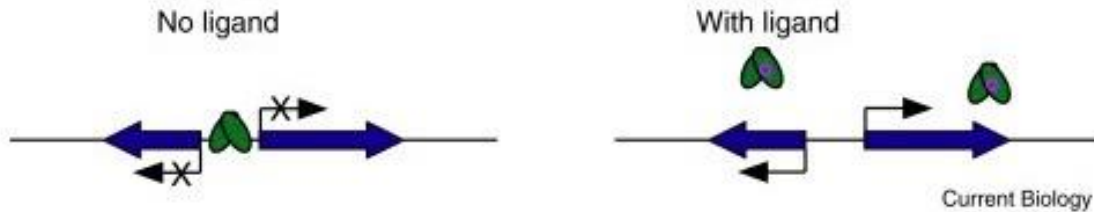
TETRACYCLINES

- Inhibits protein synthesis
 - Bind to the 30S ribosomal subunit in the mRNA translation complex
- Our protein: TetR family transcriptional regulator
 - Encodes repressor which binds to TetA Operon
 - Released in the presence of Tetracycline
 - TetA is a drug efflux pump



MULTIPLE ANTIBIOTIC RESISTANCE REGULATOR

- Transcriptional regulator of drug efflux pumps
- Responds to antibiotics and toxins
- Usually repressors



CLOSING THOUGHTS

- By assigning functionality to the protein sequences, we were able to detect key genes that relate directly to antibacterial resistance.
- Genome sequencing projects are important to understand complex cellular functions and offer insights in determining new treatments for various diseases.
- Genome sequences have served as the foundation for the advancement in investigating protein functionality.

REFERENCES

1. Cock PA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B and de Hoon MJL (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25, 1422-1423
2. Bakan A, Meireles LM, Bahar I. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* **2011** 27(11):1575-1577.
3. Thomas Cokelaer, Dennis Pultz, Lea M. Harder, Jordi Serra-Musach, Julio Saez-Rodriguez; BioServices: a common Python package to access biological Web Services programmatically. *Bioinformatics* 2013; 29 (24): 3241-3242. doi: 10.1093/bioinformatics/btt547