Computational Identification of novel antibiotics for *Klebsiella pneumoniae* from DNA adenine methyltransferase

Umairah Natasya Bt Mohd Omeershffudin
Bachelor in Bioinformatics (Final Year)
Department of Diagnostic & Allied Health Science (DHS)
Faculty of Health and Life Sciences (FHLS)
Management and Science University, Shah Alam, Malaysia

Supervisor: Dr. Suresh Kumar

Correspondence mail: Sureshkumar@msu.edu.my
The economic cost of AMR, in terms of lost global production, between now and 2050 is estimated at US$100 trillion. The O’Neill report estimates that almost 700,000 people worldwide die annually from AMR. The global impact of AMR could be 10 million deaths annually by 2050.

Figure 1: Antimicrobial resistance overview

Deaths attributable to antimicrobial resistance every year by 2050

North America 317,000
Latin America 392,000
Europe 390,000
Africa 4,150,000
Asia 4,730,000

Source: Review on Antimicrobial Resistance 2014

Antibiotic discovery and resistance timeline

Antibiotic class
- PENICILLINS
- MACROLIDES
- CARBAPEMEMS
- TETRACYCLINES
- FLUOROQUINOLONES

Date of resistance identified
- 1940
- 1953
- 1985
- 1993

Date of discovery
- 1928
- 1948
- 1985

30 years since a new class of antibiotics was last introduced

Figure 1: Antimicrobial resistance overview

# Introduction

**Klebsiella Pneumoniae**

<table>
<thead>
<tr>
<th>Class:</th>
<th>Gammaproteobacteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order:</td>
<td>Enterobacteriales</td>
</tr>
<tr>
<td>Family:</td>
<td>Enterobacteriaceae</td>
</tr>
<tr>
<td>Genus:</td>
<td>Klebsiella</td>
</tr>
<tr>
<td>Species:</td>
<td>K. pneumoniae</td>
</tr>
</tbody>
</table>

**Carbapenem-Resistant Enterobacteriaceae**

<table>
<thead>
<tr>
<th>Threat Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urgent</td>
<td>High-consequence antibiotic-resistant threats because of significant risks identified across several criteria. These threats may not be currently widespread but have the potential to become so and require urgent public health action and containment efforts.</td>
</tr>
<tr>
<td>Serious</td>
<td>Significant antibiotic-resistant threats. For varying reasons (e.g., low or declining domestic incidence or reasonable availability of therapeutic agents), they are not considered urgent, but these threats will require ongoing public health monitoring and prevention activities.</td>
</tr>
<tr>
<td>Concerning</td>
<td>These are bacteria for which the threat of antibiotic resistance is low, and/or there are multiple therapeutic options for resistant infections. These bacterial pathogens cause severe illness, and threats in this category require monitoring and in some cases rapid incident or outbreak response.</td>
</tr>
</tbody>
</table>

**Carbapenem-Resistant Enterobacteriaceae**

- **9,000** drug-resistant infections per year
- **600** deaths

**CRE have become resistant to all or nearly all available antibiotics**

[Figure 2: Klebsiella antimicrobial resistance]
Most antibiotics either target the three processes of the central dogma or the cell wall metabolism.

Current pace of antibiotic discovery is not sufficient to combat antibiotic resistance (Petchiappan & Chatterji, 2017)

Figure 3: Current antibiotic target approach
DNA methylation is carried out by enzymes known as DNA methyltransferases. Base methylation of cytosine and adenine is present in both bacteria and eukaryote. C5-Methyl-cytosine (m5C) are often found in eukaryotes while N6-methyl-adenine (m6A) are seen more impactful in bacteria. Methyl-directed mismatch repair, replication initiation, and gene expression (Low et al., 2001; Snyder & Champness, 2003).
Retrieve all 32 proteins of *Klebsiella Pneumoniae* that contain DNA adenine methyltransferases.

**Phase I**

**Non-homology analysis**

- BLASTp (NCBI) of selected proteins against *Homo sapiens* (nr database)

**Hit**

- E-value < 0.005
  - Yes: Excluded
  - No: Selected
METHODOLOGY

PHASE II

Essentiality analysis

BLASTp(DEC) of selected proteins against essential genes

Yes

Hit E-value < 0.0001

Selected

Excluded

No

BLASTp(NCBI) search against pathogenic organism (nr database)

Yes

Hit E-value < 0.0001

Selected

Excluded

No

Broad spectrum analysis
METHODOLOGY

Identification of drugability (Drugbank)

- Hit E-value < 0.0001
  - Yes: Selected
  - No: Excluded

Potential target (26)

- BLASTp (NCBI) of selected proteins against Gut Flora (nr database)

- Hit E-value < 0.005
  - Yes
  - No: Selected
  - Excluded

Non-homology analysis against gut microflora
METHODOLOGY

19 protein predicted for cellular localization

9 protein located in cytoplasmic

- MW < 100 kDa
- mean PI < 7.2
- Hydrophobicity (GRAVY) > 0.142
- Length > 550 (amino acids)
- Signal peptide present (likelihood > 0.5)
- Transmembrane helix > 1
- N-glycosylated ser more than 2
- O-glycosylated ser not more than 1

3 potential drug target
METHODOLOGY

BLASTp (NCBI) of selected protein against Anti-Targets (Human)

- Hit for e-value < 0.005
  - No: Selected
  - Yes: Excluded

BLASTp against protein sequences from VFDB

- Present in NCBI
  - Yes: Selected
  - No: Excluded
METHODOLOGY

Figure 5: Overview methodology of identifying drug target and prioritization
Figure 6: Identification of drug target using subtractive genomics and drug property analysis
Non-homology analysis: homology proteins contain unwanted toxicity that can later cause adverse pharmacokinetics effects and cross-reactivity in human cell.

Essentiality analysis: determine the essential genes required for the survival of the pathogen.

Broad spectrum analysis: select targets that have homologs in multiple pathogenic organisms.

Druggability analysis: ‘druggable’ target should have potential to bind to the drug-like molecules with high affinity.

Non-homology analysis against gut micro biota:
remove homologous protein that have similarities to the gut micro biota

Anti-target non-homology analysis: Identify of protein that is non-homologous to anti-target

Virulence factor analysis: identify potential targets that have similarities in VFDB

Figure 7: Flow chart of subtractive genomics
RESULTS & DISCUSSION

Drug Target Prioritization

MW < 100kDA
Increases absorption rate

Not more than one o-glycosylated

PPI interaction of >0.7

Mean PI (isoelectric points) < 7.2

Table 1: Drug target prioritization among 3 drug targets identified

<table>
<thead>
<tr>
<th>Uniprot Id</th>
<th>ChEMBL -targets identified</th>
<th>DoGSiteScorer</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0A3P4EC49</td>
<td>3</td>
<td>0.81</td>
</tr>
<tr>
<td>A0A3P4UG76</td>
<td>5</td>
<td>0.81</td>
</tr>
<tr>
<td>A0A2U0NNR3</td>
<td>10</td>
<td>0.82</td>
</tr>
</tbody>
</table>


Prioritized drug target: A0A2U0NNR3

Figure 8: STRING- PPI interaction analysis of drug target- A0A2U0NNR3
RESULTS & DISCUSSION

Functional annotation- Interproscan

Prioritized drug target: **A0A2U0NNR3**

**Figure 9: Functional annotation of drug target- A0A2U0NNR3**
## PATHWAY ANALYSIS

<table>
<thead>
<tr>
<th>Uniprot ID</th>
<th>KEGG pathway</th>
<th>EC</th>
<th>Gene ontology</th>
<th>Virulent factor prediction</th>
</tr>
</thead>
</table>
| A0A2U0NNR3   | ko03430 - DNA Mismatch repair | 2.1.1.72 -site-specific DNA-methyltransferase (adenine-specific); modification methylase | GO:0009007-Site-specific DNA-methyltransferase (adenine-specific) activity | • VFG010749 (sdhB) Dot/Icm type IV secretion system effector  
• VFG001959 (hddC) capsular polysaccharide heptosyltransferases |

**Table 2:** Pathway analysis of drug target- A0A2U0NNR3
Identified drug target occurs in DNA mismatch repair (MMR) biological pathway that plays a key role in maintaining genomic stability of bacteria. In this pathway, the GATC sites are methylated by Dam in *klebsiella pneumoniae* but absent in human which shows the promising target for inhibitor study for antibiotic development.

Figure 10: Pathway analysis of drug target-A0A2U0NNR3
Figure 11: (a) Homology modelling (B) Ramachandran Plot (C) Pocket identification (d) Binding pocket conservation

Prioritized drug target: A0A2U0NNR3

Ramachandran Plot
1078834

Plot statistics

| Residues in most favoured regions [A,B,L] | 211 | 89.3% |
| Residues in additional allowed regions [A,B1,L1] | 22 | 9.6% |
| Residues in generously allowed regions [-a,-b1,-l1] | 1 | 0.4% |
| Residues in disallowed regions | 1 | 0.4% |
RESULTS & DISCUSSION

CONCLUSION

We identified and prioritized one DAM drug target from 32 proteins of *klebsiella pneumonia* based on subtractive genomics, based on drug property, pocket analysis, pathway analysis and structure analysis. Our proposed new computational pipeline approach helps to find out the drug target in rapid way.

FUTURE WORK

Identification DAM inhibitors from ZINC database
To identify out best inhibitors for our identified drug target- **Prioritized drug target: A0A2U0NNR3** by using cheminformatics methods like ADMET property, Lipinski rule of 5 and molecular docking approach
Acknowledgment:
To my supervisor Dr. Suresh Kumar and to my parents and friends
To Department of Diagnostic & Allied Health Science (DHS), Faculty of Health and Life Sciences (FHLS), Management & Science University for providing computational facilities

My Publications related to this work


Molecular Docking Analysis between 2UZR protein from AKT1 gene of ovarian cancer and ligands of bioactive chemical compound of Frankincense Boswellia, Indian Conference on Bioinformatics Inbix’19, Jalandhar India 22nd -23rd April 2019

Thank you