Computational Prediction of Host-Pathogen Protein interactions between Melioidosis pathogen Burkholderia pseudomallei and Human

Chathumadavi Ediriweera
Department of Computer Science
Utah State University
Pathogen: Burkholderia pseudomallei

- Aerobic, Gram-negative, Motile bacillus
- Causes Melioidosis
- Opportunistic pathogen
- Produces exotoxins
- Found in water and moist soil
- Most commonly found in Southeast Asia, and in Northern Australia.
Why study about Burkholderia pseudomallei?

- Critical biothreat agent
- Causes Melioidosis in:
  - Human: pneumonia and sepsis
  - Animals
  - Plants: tomato
- Mortality is up to 50%

Goal

- Discover potential virulence factors using tomato as a host model.
- Apply those knowledge in drug discovery for melioidosis in human.
Role of Protein Interactions in disease

- To understand the biological details of pathogenesis and disease progression and to subsequently develop methods for prevention, treatment and even diagnosis
  - to identify the molecules and the mechanisms triggering, participating, and controlling the perturbed biological process
- Provide a vast source of molecular information
- Involved in metabolic, signaling, immune, and gene-regulatory networks
- Mediate the healthy states in all biological processes
- They should be the key targets of the molecular-based studies of biological diseased states
Host Pathogen Protein Protein Interactions

- Illuminate infection mechanisms
- For developing better prevention measures
- Knowledge of intra-species PPIs important in understanding the functional role of the proteins that are involved in host–pathogen PPIs

Fig. 1 Graphic representation of Burkholderia mallei and B. pseudomallei pathogenesis (Recent Advances in Burkholderia mallei and B. pseudomallei Research, Hatcher C., Muruato L. et al.)
Challenges with Host-Pathogen Interaction data

- Experimental Data are very scarce in most host–pathogen systems
- Experimentally verifying HPI is very scarce

- Computational prediction of host–pathogen PPIs
- Important strategy to fill in the gap
- If found Novel virulence factors can do computational prediction and assessment approaches for therapeutics drug design
Methodology

HPIDB

Interolog (Blastp)

MINT, IntAct, BioGRID, DIP

Subcellular localization (Uniprot annotations, SignalIP)

GO Enrichment Analysis (TopGO, DAVID)

Pathway Analysis (DAVID)

Host-Pathogen Interaction Database

Protein Interaction Database

Functional Analysis
Results : Interolog

At $e^{-20}$ ;

<table>
<thead>
<tr>
<th>Database</th>
<th>HPI Count</th>
<th>Host proteins</th>
<th>Pathogen proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPIDB</td>
<td>$\approx 9000$</td>
<td>2439</td>
<td>1424</td>
</tr>
<tr>
<td>Other Databases</td>
<td>$\approx 7600$</td>
<td>3324</td>
<td>1232</td>
</tr>
<tr>
<td>Consensus</td>
<td>1341</td>
<td>730</td>
<td>511</td>
</tr>
</tbody>
</table>
123 Bps are predicted using SignalP for the presence of signal peptide cleavage sites involved in Type VI, III and II/III secretion systems.
GO:0008150 biological_process
  └── GO:0009987 cellular_process
  └── GO:0008152 metabolic process
      ├── GO:0044237 cellular metabolic process
      ├── GO:0044281 small molecule metabolic process
      ├── GO:0071704 organic substance metabolic process
      └── GO:0006807 nitrogen compound metabolic process
          └── GO:1901564 organonitrogen compound metabolic process
              └── GO:0044238 primary metabolic process
                  └── GO:0006520 cellular amino acid metabolic process

GO Enrichment Analysis: Cellular Component

Human: Cellular Components

Top 20 GO Terms:
- monolayer-surrounded lipid storage body
- eukaryotic translation initiation factor 4F complex
- microsome
- clathrin-coated vesicle
- soluble fraction
- lipid particle
- filamentous actin
- vesicle membrane
- membrane raft
- clathrin adaptor complex
- Golgi transport complex
- Golgi apparatus part
- extracellular region
- contractile fiber
- perinuclear region of cytoplasm
- AP-type membrane coat adaptor complex
- cell surface
- dendrite
- cytoplasmic vesicle membrane
- heterogeneous nuclear ribonucleoprotein complex

Count
GO Enrichment Analysis: Cellular Component

Burkholderia p.mallei: Cellular Components
Top 20 GO terms

- intracellular part
- plasma membrane
- catalytic complex
- DNA repair complex
- chromosome
- bacterial-type flagellum
- envelope
- ATPase complex
- excinuclease repair complex
- endodeoxyribonuclease complex
- endonuclease complex
- ATP-binding cassette (ABC) transporter complex
- ATPase dependent transmembrane transport complex
- cell envelope
- bacterial-type flagellum basal body, rod
- plasma membrane part
- transmembrane transporter complex
- transporter complex
- bacterial-type flagellum part
- cell projection part

Count

0 50 100 150 200
GO Enrichment Analysis: Molecular Function

Human: Molecular Function

Top 20 GO terms:

- Diphosphoinositol-polyphosphate diphosphatase activity
- Insulin-like growth factor II binding
- O-palmitoyltransferase activity
- Protein heterodimerization activity
- Carboxypeptidase activity
- Cysteine-type endopeptidase inhibitor activity
- Carboxylesterase activity
- SH3 domain binding
- Serine hydrolase activity
- C-X-C chemokine binding
- General transcriptional repressor activity
- Hydrolase activity, acting on ether bonds
- Notch binding
- Metalloexopeptidase activity
- Thyroid hormone receptor binding
- Calcium-dependent protein binding
- Glycosaminoglycan binding
- Sequence-specific DNA binding
- Transferase activity, transferring sulfur-containing groups
- Protein C-terminus binding
GO Enrichment Analysis: Molecular Function

Bukholderia p.mallei: Molecular Function

Top 2 GO terms

- adenyl nucleotide binding
- ATP binding
- purine nucleotide binding
- nucleotide binding
- nucleoside phosphate binding
- adenyl ribonucleotide binding
- purine ribonucleotide binding
- ribonucleotide binding
- carbohydrate derivative binding
- nucleoside-triphosphatase activity
- pyrophosphatase activity
- hydrolase activity, acting on acid anhydri...
- hydrolase activity, acting on acid anhydri...
- ATPase activity
- vitamin binding
- organic cyclic compound binding
- heterocyclic compound binding
- ligase activity
- binding
- molecular transducer activity

Count

0 200 400 600
Pathway Analysis

Human: KEGG Pathway Analysis

Top 20 Terms

- Type I diabetes mellitus
- Fc gamma R-mediated phagocytosis
- Intestinal immune network for IgA production
- Autoimmune thyroid disease
- Methane metabolism
- Glycosaminoglycan degradation
- Calcium signaling pathway
- Metabolism of xenobiotics by cytochrome P...
- Fructose and mannose metabolism
- Ether lipid metabolism
- Sulfur metabolism
- Complement and coagulation cascades
- GnRH signaling pathway
- Adherens junction
- Systemic lupus erythematosus
- Viral myocarditis
- Notch signaling pathway
- Riboflavin metabolism
- Insulin signaling pathway
- Nitrogen metabolism
Pathway Analysis

KEGG Pathway: Burkholderia p.mallei

Top 20 Terms

- D-Glutamine and D-glutamate metabolism
- Bacterial chemotaxis
- Glyoxylate and dicarboxylate metabolism
- Fructose and mannose metabolism
- Two-component system
- Pantothenate and CoA biosynthesis
- Nicotinate and nicotinamide metabolism
- Biotin metabolism
- Arginine and proline metabolism
- Nicotinate and nicotinamide metabolism
- C5-Branched dibasic acid metabolism
- Biotin metabolism
- Bacterial chemotaxis
- Glycine, serine and threonine metabolism
- Valine, leucine and isoleucine biosynthesis
- Lipopolysaccharide biosynthesis
- Starch and sucrose metabolism
- Glyoxylate and dicarboxylate metabolism
- D-Glutamine and D-glutamate metabolism
- Biotin metabolism
Conclusion

**Two-component systems**: Two-component system in *Burkholderia cenocepacia*, an opportunistic pathogen that causes pneumonia in cystic fibrosis (CF) patients.

- Important regulatory systems that allow bacteria to adjust to environmental conditions, and in some bacteria are used in pathogenesis.
**Bacterial Chemotaxis**: suggests that bacteria can actively search for and target phagocytic cells, which may help us better understand how bacteria interact with immune systems.

- Initiate stable associations with host cells.

*Burkholderia* bacteria use chemotaxis to find social amoeba *Dictyostelium discoideum* hosts

Longfei Shu, Bojie Zhang, David C. Queller & Joan E. Strassmann

Future Directions

- In Progress:
  - Domain-domain interaction method.
  - Subcellular localization on data acquired from domain-domain interaction.

- Next Step:
  - Machine Learning approach
Thank you