Mahmoud and Aziz *BMC Bioinformatics* 2010, **11**(Suppl 4):P28
http://www.biomedcentral.com/1471-2105/11/S4/P28

BMC Bioinformatics

POSTER PRESENTATION Open Access

Do pathogenic bacteria encode more secreted proteins than their non-pathogenic relatives?

Ahmed Abo-Bakr Mahmoud[1], Ramy Karam Aziz[2*]

From UT-ORNL-KBRIN Bioinformatics Summit 2010
Cadiz, KY, USA. 19-21 March 2010

Background

Pathogenic and non-pathogenic bacteria secrete proteins for nutrient acquisition, cell-cell communication, and niche adaptation [1]. We hypothesized that pathogenic bacteria may encode larger fractions of secreted proteins (fsp) than their non-pathogenic relatives, assuming that pathogens might be under selective pressure to secrete virulence proteins involved in host immune evasion, invasion, and toxigenesis. To test this hypothesis, we compared the Sec-dependent fsp of various gram-positive and gram-negative bacteria and investigated the relation between the fsp and pathogenic potential of an organism.

Methods

We developed a pipeline that starts by a Perl script that truncates protein sequences to 70 amino acids or fewer followed by the application of existing signal prediction tools [2-4] and ends by the statistical analysis of the prediction data. For subsequent comparative secretome analyses, we used both the hidden Markov models- and the neural networks-based methods implemented in the SignalP 3.0 algorithm [2] (URL: http://www.cbs.dtu.dk/services/SignalP) with modified thresholds. We used DataDesk (Data Description, Inc., Ithaca, NY; URL: http://www.datadesk.com) for all statistical analyses (including correlation analysis, analysis of variance, and multivariate analysis) and for plotting the results.

Results

We determined the theoretical secretomes of 176 chromosomes and 115 plasmids in five gram-positive and five gram-negative bacterial genera containing

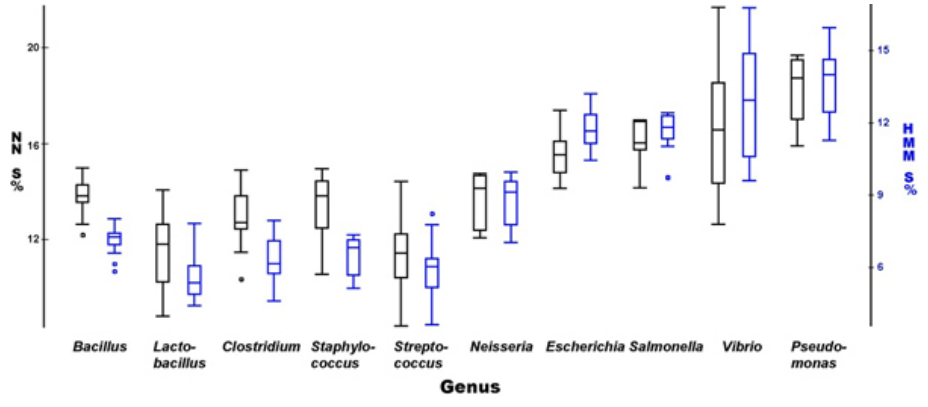

Figure 1 Predicted fractions of secreted proteins by the chromosomes of ten bacterial genera with pathogenic and non-pathogenic member. NN = neural networks method; HMM = hidden Markov models method.

* Correspondence: ramy.aziz@salmonella.org
[2]Department of Microbiology and Immunology, Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt

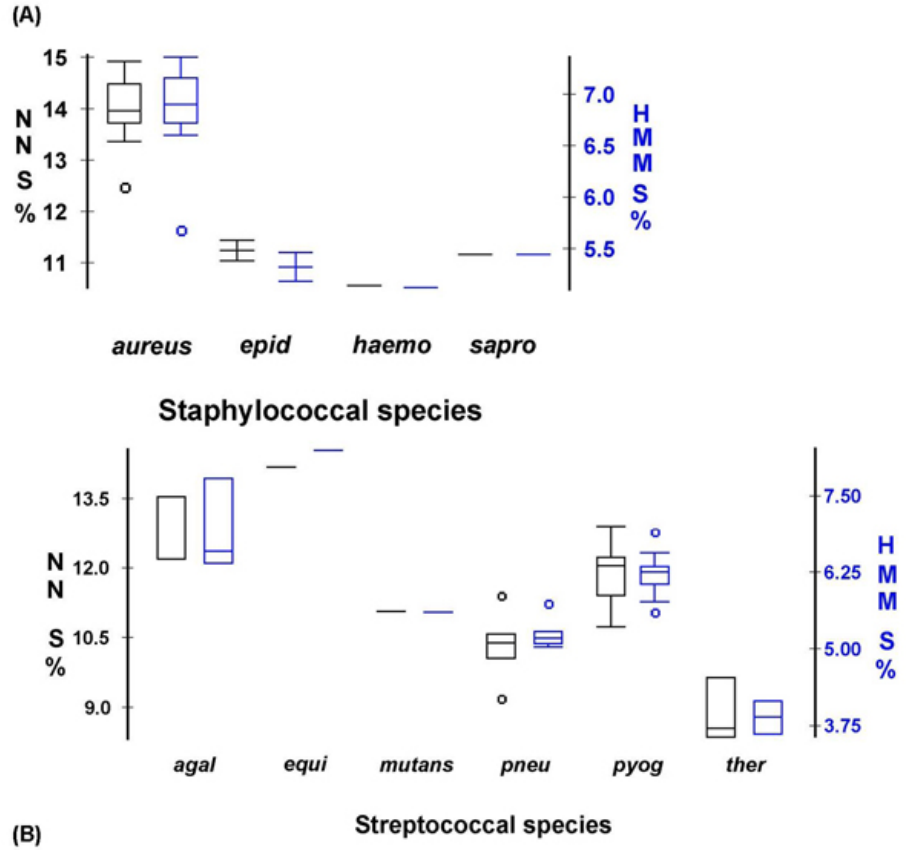

Figure 2 Pathogenic Gram-positive cocci encode larger fractions of secreted proteins than non-pathogenic relatives. NN = neural networks method; HMM = hidden Markov models method. (A) Staphylococcal species (*epid* = *epiderimidis; haemo* = *haemolyticus; sapro* = *saprophyticus*). (B) Streptococcal species (*aga* = *agalactiae; pneu* = *pneumoniae; pyog* = *pyogenes; ther* = *thermophilus*) *P* values: *aureus* vs. all: $P < 10\text{-}6$; *ther* vs. all: $P = 0.002$ (NN), 0.0001 (HMM); *agal* vs. *pyog*: $P = 0.012$ (NN), 0.003 (HMM).

pathogenic and non-pathogenic members (Figure 1). Our analysis showed significant differences in chromosomally encoded fsp between gram-positive and gram-negative bacteria (chromosomes of gram-negative bacteria have larger fsp), while there was no particular pattern in plasmid-encoded fsp. Whereas the overall difference between pathogenic and non-pathogenic species was not statistically significant, significant correlation was observed between fsp and pathogenesis in gram-positive cocci. For example, pathogenic *Staphylococcus aureus* have higher fsp than other staphylococci, while the non-pathogenic *Streptococcus thermophilus* has the lowest fsp of all streptococci (Figure 2).

Conclusion

We developed a pipeline for the determination and comparison of fractions of secreted proteins in bacterial genomes, and observed significant differences between pathogenic and non-pathogenic species of staphylococci and streptococci.

Author details
[1]Maternity Governmental Hospital, Kuwait City, 85860, Kuwait. [2]Department of Microbiology and Immunology, Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt.

Published: 23 July 2010

References

1. Gennity JM, Inouye M: **Protein secretion in bacteria.** *Curr Opin Biotechnol* 1991, **2**:661-667.
2. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**:783-795.
3. Emanuelsson O, Brunak S, von Heijne G, Nielsen H: **Locating proteins in the cell using TargetP, SignalP and related tools.** *Nat Protoc* 2007, **2**:953-971.
4. Zhou M, Boekhorst J, Francke C, Siezen RJ: **LocateP: genome-scale subcellular- location predictor for bacterial proteins.** *BMC Bioinformatics* 2008, **9**:173.

doi:10.1186/1471-2105-11-S4-P28
Cite this article as: Mahmoud and Aziz: Do pathogenic bacteria encode more secreted proteins than their non-pathogenic relatives?. *BMC Bioinformatics* 2010 **11**(Suppl 4):P28.