

## Abstract

*Coxiella burnetii* is the causative agent of Q fever and listed as a category B bioterrorism agent by the Centers for Disease Control and prevention (CDC). It is an obligate intracellular gram-negative bacterium of the gamma subdivision of Proteobacteria. Proteins that are secreted can be essential for the lifestyle of parasites/infectious agents. These include penetration of host tissue barriers, feeding and evasion or modulation of the host immune response. Elucidation of secreted proteins may provide insight into the dynamics of the host-pathogen relationship and may identify potential vaccine or therapeutic targets. Thus far there has been no published evidence for the presence of secreted proteins in *Coxiella*. As a test case, we are mining the annotated genome of *Coxiella burnetii* to identify proteins for potential vaccine development based on the possibility of them being expressed at the bacterial surface. We have developed a prototype pipeline that predicts potential surface-exposed outer membrane proteins. The pipeline has yielded a set of *Coxiella burnetii* proteins that contain a putative signal peptide required for entry into a secretory pathway. Identification of these proteins showed that they were components of a Type I secretion system. Searches over the database for the other components of this pathway were positive. Follow up studies will include structural analysis to determine surface exposed regions that may serve as epitopes.

## Introduction

The PathoSystems Resource Integration Center (PATRIC) was established at the Virginia Bioinformatics Institute (VBI), in July 2004. It is one of the eight Bioinformatics Resource Centers (BRC) that have been set up by the National Institute of Allergy and Infectious Disease (NIAID) with a mandate to integrate biological information on NIAID/CDC category A-C microbial pathogens, as well as to aid research on disease detection and prevention and development of therapeutic solutions. In addition to *Coxiella*, PATRIC contains annotated and curated genomic and related biological data for two additional bacteria, *Brucella* and *Rickettsia*, as well as five viral pathogens, Corona-, Lyssa-, and Caliciviruses plus the viruses that cause Hepatitis A and Hepatitis E. The website contents are available at our website and the website address is: <http://patric.vbi.vt.edu/>

Q fever typically comes from occupational exposure involving meat processing, sheep and dairy workers, but because these spores can be deliberately released into the environment, it is listed as a category B bioterrorism agent. Though vaccines exist for *Coxiella*, there is a need for a more efficacious vaccine. In keeping with PATRIC's overall goals of using pathogen genome data and bioinformatic tools to facilitate the development of drugs, vaccines and diagnostics, this project has been undertaken. This project aims to use bioinformatic tools to identify proteins that had the potential of being secreted proteins and their subcellular localization. Until now, no prediction of secreted proteins has been made for *Coxiella*, and no proteins have been identified as being part of a secretory apparatus.

## Material and Methods

**Input Data:** *Coxiella burnetii* RSA 493 Genome, Chromosome I (2016 protein coding genes)

### Tools:

SignalP<sup>1</sup>, predicts the presence of a signal peptide, and signal peptide cleavage site (accepted  $\geq 80\%$  confidence)  
BOMP<sup>2</sup>, predicts the presence of beta-barrel structures (Beta-barrel integral Outer Membrane Protein -BOMP)  
SigCleave<sup>3</sup>, predicts signal peptide cleavage sites (score of 3.5 and above)  
psortB<sup>4</sup>, predicts subcellular localization.

### Signal P positive with predicted beta barrel

- 1&2 Outer Membrane Protein A (OmpA)
3. Hypothetical protein
4. Outer Membrane Protein A-like (OmpA-like)
5. DotA protein, putative

### Signal P negative with predicted beta barrel

- 1-3. Outer membrane protein A-like (OmpA)
- 4&5. Hypothetical proteins
6. Hypothetical protein
7. Hypothetical protein
8. Enhanced entry protein
9. Organic solvent tolerance protein

### Signal P positive identified by PATRIC

- 1-4. Hypothetical proteins
5. Na<sup>+</sup>/H<sup>+</sup> antiporter, putative
6. Hypothetical protein

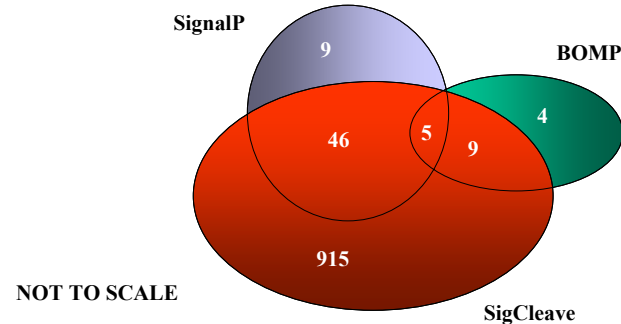
**Table 1 Potential secretory proteins.** The prediction programs SignalP, SigCleave and BOMP were run on the *C. burnetii* proteome to identify putative secretory proteins. These proteins may serve as possible drug and vaccine development targets.

## Acknowledgements

This project is funded by NIAID / NIH contract HHSN26620040035C awarded to Bruno Sobral. Our thanks to Abdu Azad, Organism Expert for *Coxiella*, and Joseph Gillespie for critically evaluating the poster and offering excellent suggestions.

## Results

Out of a total of 2016 protein coding genes (chromosome I alone), SignalP predicted 60 proteins with a putative signal peptide (80% confidence level and above). SigCleave predicted 980 proteins of which 51 overlapped with the SignalP (Fig 1). BOMP predicted 18 proteins with a beta-barrel structure, nine of which overlapped with SigCleave, and five were common to all three prediction programs. Hence, our best candidates for protein secretion under the classical (type V) and ABC autotransporter (type I) secretory systems were the five (intersection of all three prediction programs-Fig 1) that had a signal peptide (with cleavage sites) and had a beta-barrel structure. The next set of nine proteins were at the intersection of SigCleave and BOMP and finally the six 46 that were at the intersection of SignalP and SigCleave alone. Six of these, are part of a group of genes that were added during re-annotation at VBI. (For original set, see Seshadri et.al.<sup>5</sup>). Psorth was run on all of these proteins for their subcellular localization (Table 2). One of the aims of this project was not only to identify potential signal peptide bearing proteins, but also to identify other components of a secretory system. Our initial search identified multiple outer membrane protein A (OmpA)-like proteins (Table 1) which are associated with the Type I secretory apparatus (Sec-independent). Hence, we did a search to identify the other two members of this apparatus (See Fig.2 and Ref.6) viz. an ABC protein superfamily of transporters and a membrane fusion protein (MFS). This search was performed in the PATRIC database which had the original RefSeq annotation as well as genes added during the re-annotation by VBI. We found numerous ABC transporters, and one membrane fusion protein of the HlyD family, suggesting a probable Type I secretory apparatus.



NOT TO SCALE

Figure 1 Overlap between different prediction programs Proteins that are common between SignalP, SigCleave and BOMP programs

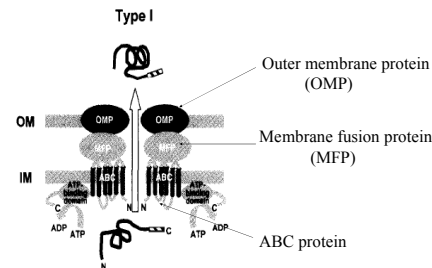
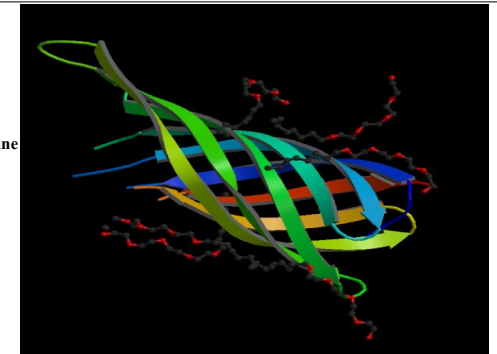


Figure 2 Schematic view of the Type I secretion system<sup>6</sup>  
Outer membrane protein, membrane fusion protein and ABC Protein make up the 3 components of the secretion system

Figure 3 High resolution structure of the outer membrane protein A (OMP) transmembrane domain<sup>7</sup> PDB X-ray diffraction structure from *E.coli* shown as a model only



## Summary and future direction

We have identified potential drug targets for vaccine development from our search using bioinformatic tools. These potential proteins will be further characterized (especially hypothetical proteins), both by ourselves (bioinformatic studies) as well as using expression studies in a laboratory. We have also potentially identified the components that make up a Type I secretory apparatus (Sec-independent), in *Coxiella burnetii*. Further, these bioinformatic tools will be made available on our website.

**References:** <sup>1</sup>J.Mol.Biol., 340:783 (2004) <sup>2</sup>Nucleic Acids Res., 32:W394 (Web Server issue) (2004) <sup>3</sup>Nucleic Acids Res., 14:4683 (1986) <sup>4</sup>Bioinformatics 21:617 (2005) <sup>5</sup>PNAS 100:5455 (2003) <sup>6</sup>J.Bioscience Bioengineering 95:1 (2003) <sup>7</sup>J.Mol.Biol., 298:273 (2000)