Step I: White Paper Application

Application Guidelines

1. *The application should be submitted electronically per requirements via the web site of any of the NIAID Genomic Sequencing Centers for Infectious Diseases. Include all attachments, if any, to the application.*
2. *There are no submission deadlines; white papers can be submitted at anytime.*
3. *GSC personnel at any of the three Centers can assist / guide you in preparing the white paper.*
4. *Investigators can expect to receive a response within 4-6 weeks after submission.*
5. *Upon approval of the white paper, the NIAID Project Officer will assign the project to a NIAID GSC to develop a management plan in conjunction with the participating scientists.*

White Paper Application

Project Title:

Authors:
Primary Investigator Contact:

Name	Julia Vipond
Position	Operational Lead
Institution	Health Protection Agency – London
Address	Porton Down Salisbury SP4 0JG, UK
State	
ZIP Code	
Telephone	+44(0) 1980 612584
Fax	+44(0) 1980 619894
E-Mail	julia.vipond@hpa.org.uk

1. Executive Summary *(Please limit to 500 words.)*

Provide an executive summary of the proposal.

It has been observed for nearly 80 years that *Burkholderia pseudomallei* can change colony morphology (1). In 2007, Chantratita et al. presented an interesting analysis and cataloguing of several colony morphologies found both between patients and within patients infected with *B. pseudomallei* (2). They have hypothesized that the bacteria undergo a type of adaptation or switching in response to stresses encountered either in the environment or during an infection. They went on to show that these different colony morphotypes can have different associations with epithelial cells and macrophages.

A group at the HPA in the UK has observed colony morphotypes arising after mice are infected with *B. pseudomallei* K96243. Specifically, these colony morphotypes were seen in bacteria cultured from the lungs. By PCR analysis, these different colonies are not caused by contamination, and are likely "switched" to another state. It is known that *B. pseudomallei* can lie dormant in humans and produce an active infection years after initial exposure, and the lung may be a place where the dormancy occurs. Although it could very likely be a far stretch to think that one of the morphotypes could be a form of the bacteria that could lie dormant, it is of interest to characterize the morphotypes.

To date, to our knowledge, no one has characterized the molecular mechanisms of this *B. pseudomallei* switching. It seems that deep sequencing of the genomes of the colony morphotypes would be the most efficient way to learn what the genetic changes are. Sequencing needs to be of sufficient coverage to discern minor differences such as inversions (it could be entire operons or simply a promoter or regulatory region), transpositions, or duplications. It is unlikely that the genetic differences are caused by loss of genetic material since it appears from the Chantratita paper and other anecdotal evidence that morphotypes can be forced to become another type.

We propose sequencing of the DNA extracted from two colony morphotypes. One of these types is from the positive control of *B. pseudomallei* K96243 created for use in the enumeration procedure and is therefore typical of the morphologies we have seen to date. The second type was isolated from two animals which had received a low dose stationary phase aerosol infection. This morphotype was atypical of the morphologies seen to date.

Both types have been characterized as Gram negative rods and are PCR positive for *B.pseudomallei* (*lpxO* gene target).

2. Justification

Provide a succinct justification for the sequencing or genotyping study by describing the significance of the problem and providing other relevant background information.

This section is a key evaluation criterion.

1. *State the relevance to infectious disease for the organism(s) to be studied; for example the public health significance, model system etc.*

2. *Are there genome data for organisms in the same phylum / class / family / genus? What is the status of other sequencing / genotyping projects on the same organism including current and past projects of the NIAID GSC? Provide information on other characteristics (genome size, GC content, repetitive DNA, pre-existing arrays etc.) relevant to the proposed study. Have analyses been performed on the raw data already generated/published? If additional strains are proposed for a species, please provide a justification for additional strains?*

3. *If analyses have been conducted, briefly describe utility of the new sequencing or genotyping information with an explanation of how the proposed study to generate additional data will advance diagnostics, therapeutics, epidemiology, vaccines, or basic knowledge such as species diversity, evolution, virulence, etc. of the proposed organism to be studied.*

Burkholderia pseudomallei is the causative agent of melioidosis, a highly fatal disease common in South East Asia and tropical areas of Australia. There currently is no protective vaccine, and having been infected does not lead to immunity in indivuduals. *B. pseudomallei* is resistant to many antibiotics; indeed, the treatment for melioidosis are drug cocktails for an extended period of time. Due to its very low infectious and lethal doses, it multi-drug resistance, and lack of immunity, *B. pseudomallei* is classified as a potential biothreat agent in the United States and the United Kingdom. As of December 2010, there exists only a limited number of *B. pseudomallei* strains that are sequenced. The genomic analysis of this genome, and its close relative, *B. mallei*, demonstrates that these genomes are highly variable – not only in gene content, but also in the quantity of short tandem repeats and mobile genetic elements. We do not yet understand whether this genome plasticity is involved in the phenotypic diversity observed in what amount to clonal populations. Furthermore, our own and other published studies have demonstrated a direct link between different morphologies and virulence. Combined, these facts underscore the need to obtain genomic sequences from different colony morphotypes from the same *B. pseudomallei* isolate.

3. Rationale for Strain Selection

4. *Provide the rationale behind the selection of strains and the number of strains proposed in the study. The focus of the program is on potential agents of*

bioterrorism or organisms responsible for emerging or re-emerging infectious diseases. Non-select agents or non-pathogenic organisms will be considered when they can provide insight into these scientific areas.

The strains to be sequenced are derivatives from a single isolate that was used to infect a mouse. After infection, different *B. pseudomallei* colony morphotypes were isolated from the spleen, liver, lungs, and blood of the same animal. The genomic DNA of these isolates will be prepared for sequencing.

4a. Approach to Data Production: Data Generation

5. *State the data and resources planned to be generated. (e.g draft genome sequences, finished sequence data, SNPs, DNA/protein arrays generation, clone generation etc.)*

The resulting genomic sequencing results will be used to generate a draft genome assembly of the wild-type morphotype. The remaining sequences will be used to identify polymorphisms associated with each type of morphology. These will include SNPs, Deletions/Insertions, amplications or large structural variations in the genomes.

4b. Approach to Data Production: Data Analysis

6. *Briefly describe the analysis (value-add) envisioned to be performed subsequently by the community and the potential to develop hypotheses driven proposals given the datasets and resources produced by this work.*

Identified genomic changes will be used to make direct mutations or changes in the wild-type morphotype and/or other *B. pseudomallei* strains to confirm the relevance of the changes in the phenotype.

5. Community Support and Collaborator Roles:

7. *Provide evidence of the relevant scientific community's size and depth of interest in the proposed sequencing or genotyping data for this organism or group of organisms. Please provide specific examples.*

8. *List all project collaborators and their roles in the project*

9. *List availability of other funding sources for the project.*

The *Burkholderia pseudomallei* community is fairly large, with over 350 participants in the scientific meeting dedicated to this organism and the disease it causes. All members of that community would benefit from additional sequenced isolates, as well as from information related to phenotype/genotype relationships.

Dr. Susan Garges at the NIAID has been an important intellectual contributor to this white paper.

6. Availability & Information of Strains:

10. *Indicate availability of relevant laboratory strains and clinical isolates. Are the strains/isolates of interest retrospectively collected, prepared and ready to ship?*

Note: If samples are prospectively prepared the GSC can provide protocols and recommendation based on the Centers past experiences. The samples must however meet minimum quality standards as established by the Center for the optimal technology platform (sequencing/ genotyping) to be used in the study.

11. *Attach relevant information, if available in an excel spreadsheet for multiple samples: e.g*
    - *Name*
    - *Identifier*
    - *Material type (DNA/RNA/Strain)*
    - *Genus*
    - *Species*
    - *Specimen / Strain*
    - *Isolation source*
    - *Isolated from*
    - *Select agent status*
    - *International permit requirement*
    - *BEIR/ATCC repository accession number*
    - *Other public repository location*
    - *Other public repository identifier*
    - *Sample provider's name*
    - *Sample provider's contact*

12. *What supporting metadata and clinical data have been collected or are planned on being collected that could be made available for community use?*

All data pertaining to the experimental infection of the mouse, along with the site of isolation and the colony morphology will be made available to the community. The wild-type strain and other morphotpes are available through the HPA's public stock repository.

7. Compliance Requirements:

7a. Review NIAID's Reagent, Data & Software Release Policy:

NIAID supports rapid data and reagent release to the scientific community for all sequencing and genotyping projects funded by NIAID GSC. It is expected that projects will adhere to the data and reagent release policy described in the following web sites.

http://www.niaid.nih.gov/labsandresources/resources/mscs/data.htm

http://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-013.html

Once a white paper project is approved, NIAID GSC will develop with the collaborators a detailed data and reagent release plan to be reviewed and approved by NIAID.

Accept ☒ Decline ☐

7b. Public Access to Reagents, Data, Software and Other Materials:

13. *State plans for deposit of starting materials as well as resulting reagents, resources, and datasets in NIAID approved repositories. Sequencing projects will not begin until the strain is deposited into NIAID funded BEI repository (http://www.beiresources.org/). This includes web based forms are completed by the collaborator and received by the NIAID BEI (http://www.beiresources.org/).*

All data will be released to NCBI. Strains are already available through the HPA's public stock repository in London.

7c. Research Compliance Requirements

Upon project approval, NIAID review of relevant IRB/IACUC documentation is required prior to commencement of work. Please contact the GSC Principal Investigator(s) to ensure necessary documentation are filed for / made available for timely start of the project.

Investigator Signature:

Investigator Name: **Julia Vipond** **Date December, 2010**