Data Management and sharing plan Principal Investigator / Researcher: Mohlopheni Marakalala
Funder: Wellcome Trust
Institution University of Cape Town

1. Proposal Name

Human TB granulomas as targets for host-directed therapies and prediction of disease progression

2. Introduction/Type of Study

Tuberculosis remains a leading global health problem that is responsible for 1.5 million deaths annually. Better therapeutic and diagnostic tools are urgently needed to combat the disease. Understanding of immune factors associated with TB pathological progression may accelerate discovery of adjunctive host-directed therapies. The goal of this project is to identify and validate protein bio-signatures that are causally or predictively connected to increased risk of pathological damage during infection.

3. Data Outputs and Value

What data outputs will your research generate and what data will have value to other researchers?

The research will generate datasets of protein biosignatures that will be available as an invaluable resource accessible to researchers interested in biomarkers and potential host-directed therapies for TB. The dataset will be available in TXT file or Excel format that is easy to access by other researchers. The data will also be accessible through a link under manuscript figures reporting appropriate research findings.

4. Timeframes for Data Sharing: When will I share the data?

The generated data will be made widely available to user communities at the earliest feasible opportunity or on publication. According to the timeline of my proposed work, I plan to have a dataset that can be shared during third year of funding period (2020) and possibly another data deposit at the end of the fifth year of the funding period (2022).

5. Data Access Where will I make the data available?

The proteomics data will be deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE (PRoteomics IDEntifications) partner repository, which is a centralised, standards compliant, public data repository for proteomics data. The repository has been developed to provide the proteomics community with a public repository for protein and peptide identifications together with the evidence supporting these identifications.

How will others be able to access the data?

On uploading the data to ProteomeXchange Consortium via the PRIDE partner repository, a data set identifier number will be assigned to the deposited data. If possible access to the dataset will be made available to other researchers through a

publication of a resource paper that makes data users to formally cite their usage of the resource.

6. Data Sharing Are any limits to data sharing required?
The are no limits to data sharing.

7. Budget What resources will i require to deliver my plan?
No additional resources are required to deliver the data sharing plan.