Identifying protein subcellular localization in Leishmania using spatial proteomics.

Eden R. Ferreira[1], Adam Dowle[1], Ulrich Dobramysl[2], Richard Wheeler[2], Jeremy Mottram[1]

[1]University of York, [2] University of Oxford.

Leishmania is a flagellated kinetoplastid protozoan and causative agent of Leishmaniasis, a neglected tropical disease with 0.7 to 1 million new cases each year. To understand the biology of Leishmania parasite, information on subcellular localization of proteins is key to making testable predictions about protein function, complex formation, organelle composition and cellular architecture. For this purpose, we are using Localisation of Organelle Proteins by Isotope Tagging after Differential ultracentrifugation (LOPIT-DC). To obtain suitable cell lysates for LOPIT-DC cytoplasmic membrane needs to be ruptured but organelles and compartments should be preserved. In this sense, we successfully obtained cell lysates in eleven fractions by mechanical disruption using nitrogen cavitation associated with differential ultracentrifugation. To test the fractionation effectiveness, we submitted the samples to western blotting using antibodies for proteins known to be localized in different compartments which demonstrated we could successfully separate organelles by weight/volume. Next, fractions from four independent experiments were labelled with TMT-11plex reagents for relative quantification and analysed by LC-MS using an Orbitrap Fusion Tribrid mass spectrometer with multi-notch MS^3 acquisition to minimise chimeric interference. Mass spectrometry analysis and database searching of the *Leishmania mexicana* promastigotes resulted in the identification and relative per-fraction quantification of 3782 proteins common in all four experiments. Using the resulting set of abundance data across all eleven fractions in the four experiments, we employed a Gaussian Mixture model to infer the localisation of previously un-annotated proteins. We trained this model on known protein localisation annotations in *Trypanosoma brucei* (tryptag.org) carried over to *L. mexicana* using gene orthology and thresholding by a sequence similarity of at least 30%. We then added known ribosome, flagellar pocket and Golgi apparatus to the list of *L. mexicana* marker proteins (from TriTrypDB). Dimensionality reduction using t-Stochastic Neighbourhood Embedding (t-SNE) and hierarchical density based clustering (HDBSCAN) showed that we could effectively visualize protein clusters according to the organelle and sub-organelle compartments. This will provide a comprehensive understanding of the proteomic organization, function and evolution including adaptive interaction with host and assumptions on protein-protein interaction within organelles.

This project is part of LeishGEM a Wellcome Trust Collaborative Award.