Developing multi-omic informatics workflows within the Galaxy platform to uncover cancer-microbiome interactions

Pratik Jagtap1, Caleb Easterly1, Subina Mehta1, Praveen Kumar1, Ray Sajulga1, Marie Crane2, Emma Smith1, Berenice Batut3, James Johnson1, Magnus Arntzen4, Timothy Griffin1

1Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota; 2Macalester College, Saint Paul, MN
3Bioinformatics Group, University of Freiburg, Freiburg, Germany; 4Faculty of Chemistry, Biotechnology and Food Science, NMBU, As, Norway

- Microbiomes play a critical role in health and disease in human hosts and in environmental ecosystems. In cancer studies, although host microorganisms have been implicated for their role in cancer progression, challenges exist in understanding the mechanism of cancer host-microbiome interaction.
- We seek to use complementary microbiome -omics techniques to provide functional information by quantifying microbial RNA and protein expression levels. By using these functional microbiome analysis methods along with host expression data, we seek to unravel the microbiome dysbiosis linkage to cancer progression.

METAPROTEOMICS

The Galaxy-P team has developed contemporary metaproteomics workflows that analyzes microbial proteins to enable quantitative analysis of protein functional group expression, taxonomic abundance and taxonomy -function interaction analysis.

METATRANSCRIPTOMICS

For metatranscriptomics data, the ASaIM workflow (https://doi.org/10.1093/gigascience/giy057) was modified to enable assembly, extraction, functional and taxonomy visualization, and taxonomy-functional interaction data analysis.

When functional microbiome expression data is compared with host expression data, we anticipate insights into the mechanistic nuances of complex interactions amongst the host and microbiome during cancer progression.

PROTEIN BOXPLOTS

HOST PROTEOGENOMICS

Transcriptomics workflows within Galaxy are used to generate customized protein databases, estimate gene expression & detect variant genes expressed during cancer progression.

PROTEIN ABUNDANCE VOLCANO PLOTS

TRANSCRIPT ABUNDANCE VOLCANO PLOTS

HOST TRANSCRIPTOMICS

EXAMPLE OUTPUTS