Quantifying Host and Viral Protein Remodeling during HIV Reactivation from Latency

RICARDO ROURE[1,2,3,4], Dain Ryan Brademan[5], Prashant Kaushal[2,3], Ruth Huttenhain[5], and Mehdi Bouhaddou[2,3]

[1] Bruins In Genomics (BIG) Summer Program, University of California Los Angeles, Los Angeles, CA
[2] Department of Microbiology, Immunology, and Molecular Genetics (MIMG), University of California Los Angeles, Los Angeles, CA
[3] Institute for Quantitative and Computational Biosciences (QCBio), University of California Los Angeles, Los Angeles, CA
[4] Department of Microbiology and Cell Science, University of Florida, FL
[5] Department of Molecular & Cellular Physiology, Stanford University, CA

Abstract

After infection, human immunodeficiency virus 1 (HIV-1) enters a latent stage within the host genome. One approach in developing a cure for HIV-1 is reactivating the latent virus, allowing the host immune system to sense and destroy the virus. During reactivation, IV produces viral proteins for transmission by utilizing host proteins and post-translational modifications, of which the full extent remains mysterious. This project seeks to quantitatively compare the host proteins affected during HIV reactivation and identify their pathways. We developed a quantitative analysis pipeline using MSstats to assess quality control of mass spectrometry proteomics and phosphoproteomics during reactivation using Phorbol-12-myristate-13-acetate (PMA), conduct statistical analyses between conditions (e.g. PMA vs Mock), and perform gene set overrepresentation analysis to reveal the biological pathways regulated. Our results from the proteomics showed leukocyte activation is most regulated while the phosphoproteomics revealed the cell adhesion pathway, possibly due to cytoskeleton modifications during egress.

Introduction

- HIV: Human immunodeficiency virus is an retrovirus that infects the host's immune system, specifically the white blood cells. Once infected by the virus, HIV enters a latent stage and integrates its RNA genome into the host cell's genome, where the virus can evade the immune system or any antiretroviral treatments.
- Shock and Kill: A therapeutic strategy using drugs to reactivate latent HIV, allowing the host immune system to destroy the virus and eliminate host reservoirs.
- PMA: Phorbol-12-myristate-13-acetate is a latency reversal agent that prompts HIV to exit the latency stage by activating protein kinase C. High levels of PMA can be toxic to the cells.
- Mass Spectrometry: An analytical tool that uses mass to charge ratio of ions to identify proteins. Mass spectrometry can measure either protein level abundance or phosphorylated peptides in a sample.
  - Abundance Data: Quantification of overall protein levels.
  - Phosphoprotomics Data: Measure of phosphorylated sites after enrichment.
- MSstats: Statistical software in R programming used for pairwise comparison of proteins and peptides from two or more conditions.
- Gene Set Overrepresentation Analysis: A statistical method to determine the genes that are over or under represented in a dataset and the biological pathways these genes are involved in.

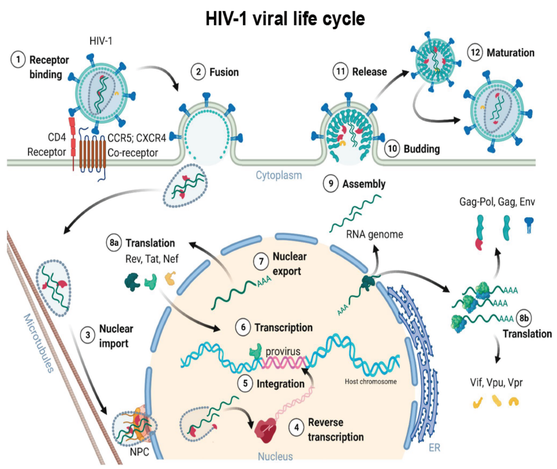
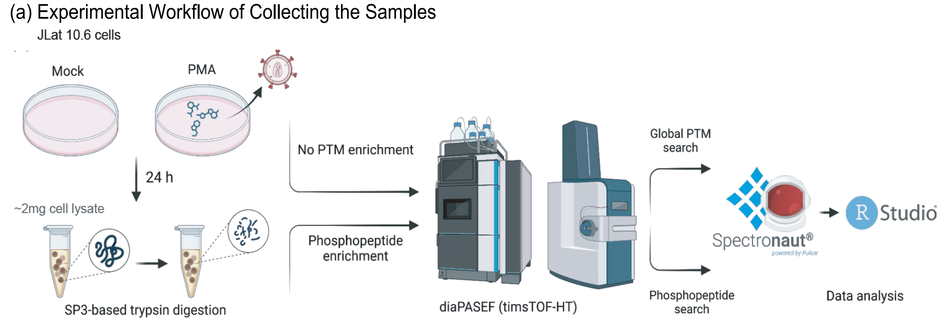


Fig. 1. Depiction of HIV using host machinary for post translational modification during HIV reactivation. (Reproduced from Heuvel et al., 2022).

Methodology



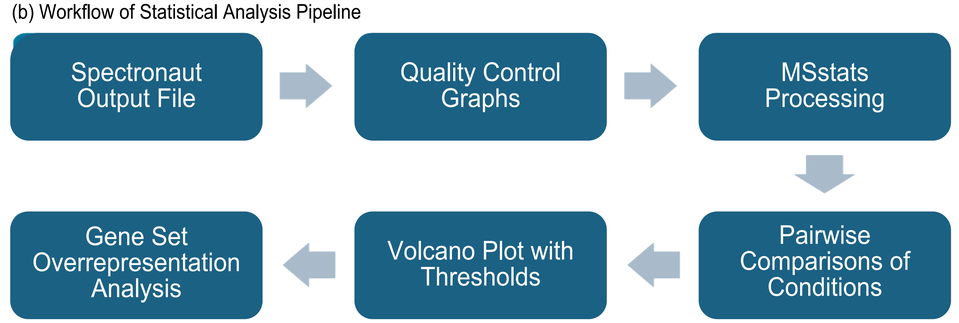
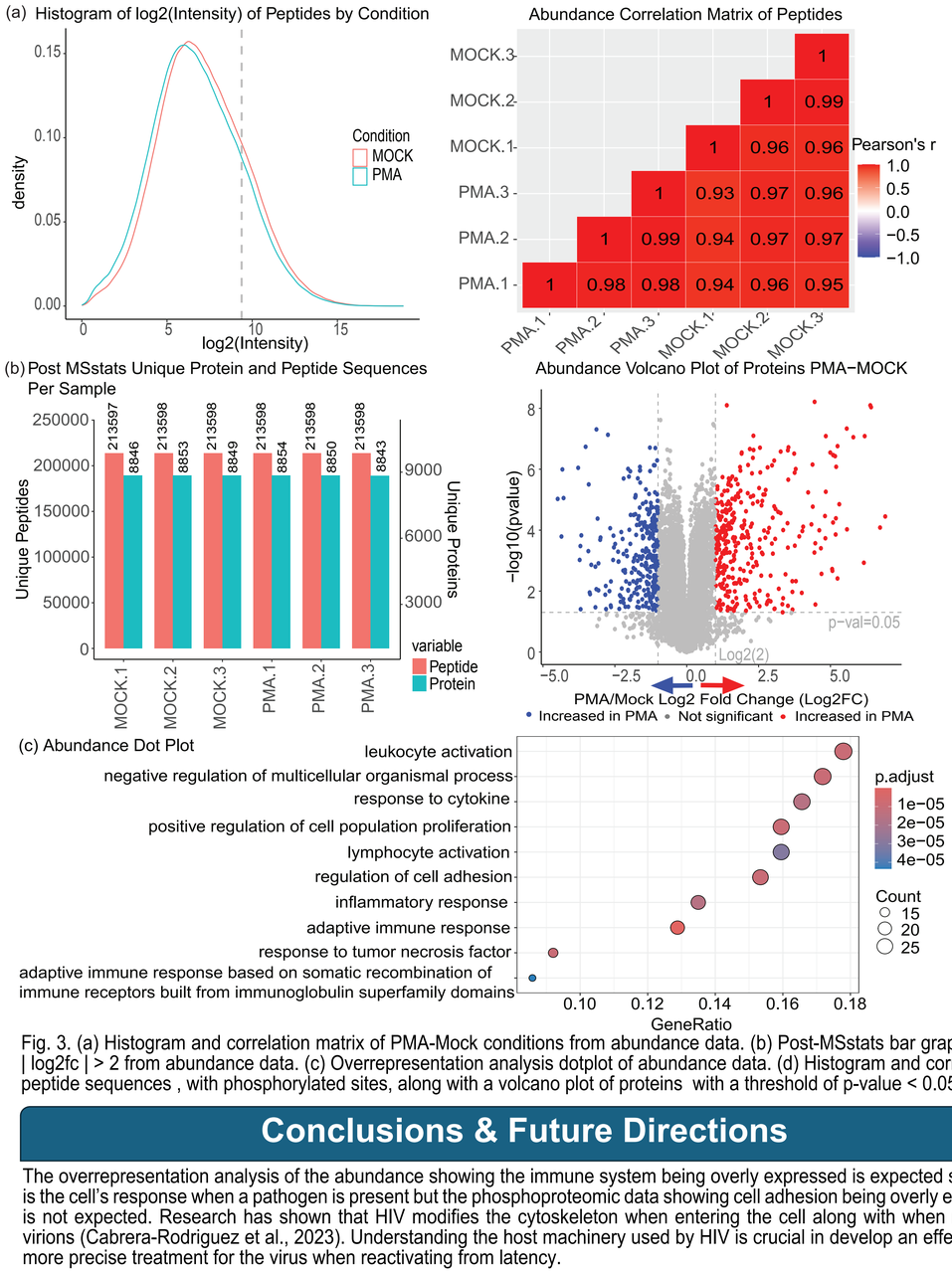


Fig. 2. (a) Methodology of preparing the conditions of the samples and performing abundance and phosphoproteomic analysis with mass spectrometry. (b) Workflow of the quantitative analysis pipeline starting from the output file of the mass spectrometer to the volcano plot of the statistically significant proteins and overrepresentation analysis of the associated genes.

Results

Abundance Data



Phosphoproteomics Data

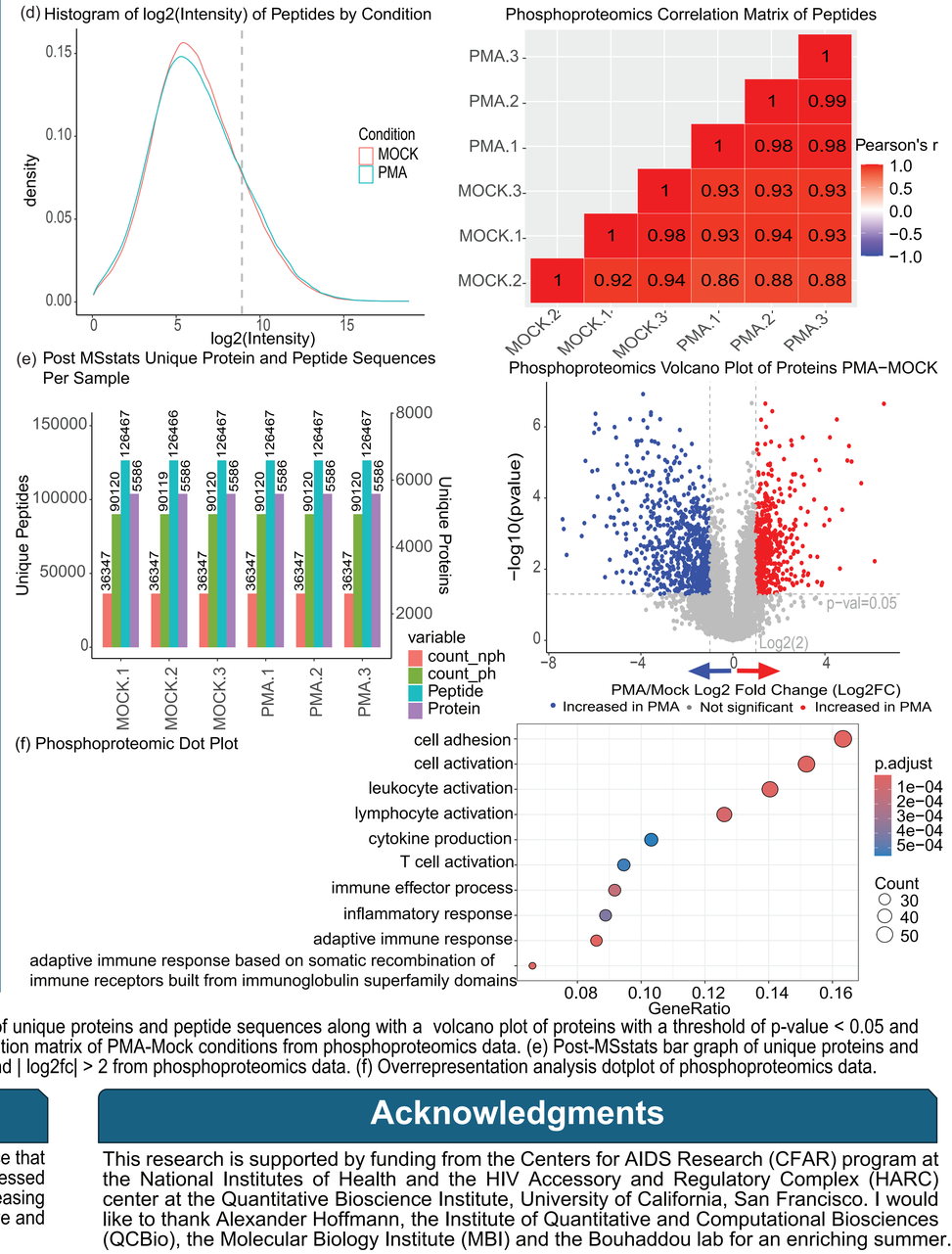


Fig. 3. (a) Histogram and correlation matrix of PMA-Mock conditions from abundance data. (b) Post-MSstats bar graph of unique proteins and peptide sequences along with a volcano plot of proteins with a threshold of p-value < 0.05 and | log2fc | > 2 from abundance data. (c) Overrepresentation analysis dotplot of abundance data. (d) Histogram and correlation matrix of PMA-Mock conditions from phosphoproteomics data. (e) Post-MSstats bar graph of unique proteins and peptide sequences , with phosphorylated sites, along with a volcano plot of proteins with a threshold of p-value < 0.05 and | log2fc| > 2 from phosphoproteomics data. (f) Overrepresentation analysis dotplot of phosphoproteomics data.

Conclusions & Future Directions

The overrepresentation analysis of the abundance showing the immune system being overly expressed is expected since that is the cell's response when a pathogen is present but the phosphoproteomic data showing cell adhesion being overly expressed is not expected. Research has shown that HIV modifies the cytoskeleton when entering the cell along with when releasing virions (Cabrera-Rodriguez et al., 2023). Understanding the host machinery used by HIV is crucial in develop an effective and more precise treatment for the virus when reactivating from latency.

Acknowledgments

This research is supported by funding from the Centers for AIDS Research (CFAR) program at the National Institutes of Health and the HIV Accessory and Regulatory Complex (HARC) center at the Quantitative Bioscience Institute, University of California, San Francisco. I would like to thank Alexander Hoffmann, the Institute of Quantitative and Computational Biosciences (QCBio), the Molecular Biology Institute (MBI) and the Bouhaddou lab for an enriching summer.