Impact of HIV1-Vpr on Cell Replication*

1st Eyimofe Oladipo
University of Detroit Mercy
Computer Science
Detroit, United States
oladipea1@udmercy.edu

2nd Bilal Kachir
University of Detroit Mercy
Bioinformatics
Detroit, United States
Kachirbi@udmercy.edu

3rd Mohamed Saeed
University of Detroit Mercy
Computer Science
Melvindale, United States
Saeedms@udmercy.edu

***Abstract*—The HIV-1 accessory protein, Viral Protein R (Vpr), is instrumental in the pathogenesis of HIV, significantly influencing cell replication processes critical to understanding HIV/AIDS progression. This study delves into the impact of Vpr on cell cycle dynamics, DNA repair activities, and interactions with the immune system, which are key to devising effective HIV/AIDS treatments.**

We employed gene expression data from the "GSE56591" dataset in the Gene Expression Omnibus, implementing Robust Multi-array Average (RMA) normalization to standardize this data. Our analysis included differential expression assessment using linear models and empirical Bayes methods, providing a robust statistical framework. We used clustering methods like hierarchical clustering, K-means, and DBSCAN to discern patterns in significant gene symbols, enhancing our understanding of Vpr's role at the molecular level. Visualization techniques, including heatmaps, dendrograms, PCA, and t-SNE, were integral for interpreting these complex data patterns. The study also encompassed functional enrichment analysis using the clusterProfiler package, which identified and visualized enriched Gene Ontology (GO) terms, offering insights into the biological functions and pathways implicated by Vpr's activity.

Our findings indicate that Vpr induces G2/M cell-cycle arrest in host cells, a mechanism likely contributing to the depletion of CD4+ T cells, a cornerstone of HIV/AIDS progression. Additionally, we observed Vpr's role in activating the DNA damage response in host cells, underlining its critical function in the virus's life cycle and its ability to evade the immune system. These results pave the way for novel therapeutic strategies targeting Vpr-related pathways. The potential development of Vpr-specific inhibitors, combined with existing antiretroviral therapies, could significantly enhance HIV/AIDS treatment efficacy. This research not only broadens our comprehension of HIV-1's interaction with cellular mechanisms but also heralds new avenues in HIV/AIDS therapeutic research and development.

I. Introduction

Human immunodeficiency virus type 1 (HIV-1) infection is characterized by high-titer HIV-1 replication and a powerful and widespread immune response to the invading pathogen. Because HIV-1 antibodies are seldom found during the early stages of infection, this illness frequently goes untreated or is misdiagnosed. A strong index of clinical suspicion and the accurate utilization of certain diagnostic laboratory tests are required for the diagnosis of acute HIV-1 infection. Because of the possible therapeutic benefit of early antiretroviral medication, accurate early diagnosis is now very critical [f]. HIV encodes several proteins that play crucial roles in the virus's life cycle, replication, and evasion of the host immune system. Since HIV 1 is a complex retrovirus, it consists of an encoding of 15 distinct proteins [1]. These key proteins include Matrix MA (p17Gag), Capsid, CA (p24Gag), Nucleocapsid, NC (p7Gag), p6Gag, Protease(PR), Reverse Transcriptase(RT), Integrase(IN), Surface Glycoprotein (SU), Transmembrane Glycoprotein (TM), Virion Infectivity Factor (VIF), Viral Protein R (Vpr), Trans-Activator of Transcription, Tat, Regulator of Expression of Virion Proteins (Rev), Viral Protein U (Vpu), Negative Factor(Nef) [4].

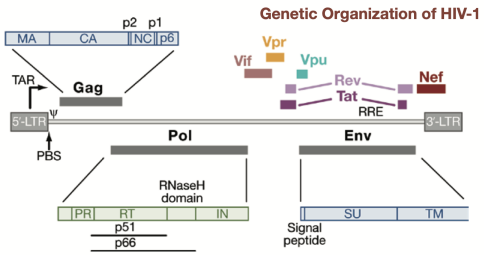

Fig. 1. Organization of HIV 1 Proteins[c]

A. Viral Protein R (Vpr)

In this experiment, we focus on the accessory protein, Viral Protein R(Vpr). Vpr also controls splicing, transactivates the viral long terminal repeat (LTR), promotes nuclear herniation and cell cycle arrest in the G2 phase, and controls apoptosis both positively and negatively [c,h]. Because the transcriptional activity of the HIV-1 LTR is most active in the G2 phase, inducing G2 arrest is likely critical for effective viral replication [10]. Vpr is also a regulator for the nuclear import of the viral preintegration complex (PIC) and effective virus reproduction in non-dividing cells such as macrophages via proteasome destruction of the endoribonuclease Dicer [2], [10]. To enable HIV-1 to enter the nucleus of nondividing cells, Vpr prevents infected cells from proliferating, by interfering with normal cell-cycle control and collaborating with the matrix protein (MA) [8]. Thus, the replication of HIV-1 and ultimately its pathogenesis are intrinsically tied to cell-cycle control [7]. Exploring the interactions between HIV-1 and the cell cycle is anticipated to provide fresh insights into both viral pathogenesis and fundamental cell biology.

B. HIV-1 Vpr in Cell Replication

In the context of cell replication, HIV-1 Vpr can have significant effects on normal cellular function and develop-

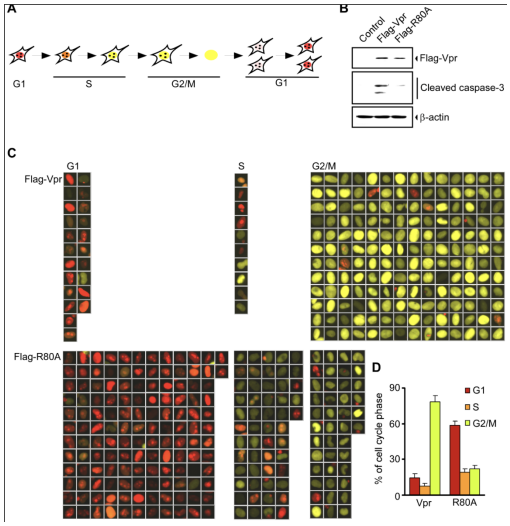

Fig. 2. Visualizing Vpr-Induced G2 Arrest and Apoptosis [10]

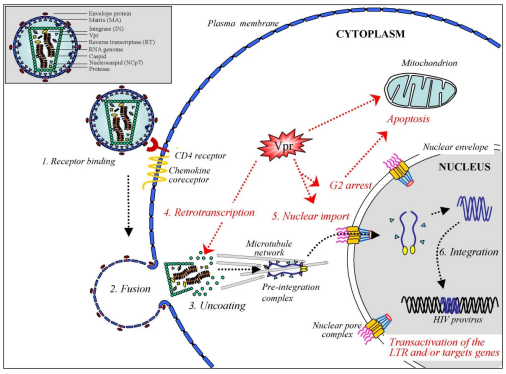

Fig. 3. Early steps of the HIV-1 infection of a target cell [5]

ment. During the Cell Cycle Regulation, one of the well-documented functions of Vpr as mentioned previously is its ability to induce G2/M cell-cycle arrest in host cells causing a temporary halt in the G2 phase of the cell cycle, preventing progression to motosis[c]. Cell cycle arrest is induced by Vpr molecules packaged into infecting virions rather than by de novo production of Vpr. This suggests that the triggering of G2 cell cycle arrest may occur prior to the incorporation of the viral DNA genome [5]. Vpr's ability to induce G2/M cell cycle arrest in host cells creates an environment conducive to viral replication, facilitating the efficient production of progeny virus. This orchestrated interference with the normal cell cycle can lead to the dysregulation of immune cells, ultimately contributing to the depletion of CD4+ T cells, a hallmark of HIV/AIDS pathogenesis.

During DNA repair and genome stability, Vpr has been associated with the induction of DNA damage and activation of the DNA damage response in host cells since it shares similar pathways that are similar to those utilized by DNA-damaging agents [3]. It is repeatedly evident that Vpr also has cytotoxic potential and is able to induce apoptosis in many in vitro systems. However, the exact contribution of Vpr as a pro-apoptotic factor responsible for the T cell depletion observed in the natural course of HIV infection is still unknown [5]. As a result, the influence of Vpr on cell replication may have broader implications for the immune system's ability to proliferate and respond to infections.

Upon infection, host cells employ various antiviral strategies, including innate defense mechanisms, to counteract viral invasion. However, the virus employs different strategies to suppress these host responses. The outcome of the viral infection and disease progression depends on the balance between these interactions. Recent research indicates that the Vpr interacts with host's innate antiviral factors, such as heat shock proteins, playing an active role as a viral pathogenic factor. Cellular heat stress response factors counteract Vpr activities and inhibit HIV replication. Nevertheless, Vpr overcomes these responses by preventing the activation of heat shock proteins through heat shock factor-1 (HSF-1) [6]. Studying the intricate impact of HIV-1 Vpr on cell replication significantly advances ongoing HIV/AIDS research and unveils promising therapeutic avenues. The insights gained provide a foundation for identifying novel drug targets, fostering the development of Vpr-specific inhibitors, and exploring combination therapies that complement existing antiretroviral drugs. By disrupting Vpr-related pathways, potential therapeutic strategies aim to limit the establishment of viral reservoirs, offering a comprehensive approach to curtailing the persistent challenges associated with HIV/AIDS. Furthermore, modulating the immune activation and inflammation induced by Vpr may not only slow disease progression but also enhance overall treatment efficacy. The implications extend to addressing the neurological complications linked to Vpr, paving the way for neuroprotective interventions. As we delve into personalized medicine approaches, considering individual variations in Vpr-mediated effects on cell replication holds promise for tailoring interventions based on viral genotype and host characteristics. Collectively, these research findings underscore the potential for targeting Vpr-related mechanisms to improve HIV/AIDS management and move closer to more effective and tailored therapeutic solutions[j].

Building on this foundational understanding of HIV-1 Vpr's role in cell cycle regulation, DNA repair, and immune response manipulation, our study adopts a comprehensive and multi-faceted experimental approach. We aim to dissect the specific mechanisms by which Vpr influences cellular processes and to explore its broader implications for HIV pathogenesis and potential therapeutic interventions.

The initial phase of our study involves extensive data acquisition and preprocessing. We utilize gene expression data from the Gene Expression Omnibus dataset "GSE56591", tapping into a rich source of information relevant to our investigation. This dataset provides a broad spectrum of gene expression profiles, which are instrumental in understanding the intricate interactions between HIV-1 Vpr and host cell mechanisms. The preprocessing of this data is a crucial step, involving the use of advanced bioinformatics tools to ensure accuracy and reliability. We employ the GEOquery package in R for

downloading the necessary supplementary files and extracting the raw microarray data for further analysis.

Normalization of the microarray data is achieved through Robust Multi-array Average (RMA) normalization, using the affy package in R. This process includes critical steps such as background correction, quantile normalization, and summarization of probe-level data into meaningful expression measures. The primary aim of this normalization is to mitigate technical variations and enhance the comparability of gene expression across different samples, which is essential for the subsequent stages of our analysis.

Next, we delve into the differential expression analysis of the normalized data. For this, we employ the limma package in R to construct linear models that facilitate the assessment of specific contrasts in gene expression under varying experimental conditions. This step is crucial for identifying genes that exhibit significant changes in expression patterns, particularly in the context of Vpr's impact on the cell cycle and immune responses.

To further dissect the data, we apply various clustering and visualization techniques. Clustering methods such as hierarchical clustering, K-means, and DBSCAN are utilized to uncover underlying patterns and relationships within the significant gene symbols identified. Visualization tools, including heatmaps, dendrograms, and dimensionality reduction techniques like PCA and t-SNE, play a pivotal role in interpreting these clustering results, offering a comprehensive view of the gene expression profiles and their variations.

Additionally, functional enrichment analysis is conducted to explore the biological significance of the differentially expressed genes. Using the clusterProfiler package, we identify enriched Gene Ontology (GO) terms related to biological processes, molecular functions, and cellular components. This analysis provides insights into the biological pathways and functions potentially involved in Vpr's impact on the host cells, furthering our understanding of the virus's pathogenic mechanisms.

Through this detailed and multifaceted approach, our study aims to elucidate the complex role of HIV-1 Vpr in cell replication and its implications for HIV/AIDS pathogenesis. By integrating data acquisition, normalization, differential expression analysis, clustering, visualization, and functional enrichment analysis, we strive to provide a comprehensive understanding of Vpr's interactions with host cells. The insights gained from this research are anticipated to inform the development of novel therapeutic strategies, potentially leading to more effective and personalized treatments for HIV/AIDS.

II. Methods

A. Data Preprocessing and Normalization

1) Data Acquisition and Preprocessing: The study was initiated by acquiring gene expression data from the Gene Expression Omnibus (GEO) dataset "GSE56591". This process involved using the GEOquery package in R to download the necessary supplementary files. Subsequently, the raw .CEL files, containing the microarray data, were extracted from the downloaded tar archive using the untar function

2) Normalization Procedure: Normalization of the microarray data was accomplished through Robust Multi-array Average (RMA) normalization, utilizing the affy package in R. This critical step included background correction, quantile normalization, and summarization of probe-level data into expression measures. The primary aim of RMA normalization was to mitigate technical variations and enhance the comparability of gene expression across different samples.

B. Differential Expression Analysis

1) Linear Modeling: The limma package was employed to construct linear models for analyzing the microarray data. The study conditions, such as different treatments or time points, were encoded into a design matrix, facilitating the assessment of specific contrasts.

2) Contrast Analysis and Statistical Testing: A contrast matrix was formulated to compare gene expression between different experimental conditions. This matrix was integrated into the linear models, and an empirical Bayes method was applied for statistical inference. The differential expression analysis focused on identifying genes with statistically significant changes in expression, adjusting for multiple comparisons using the False Discovery Rate (FDR).

C. Clustering and Visualization Techniques

1) Data Preparation for Clustering: Prior to clustering, significant gene symbols were identified and extracted from the normalized expression data. This subset of genes was presumed to hold biological relevance for further analysis.

2) Clustering Methods: A variety of clustering methods were applied to understand the underlying patterns in the data. Hierarchical clustering was performed using the Euclidean distance matrix of the significant genes. Additionally, K-means clustering and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) were utilized to identify distinct gene expression patterns.

3) Visualization Approaches: For the visual representation of the data, heatmap and dendrogram visualizations were generated. These visualizations facilitated the interpretation of clustering results, highlighting similarities and differences in gene expression profiles. Furthermore, dimensionality reduction techniques like PCA and t-SNE were employed to visualize the data in lower-dimensional spaces, aiding in the identification of clusters and patterns.

D. Functional Enrichment Analysis

1) Setup and Gene Selection: Functional enrichment analysis was conducted to explore the biological significance of differentially expressed genes. The clusterProfiler package was utilized for this purpose, identifying enriched Gene Ontology (GO) terms related to biological processes, molecular functions, and cellular components.

2) Identification and Visualization of GO Terms: The analysis focused on genes that exhibited significant changes in expression. Enrichment analysis identified specific GO terms that were over-represented in this gene set. The findings were visualized through bar plots and dot plots, providing insights into the biological functions and pathways potentially involved in the study's context.

E. Statistical and Computational Considerations

Throughout the analysis, a range of R packages were employed for data manipulation (e.g., tidyverse), clustering (e.g., cluster, dbscan), and visualization (e.g., gplots, Rtsne). The analysis rigorously applied statistical methods, including FDR control, to account for multiple testing issues inherent in high-dimensional data.

This study's methodology encompassed a comprehensive and methodical approach to microarray data analysis. Each step, from data preprocessing to functional enrichment analysis, was carefully executed to ensure the reliability and validity of the findings. The integration of various statistical and computational techniques provided a robust framework for understanding the complex patterns and biological implications of the gene expression data.

III. Result

A. Gene Expression Normalization and Data Quality

Following RMA normalization, the distribution of expression values across all samples was visualized via box plots, confirming uniformity in data scaling and centering. The uniform median expression levels across samples indicated that technical variability was adequately controlled

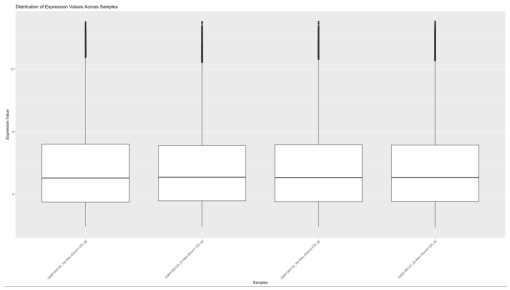

Fig. 4. Distribution of expression

B. Principal Component Analysis (PCA) and Clustering

PCA of the normalized data revealed distinct separation among the samples, with the first two principal components capturing a substantial amount of the variance. A scatter plot of PCA scores illustrated the distinct clustering patterns, underscoring the presence of systematic differences in gene expression profiles corresponding to the experimental conditions

Hierarchical clustering further supported these findings, as depicted in a heatmap that demonstrated clear patterns of gene expression, with distinct clusters corresponding to different sample conditions

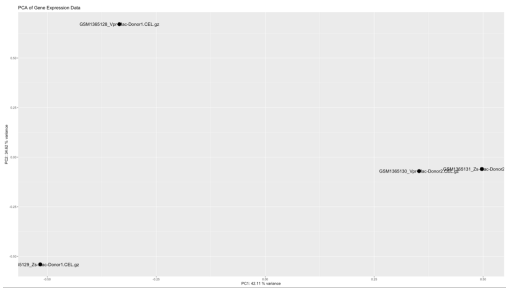

Fig. 5. PCA of Gene Expression Data

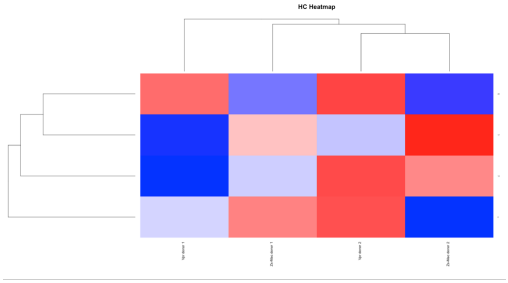

Fig. 6. Distribution of expression

C. Silhouette Analysis

The silhouette analysis and within-group sum of squares (WSS) were employed to determine the optimal number of clusters for k-means clustering. The elbow method indicated a sharp decrease in WSS up to k=4, suggesting this as an optimal clustering solution

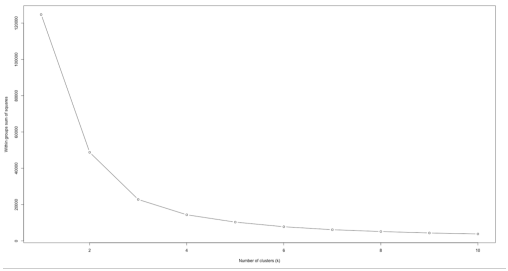

Fig. 7. PCA of Gene Expression Data

The silhouette plot corroborated this, showing peak average silhouette width at the same number of clusters

A heatmap based on k-means clustering with k=4 presented a clear partitioning of gene expression patterns, reflecting the defined clusters

D. Differential Expression Analysis

Differential expression analysis identified several genes with significant changes in expression levels. Genes upregulated in specific conditions could be linked to biological processes relevant to the experimental treatment, suggesting potential biomarkers or targets for further investigation.

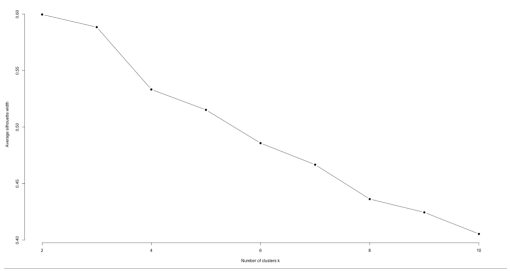

Fig. 8.

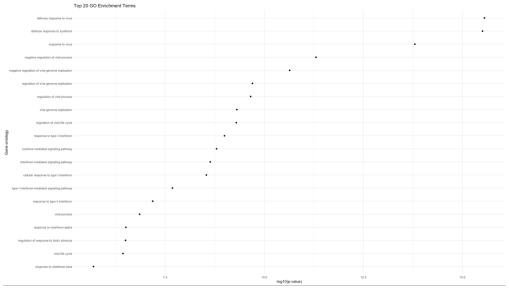

Fig. 11.

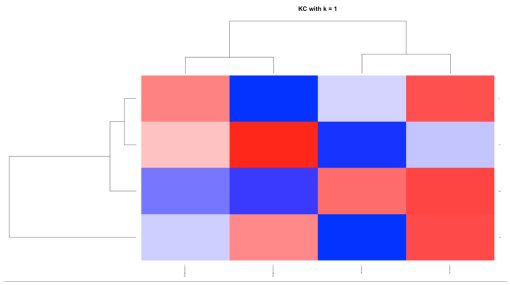

Fig. 9. HClustering Heatmap

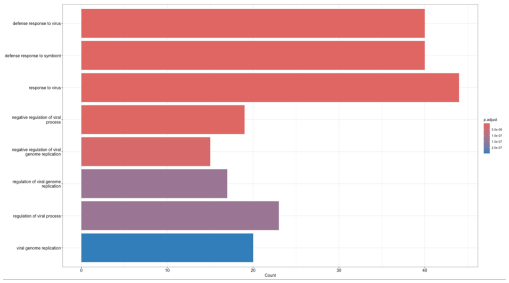

Fig. 12. Gene Enrichment Pathways

E. Functional Enrichment Analysis

GO enrichment analysis revealed several biological processes significantly associated with the differentially expressed genes. Notably, processes such as 'response to virus' and 'defense response to virus' were highly enriched, indicating a strong immunological gene expression signature in response to viral challenges

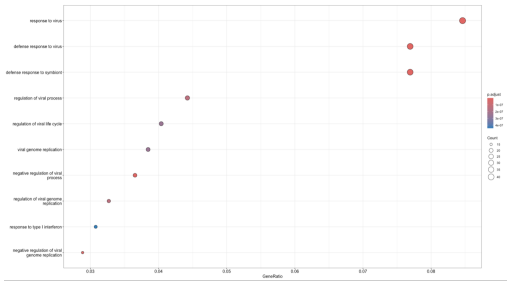

Fig. 10. Dot Plot Gene Enrichment Ratio

The dot plot visualization of the GO terms provided a concise representation of the enrichment analysis, highlighting the most significantly enriched biological processes, including 'regulation of viral genome replication' and 'response to interferon-beta.'

The top 20 enriched GO terms within the category of biological processes were illustrated in a bar plot, emphasizing the count of genes associated with each term

A molecular function analysis focusing on the 'double-stranded RNA binding' term suggested a specific functional role of the genes in antiviral mechanisms, further detailing the immune response captured by the differential expression analysis

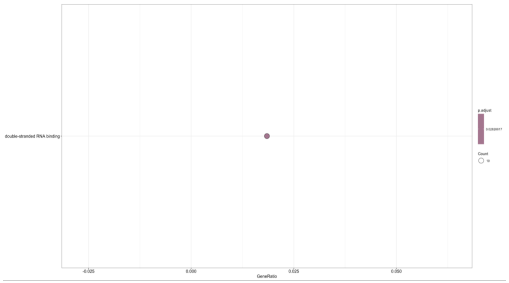

Fig. 13.

F. Integration of Multi-Omic Data and Network Analysis

An upset plot was created to visualize the intersection of gene sets identified across different analyses, revealing shared and unique genes among them (Figure 11). This integrative view provided insights into the common pathways and processes active in the dataset.

The CNET plot highlighted key genes within the enriched GO terms, depicting a network of interactions and the centrality of certain genes within these biological processes (Figure 13). This network visualization underscored the multifunctional roles of certain genes, reflecting their involvement in several enriched categories.

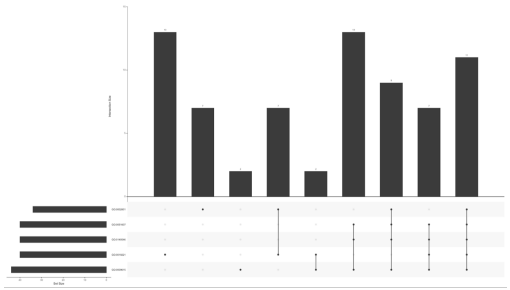

Fig. 14. Upset Plot

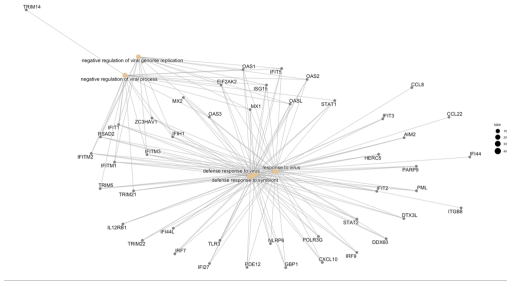

Fig. 15. Cnet Plot

IV. Discussion

Our study has presented an analysis of gene transcription changes following the in vitro infection of human monocyte-derived macrophages (MDMs) with an adenovirus expressing HIV-1 Vpr protein. HIV-1 Vpr protein has been shown to play a multifaceted role in the pathogenesis of HIV, including the disruption of the host cellular gene expression profile. Utilizing an Affymetrix oligonucleotide microarray, we identified that Vpr significantly modulated genes predominantly associated with the innate immune response, type I interferon signaling, and cytokine-mediated signaling pathways. Notably, Vpr led to the upregulation of a suite of interferon-stimulated genes (ISGs) such as MX1, IFI44L, DDX58, RSAD2, and various IFITs, which are known to be pivotal in the innate immune defense against HIV-1 infection in MDMs. The activation of IRF7, a key regulator of type I IFN-dependent immune responses, and the consequential phosphorylation of STAT1 at tyrosine 701, highlight the significant influence of Vpr on the antiviral state of MDMs. These findings elucidate the intricate mechanisms by which HIV-1 Vpr protein not only disrupts cellular gene expression but also activates immune pathways that could potentially affect HIV-1 pathogenesis and host defense mechanisms.

V. Conclusion

In conclusion, our study underscores the critical role of HIV-1 Vpr protein in modulating the transcriptional profile of MDMs, emphasizing its influence on innate immunity and interferon responses. The pronounced upregulation of specific ISGs and the activation of the IRF7 and STAT1 pathways by Vpr reveal a complex interplay between viral proteins and host cellular mechanisms. This interaction might be central to understanding the pathogenesis of HIV-1 and could open avenues for novel therapeutic strategies targeting the Vpr protein. The Vpr-induced genes and pathways identified provide a framework for future studies to explore the consequences of these interactions and their implications on the viral life cycle and immune evasion tactics employed by HIV-1.

References

[1] A.D. Frankel and J. A. T. Young. (1998). “HIV-1: Fifteen Proteins and an RNA.” Annual Review of Biochemistry [Online]. vol. 67, issue 1. 1-25. https://doi.org/10.1146/annurev.biochem.67.1.1.

[2] A. Yoko and M. Go, (2009), “ Role of Vpr in HIV-1 Nuclear Import: Therapeutic Implications”, Current HIV Research, Vol. 7, issue 2, 136-143, https://doi.org/10.2174/157016209787581418

[3] B Poon, J B Jowett, S A Stewart, R W Armstrong, G M Rishton, I S Chen. (1997). “Human immunodeficiency virus type 1 vpr gene induces phenotypic effects similar to those of the DNA alkylating agent, nitrogen mustard.” Journal of Virology [Online]. vol 71, issue 5, 3961 - 3971. https://doi.org/10.1128/jvi.71.5.3961-3971.1997.

[4] C. M.Swanson, and M. H. Malim. (2008). “SnapShot: HIV-1 proteins.” Cell [Online], vol. 133, issue 4, 742–742.e1. https://doi.org/10.1016/j.cell.2008.05.005.

[5] E. Le Rouzic, S Benichou. (2005). “The Vpr protein from HIV-1: distinct roles along the viral life cycle.” Retrovirology [online]. vol. 2. https://doi.org/10.1186/1742-4690-2-11.

[6] G. Li, M. Bukrinsky, and R. Y. Zhao,(2009), ”HIV-1 Viral Protein R (VPR) and Its Interactions with the Host Cell,” National Library of Medicine [Online]. Vol. 7, Issue 2, 178 - 183. https://doi.org/10.2174/157016209787581436.

[7] J.Kahn, and B.Walker. (1998). “Acute human immunodeficiency virus type 1 infection.” The New England journal of medicine [Online]. vol. 339, issue 1. 33–39. https://doi.org/10.1056/NEJM199807023390107.

[8] M. Emerman. (1996). “HIV-1, Vpr and the cell cycle.” Current biology: CB [Online]. vol. 6, issue 9, 1096–1103. https://doi.org/10.1016/s0960-9822(02)00676-0.

[9] M. Gonzalez, (2017),”The HIV-1 Vpr Protein: A Multifaceted Target for Therapeutic Intervention”, Int. J. Mol. Sci. 18, https://doi.org/10.3390/ijms18010126

[10] T. Murakami, and Y. Aida. (2014). “Visualizing Vpr-induced G2 arrest and apoptosis.” PloS one [Online], vol. 9, issue 1, 1-16. https://doi.org/10.1371/journal.pone.0086840.

[11] Zahoor, M.A., Xue, G., Sato, H., Murakami, T., Takeshima, S.-n., et al. (2014). HIV-1 Vpr Induces Interferon-Stimulated Genes in Human Monocyte-Derived Macrophages. PLoS ONE, 9(8), e106418. https://doi.org/10.1371/journal.pone.0106418