Cross-Tissue Biomarkers for Hepatitis B Virus-Related Hepatocellular Carcinoma

Authors: Shubham Thakur, Mayurakshi Mukherji, Fenil Parmar, Saket Choudhary

Introduction

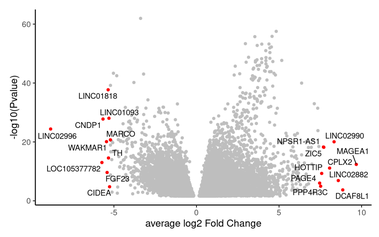
Hepatitis B Virus (HBV) infection is a leading cause of hepatocellular carcinoma (HCC) worldwide. The development of cross-tissue biomarkers enables non-invasive diagnostics for early detection of HBV-related HCC. This study aims to identify gene expression patterns in blood and liver tissues to explore potential biomarkers detectable in blood, reducing the need for liver biopsies.

Methodology

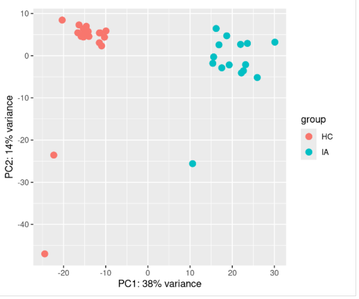
- **Data Collection**: RNA-seq data from public sources, including healthy, HBV-infected, and HBV-HCC samples.
- **Transcript Quantification**: Kallisto was used for pseudo-alignment and transcript abundance estimation.
- **Differential Expression**: DESeq2 identified differentially expressed genes (DEGs) in blood and liver.
- **Visualization**: PCA and clustering heatmaps compared gene expression patterns across tissues.
- **Functional Enrichment**: ClusterProfiler analyzed DEGs for pathways linked to immune response and metabolism.
- **Cross-Tissue Comparison**: Identified biomarkers consistently expressed in both blood and liver for non-invasive HBV-HCC detection.

Visualization

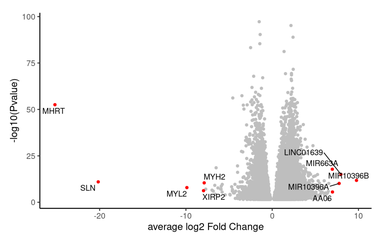
(a) PCA for Dataset 1



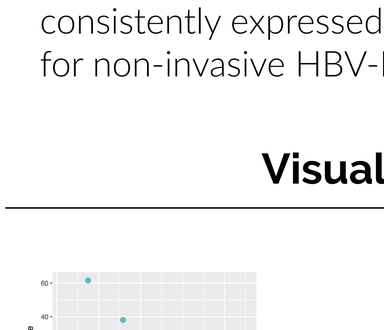
(b) Volcano Plot for Dataset 1



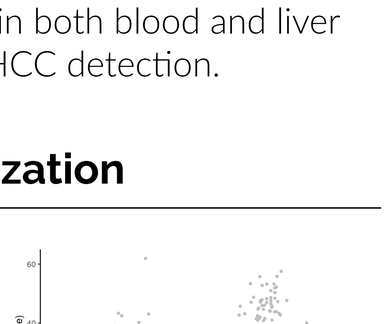
(c) PCA for Dataset 2



(d) Volcano Plot for Dataset 2



(e) PCA for Dataset 3

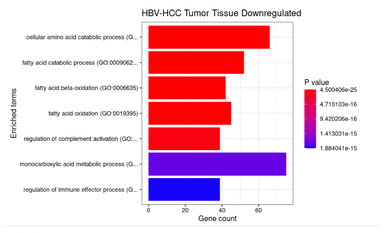


(f) Volcano Plot for Dataset 3

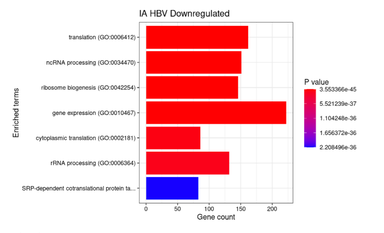
Figure: PCA plots and Volcano Plot for each of the three datasets, demonstrating gene expression patterns.

In these PCA and Volcano plots, we observe unique gene expression patterns across the datasets, indicating significant differential expression and clustering under varied conditions. This suggests potential key genes and molecular pathways that we plan to explore further.

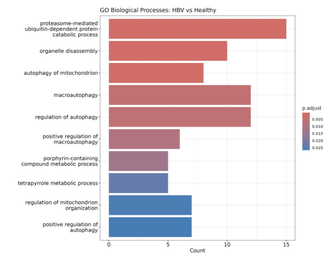
Functional Enrichment



(a) Enrichment for Dataset 1



(b) Enrichment for Dataset 2

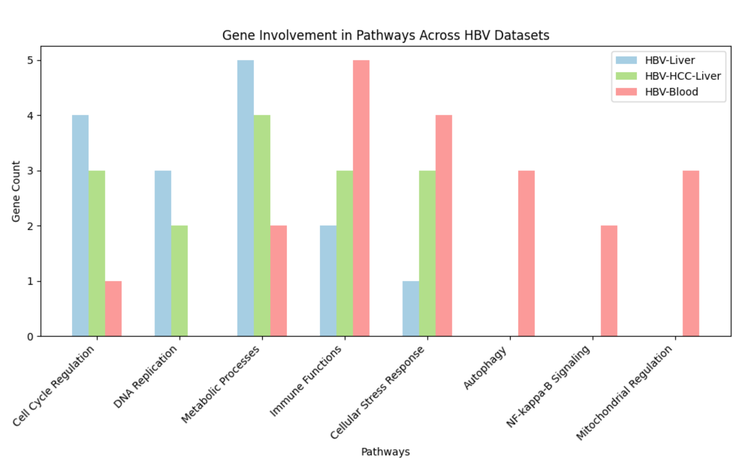


(c) Enrichment for Dataset 3

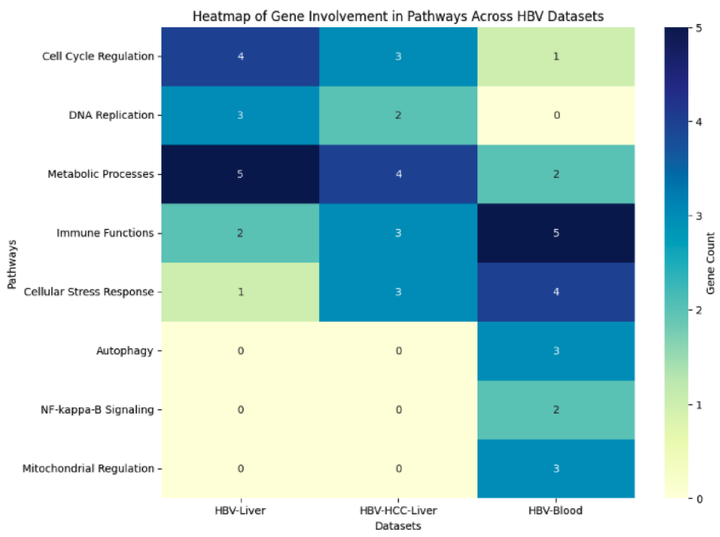
Figure: Functional enrichment plots showing key pathways for each dataset.

Pathway Visualization

Significant gene expression changes related to HBV and HBV-HCC are detectable in both liver and blood samples. This highlights the potential for identifying cross-tissue biomarkers, enabling the development of non-invasive diagnostic tools for HBV-related HCC.



(a) Pathway Visualization - Plot 1



(b) Pathway Visualization - Plot 2

Figure: Visualizations highlighting common pathways across blood and liver datasets.

From these figures, "Metabolic Processes" and "Immune Functions" appear as key cross-tissue biomarkers, while "NF-kappa B Signaling" and "Mitochondrial Regulation" show unique activity in blood samples, highlighting their potential for non-invasive diagnostics and disease progression insights.

Conclusion

This study reveals significant cross-tissue gene expression changes in HBV and HBV-HCC, enabling us to identify non-invasive blood-based biomarkers. By validating RNA-seq as a diagnostic tool, we establish blood as a viable alternative to liver biopsies, offering potential for early detection, personalized treatments, and improved disease monitoring.

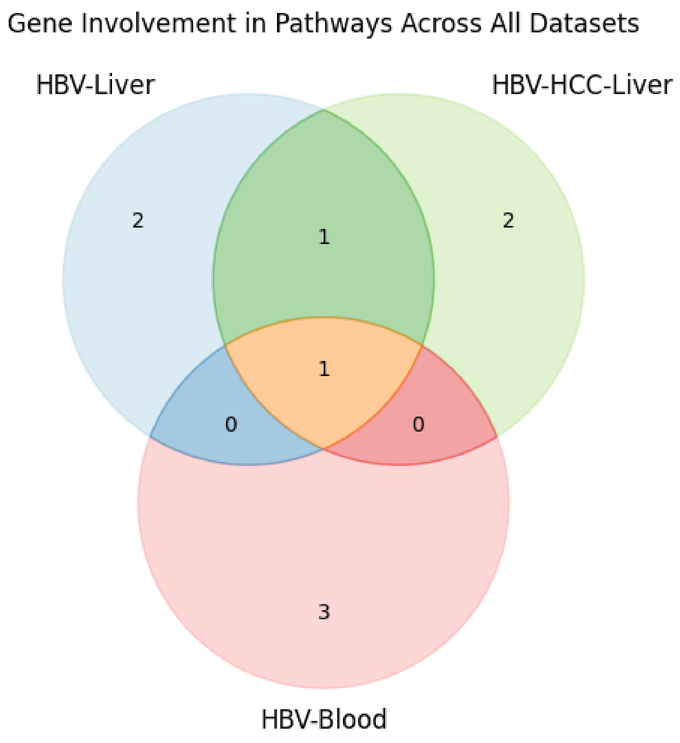


Figure: Illustrative representation of blood-based biomarkers for HBV-HCC detection.

Limitations

- Potential batch effects due to multi-source data integration.
- Kallisto may not capture novel transcripts absent from the reference transcriptome.
- Further validation is required across larger, diverse patient cohorts.

Applications

- **Non-Invasive Diagnostic Tool**: Potential for blood-based tests as alternatives to liver biopsies for early HBV-HCC detection.
- **Public Health Screening**: Screening programs could leverage these biomarkers in HBV-prevalent regions.
- **Research in HBV-HCC Pathogenesis**: Insights into disease progression and potential therapeutic targets.

Data Availability

The RNA-seq data used in this study is publicly available from NCBI Gene Expression Omnibus (GEO)].

- Dataset 1 : `https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE94660`
- Dataset 2 : `https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE230397`
- Dataset 3 : `https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE236281`

.