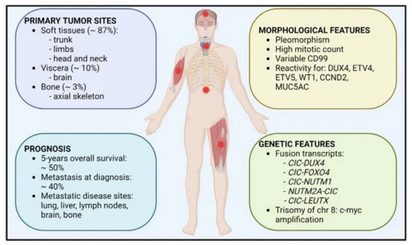
Exploring CIC-DUX4 Sarcoma Surfaceome by RNA-Sequencing

Wenxuan Zhu[1], Lauren Mills[2], Darko Bosnakovski[2]

[1] Department of MSCS, Macalester College, St. Paul, MN; [2] Department of Pediatrics, University of Minnesota, Minneapolis, MN

Abstract

CIC-DUX4 sarcoma (CDS) is an aggressive and extremely rare subtype of undifferentiated round-cell sarcoma that predominantly impacts individuals in the adolescent and young adult age group with a median survival of less than 2 years. CDS is defined by the CIC-DUX4 oncogene, a translocation between CIC and DUX4. It is poorly understood at the molecular and cellular levels, with knowing next to nothing about the CDS surfaceome, surface membrane proteins. Knowing the cancer's surfaceome is important for understanding how tumors interact with their surroundings and could be the basis of developing targeted treatments.



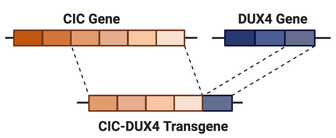
Mancarella, Caterina, Marianna Carrabotta, Lisa Toracchio, and Katia Scotlandi. "CIC-Rearranged Sarcomas: An Intriguing Entity That May Lead the Way to the Comprehension of More Common Cancers." MDPI, November 2, 2022. https://www.mdpi.com/2072-6694/14/21/5411.

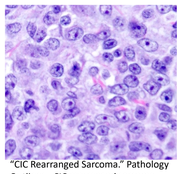
Our previous study has shown that malignancy of CDS relies on P300/CBP. It identified iP300w as a potent P300/CBP inhibitor that efficiently suppresses CIC-DUX4 oncogenic activity. Besides, from other studies, Mesenchymal stem cells (MSC), multipotent stem cells that can differentiate into various cell types, are believed to be the origin of sarcoma.

Herein, we studied the CDS surfaceome by analyzing RNASeq data from CDS cell lines treated with iP300w. We created a suite of informative visual aids to identify a distinct subset of 19 surface proteins that are dependent to CIC-DUX4/P300/CBP activity. These proteins represent an intersection of critical biological attributes, serving as potential targets of CIC-DUX4 sarcoma.

Purpose

Identify the surfaceome associated with CIC-DUX4 using RNA-Seq data.





"CIC Rearranged Sarcoma." Pathology Outlines - CIC rearranged sarcoma. Accessed August 22, 2023. https://www.pathologyoutlines.com/topic/boneCICDUX4.html.

Methodology

Datasets used in this study are listed below:

- FASTQ files of NCC-CDS1-X1 (X1) cell line and CDS tumor cells from the Gene Expression Omnibus (GEO) database with accession no. GSE165729 and GSE165032, respectively.
- FASTQ file of MSC cell line is from GEO GSE73610
- FASTQ file of Kitra-SRS (Kitra) cell line from Bosnakovski's Lab, is unpublished yet.
- CDS direct binding genes list is from the paper (DOI: 10.1172/JCI126366)
- Surfaceome list from Cell Surface Protein Atlas

The CHURP pipeline, created by the Research Informatics Solutions (RIS) from UMN, was used to produce the essential RNA reads mapping.

To discern the effects of iP300w treatment on generated cell lines, several rigorous Differential Expression analysis was executed using the `DESeq2` and `tximeta` tools. This analytical approach enabled the identification of genes exhibiting altered expression levels. The multiple test correction was undertaken using the Benjamini and Hochberg method, with a stringent criterion of adjusted p-value < 0.05.

Results

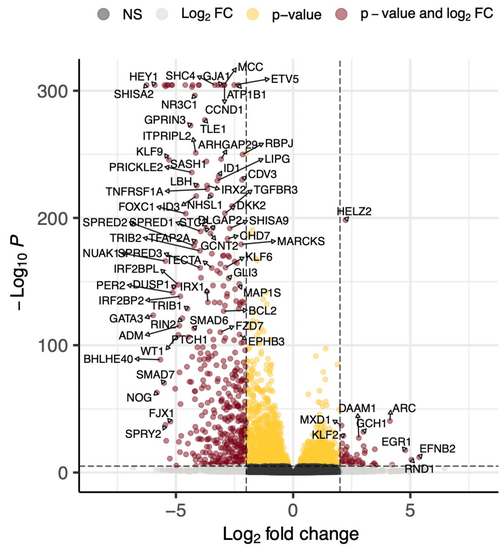


Fig. 1 *Volcano plot of CDS Cancer Kitra Cell Lines between Control and iP300w 4-hour treated, using log2FC>2, padj<0.05 as significant threshold. Genes encoding surfaceome are labeled with names.*

- In total, 132 up-regulated genes and 520 down-regulated genes, compared to the control

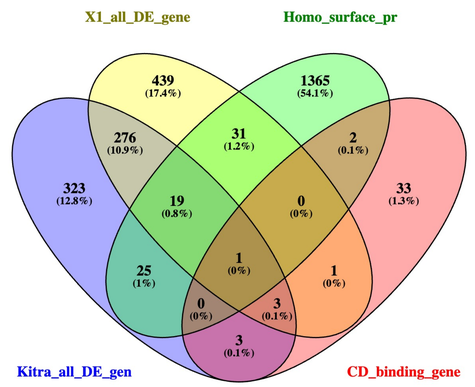


Fig. 2 *Venn diagram of Differential Expression (DE) Analysis Results for Kitra and X1 cell lines subjected to iP300w treatment, Homo sapiens surface protein genes list, and CIC-DUX4 direct binding gene list.*

- Identified 19 surface proteins dependent on CIC-DUX4/P300/CBP activity

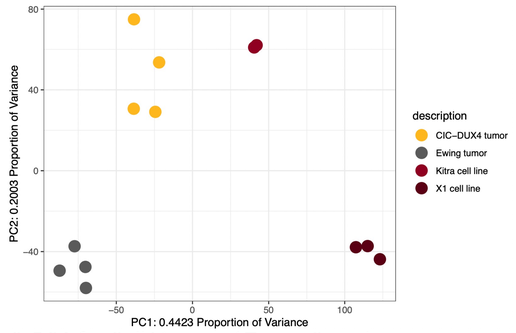


Fig. 3 *PCA for Cell Lines and Tumor Samples*

- Showed the difference in CDS and Ewing sarcoma

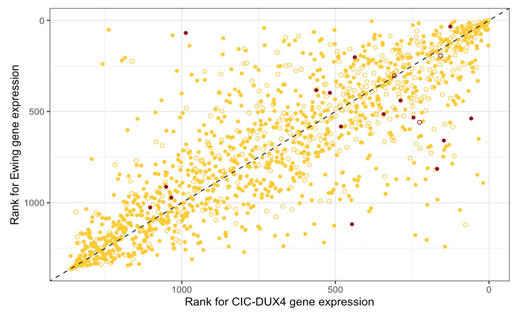


Fig. 4 *Ranking Difference of Genes Encoding Surfaceome in Tumors. Red dots: 19 identified genes; filled circle: high confidence of surfaceome; unfilled circle: putative/less-specific identifications. X and Y-axis display the average gene expression rank for tumors from the lowest (5000) to highest (0).*

- A total of 1382 genes of surfaceome were examined
- Identified 4/19 genes more likely to be surfaceome of Ewing sarcoma compared to CDS
- Revealed a potential cut-off threshold at rank 750

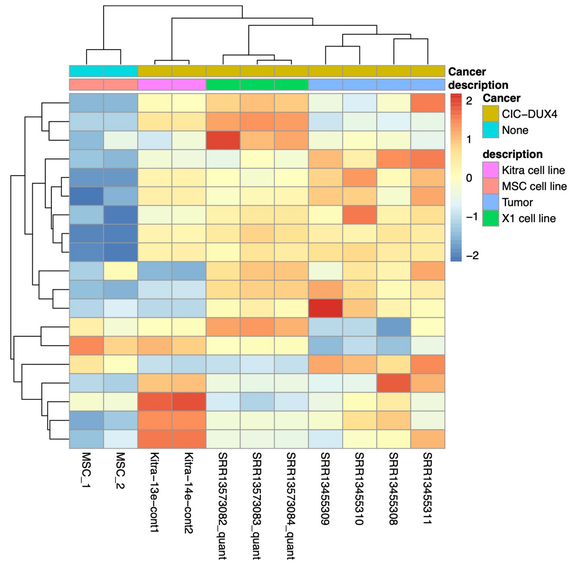


Fig. 5 *Heatmap of Gene Expression Profiling of 19 Identified CDS Genes Encoding the Cell Surfaceome Across Various Tumor Cells and Healthy MSC using Z-Scores. Red indicates high gene expression; blue represents low gene expression.*

- 16 among 19 identified genes are expressed low in MSC cell lines
- 6 identified genes have outstanding high expression in CDS cancer cell lines and tumors.

Conclusion

The discovery of 19 genes encoding cell surface proteins, a pivot encapsulating the combined effects of iP300w treatment and transcriptional regulation, is noteworthy in Fig. 2 because it holds the promise of a novel therapeutic route. Fig. 5 provides insights in verification for the expression level for the identified genes.

A thorough approach emerges as Fig. 4 explain the various gene expression profiles in the context of CDS and Ewing tumor samples and highlight the 19 identified genes from iP300w treatment, directing the medical community toward the creation of contextually relevant interventions, and providing some perspective toward potential antibody development for CIC-DUX4 sarcoma in the future.

Acknowledgements

Research reported in this publication was supported by the Children's Cancer Research Fund (CCRF).The content is solely the responsibility of the authors and does not necessarily represent the official views of Macalester College or the University of Minnesota.