Epigenetic Regulation of Head and Neck Squamous Cell Carcinoma: Insights from Transcriptome Analysis of *Nsd1* Knockout 3D Organoid Model

Kristine Lu[1], Chao Lu[2]

***Abstract* - Head and Neck Squamous Cell Carcinoma (HNSCC) are a prevalent form of cancer, affecting the oral cavity, larynx, and pharynx. HNSCC is commonly associated with tobacco use, alcohol consumption, and Human Papilloma Virus (HPV) infection. HPV Positive HNSCC generally exhibits better prognosis than HPV Negative HNSCC, which is frequently diagnosed at advanced stages with poor outcomes. Epigenetics, the study of heritable changes in gene expression that do not involve the sequence of base pairs in DNA, play a crucial role in HNSCC development. Specifically, the mutation of the gene *NSD1,* a gene that mediates histone modification, is often positively correlated with improved survival rates in HPV(-) HNSCC. This study aims to investigate the impact of *Nsd1* mutations on gene expression and pathways using murine oral-derived organoids, a 3D model that mimics head and neck squamous cancer cells' complexities. Through RNA sequencing and analysis, we compared wildtype organoids to organoids where we knocked out *Nsd1* at different stages of cancer progression. Our results suggest that *Nsd1* knockout leads to reduced inflammation, potentially contributing to weakened immune responses, and affects genes associated with Extracellular Matrix Organization, potentially influencing cancer cell migration and metastasis. Additionally, *Nsd1* may alter epithelial-mesenchymal transition, impacting cancer aggressiveness. This study provides valuable insights into the role of *Nsd1* in HNSCC and identifies potential targets for future research and therapeutic interventions.**

I. Introduction

HNSCCs (Head and Neck Squamous Cell Carcinomas) are the sixth most common cancer in the world. The majority of HNSCCs are found in Stage III and IV, where prognosis is poor, making it important to develop effective, less-toxic therapies.[1]

Our study concentrated on how HSNCCs are affected by epigenetics, the study of heritable changes in gene expression that do not involve the sequence of base pairs in DNA. Epigenetic changes consist of various chemical modifications of DNA and its associated proteins, altering gene expression patterns and phenotypic results. For example, the modification of histones, small proteins that maintain structure in chromatin, is widely considered to be a major factor in epigenetics. DNA is wound around histone octamers, keeping the DNA compact and allowing it to fit within the confines of the cell nucleus. When an epigenetic regulator chemically alters a histone protein, such as through the addition of a methyl group or acetyl group, it can affect how tightly wound DNA is, reducing or improving its transcribability and resulting in changes in gene expression, protein production, and cell state. Cancers are often associated with abnormality of certain epigenetic modifiers and modulators, making epigenetics a potential target for research and the development of new treatments.[2]

Nsd1 is a gene which plays a significant role in developmental epigenetics. One of the primary functions of *Nsd1* is to mediate the di-methylation at the 36th lysine residue of the Histone 3 protein, [H3K36me2], causing downregulation of transcription of nearby genes. Numerous studies[3, 4] have found that the *Nsd1* gene is frequently mutated in HNSCC, and that there is a positive correlation between mutations in *Nsd1* and the patient rate of recovery from HPV(-) HNSCC through the alteration of methylation levels of the H3K36me2 residue. In one such study[4], among 457 HPV(−) tumors, 13% contained alterations in the *Nsd1* gene. HNSCC patient samples with mutations in *Nsd1* exhibited reductions in H3K36me2. Recent studies[3,4] implicate that *Nsd1* mutations are associated with significantly improved patient survival rates. Mutations in *Nsd1* that lead to loss of function of the gene have also been found to correlate with the expression of several miRNAs and mRNAs important in cancer regulation and proliferation pathways. However, the mechanisms behind how these mutations in *Nsd1* ultimately lead to increased survival in HNSCC patients is still unclear, and further research is needed to clarify the pathways involved.

Our study aimed to identify genes and pathways which exhibited consistent and significant changes in expression when *Nsd1* was heavily mutated in carcinogenic head and neck squamous cells. While previous studies have been successful in analyzing the role of *Nsd1* through 2D cell cultures, we utilized oral-derived organoids (self-organized 3D tissue cultures that are generally derived from stem cells) to recapitulate features of primary mouse tongue epithelium in a 3D model encapsulating the key complexities of head and neck squamous cells. We procured samples of organoid cells at different stages of cancer growth in a mouse model of HNSCC and performed RNA sequencing. By analyzing the RNA-seq dataset, we compared the level of expression of various genes between cells with wildtype Nsd1 and cells where *Nsd1* was knocked out.

II. Materials and Methods

Deleting *Nsd1* gene in mouse tongue epithelium:

The Lu lab utilized Cre Recombinase, a specialized enzyme facilitating targeted genetic modifications by excising segments of DNA between selected loxP sites (which we introduced through genetic engineering), to delete exon 3 of the *Nsd1* gene in a conditional knockout mouse strain. Cre Recombinase was induced in the presence of a KRT5 promoter (a sequence which controls the expression of genes) by Tamoxifen (a commonly used inducer of Cre

[1]K.L. is with University High School, 4771 Campus Drive, Irvine, CA 92618, and Lu Lab @ Columbia University (corresponding author to e-mail: kaynelu921@gmail.com).
[2]C.L. is with Lu Lab @ Columbia University

Recombinase), which was administered via oral gavage 3 times a week. To ensure that the deletion and recombination of DNA had successfully occurred, the mice also carry an mTmG allele. mTmG mice allow us to determine the effectiveness of Cre-recombinase using color fluorescence.

Generation of *Nsd1* knockout organoids, RNA extraction and sequencing:

We treated *Nsd1* wildtype and *Nsd1* knockout mice with 4NQO (r-nitroquinoline 1-oxide) (a potent carcinogen) in their drinking water to induce cancer growth. Mouse tongue tissue from control and *Nsd1* KO mice following tamoxifen administration were then dissociated into single cells and sorted to allow for isolation of epithelial cell populations. Cells were seeded and added to the media, which was changed every two days, and cells were passaged to form organoids, which provide a more physiologically relevant model than cell populations for study. Passaging involves dissociating and moving cells between culture vessels in order to promote and maintain healthy organoid growth and prevent overcrowding. We established 3D organoid cultures from mice at 0, 16, or 32 weeks of 4NQO treatment, which are considered as normal, precancerous, and cancerous, respectively.

RNA extraction and sequencing analysis:

Once our organoids had developed, we disrupted the organoid cells to release the RNA using Trizol, taking appropriate measures to protect the RNA from degradation. This step aimed to capture the genetic information present in cells, particularly RNA involved in gene expression. We then extracted it using a standard RNA extraction kit and purified it to remove impurities. Extracted RNA samples were used to prepare for sequencing libraries using Illumina kits. The libraries were sequenced using NextSeq 400.

The raw sequencing data obtained from the RNA-sequencing was processed using FastQC, a widely used tool for assessing the quality of sequencing data. FastQC measured the per-base sequence quality, identifying regions of inaccurate data, as well as sequence length distribution, which ensures uniform coverage and detects anomalies in our RNA sequences. The GC content refers to the proportion of guanine and cytosine nucleotides in the RNA sequence, and discrepancies were identified and data was adjusted accordingly. FastQC also helped identify redundant sequence stretches which were removed. We performed necessary trimming and filtering to pre-process the data. The reads were then aligned to a mm10 reference genome using HISAT2 alignment algorithms. This process involved mapping the reads to their corresponding genomic locations in the mouse genome. Subsequently, the characteristic sequences in the aligned reads were quantified using the tool, featureCounts, to assess gene expression patterns.

We performed differential gene expression analysis of our processed data to determine using DESeq2, a widely used tool for RNA sequencing data analysis. DESeq2 first performs data normalization, adjusting for differences in sequencing depth between samples to ensure accurate comparisons. It then estimates dispersion to account for variability in gene expression across samples, before matching a generalized linear model framework to the dispersion-adjusted, normalized data to identify genes that show significant changes in expression levels between different batches of cells and cells of different time frames. This analysis allowed for the identification of genes that were significantly upregulated or downregulated in *Nsd1* KO cells compared to wild-type cells. We utilized DESeq2 to generate tables of differentially expressed genes, log-fold changes between data points, and p-values. Statistical tests, such as negative binomial distribution-based analysis, were applied to determine the significance of gene expression changes. These tests, particularly the negative binomial distribution-based analysis implemented in DESeq2, were selected for their suitability in handling the unique characteristics of RNA-seq data. The negative binomial distribution is well-suited for modeling RNA Sequencing data by accommodating overdispersion and accurately reflecting the variability observed in gene expression levels. A link to the R script used can be found here.

We then further analyzed the differentially expressed genes using Enrichr, a gene set analysis server widely used in gene analysis. This analysis helped identify key pathways and biological processes associated with *Nsd1* mutation and their potential implications in cancer progression.

Data Visualization:

PCA Plots, Heatmaps, and Lists of significantly upregulated and downregulated genes were generated based on statistical analysis results using data visualization libraries in R including ggplot2, DESeq2, gplots, and more.

III. Results

Our data consisted of 16 sets of RNA-seq samples from different organoid cultures at different stages of cancer development. We had 4 data points from 0 weeks into cancer development, 5 points from 16 weeks into development, and 7 points from 32 weeks into cancer development. Each data point consisted of a list of affected genes and their levels of expression, varying based on amount of time into cancer development and *Nsd1* KO vs. Wild-type. PCA plots/clustering showed that 0-week and 16-weeks organoids had similar levels of overall gene expression to each other, while 32-week organoids showed large differences from the 0/16 week data points. (Fig. 1). This is likely due to the progression of cancerous mutations over time.

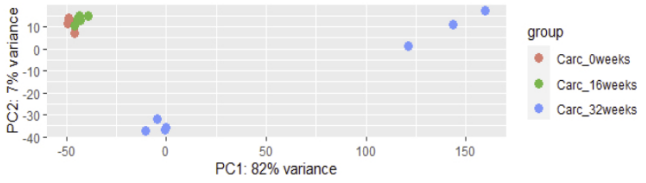

Figure 1: PCA Plot comparing organoids at 0 weeks, 16 weeks, 32 weeks of 4NQO carcinogen treatment. Generated from R

To confirm, we analyzed and compared the 0/16 week wild-type data and 32 week wild-type data as a control

group. Inputting our data into Enrichr, a gene analysis tool, we showed that at 32 weeks, there was an upregulation of genes involved in the RAC1 GTPase Cycle (associated with cytoskeletal structure and cell proliferation), inflammatory response pathways, and Allograft Rejection (rejection of foreign cells), which suggests increased activation of signaling pathways associated with cellular movement, immune response, and potential rejection-like processes (Figure 2).

On the other hand, our data also showed significant downregulation of genes related to oxidative phosphorylation, cyclin and cell-cycle regulation, and mitochondrial translation elongation pathways, indicating a shift away from normal cellular energy production and controlled cell division (Figure 3). This deviation from typical cellular processes aligns with the "hallmarks of cancer," reflecting dysregulation in energy metabolism and cell cycle control, which are common features in cancerous cells.These changes align with characteristics observed in cancer, particularly in terms of enhanced cell mobility and immune-related responses. Overall, these findings suggest a molecular signature consistent with cancer progression in the 32-week organoids compared to the 0/16-week ones. The observed alterations in gene expression patterns validate the effectiveness of the analysis methods employed, supporting the conclusion that the organoid cultures were evolving in line with the intended experimental design.

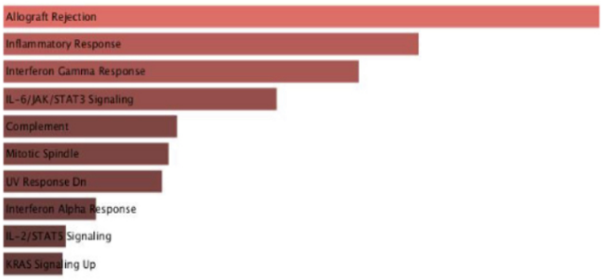

Figure 2: Significantly Upregulated Genes in MSigDB Hallmark 2020 Pathways in 32-weeks organoids, sorted by P-Value.

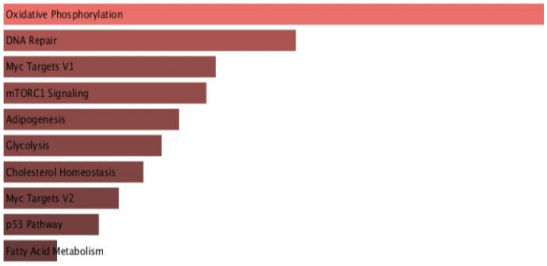

Figure 3: Significantly Downregulated Genes in MSigDB Hallmark 2020 Pathways in 32-weeks organoids, sorted by P-Value.

We also noticed in a PCA plot where the *Nsd1* KO genotype was annotated (Figure 4) that, at 32 weeks, but not at 0/16 weeks, *Nsd1* KO cell organoids varied greatly from WT organoids in their gene expression patterns.

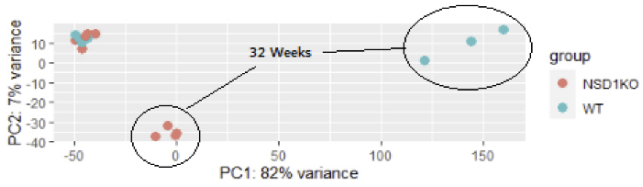

Figure 4: PCA Plot comparing Nsd1KO and WT organoids. Generated by R

A heatmap generated through unsupervised clustering, which identifies patterns in gene expression even without labeled categories, showed similar results, where wild-type 32-Weeks organoids showed high variance of gene expression compared to *Nsd1* KO organoids at 32 weeks, which clustered much closer to 0/16-week organoids (Figure 5).

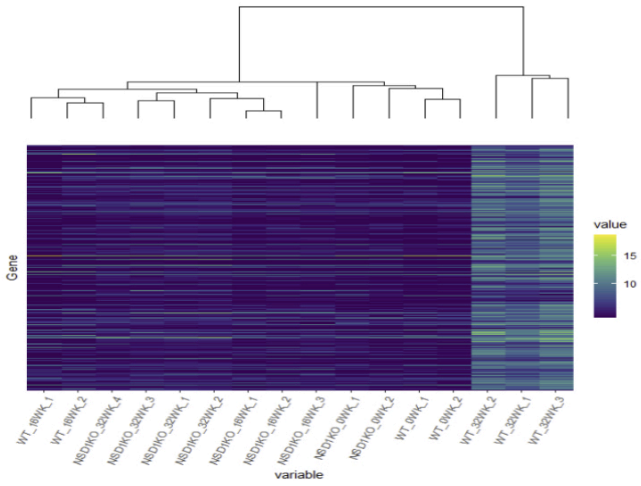

Figure 5: Gene expression heatmap of WT/Nsd1KO organoids over 0/16/32 weeks, clustered by similarity, generated by R.

We proceeded to compare differentially expressed genes of 32-weeks wild-type organoids and 32-weeks *Nsd1* KO organoids. We found significant downregulation of genes in the Allograft Rejection, inflammatory response pathways, and inflammatory response in cells when *Nsd1* was knocked out (Figure 6). This implies that inflammation, a hallmark of cancer, is lowered and tumor tissue doesn't trigger as great of an immune response in cells where *Nsd1* is knocked out. This is consistent with previous studies that report DNA methylation resulting in immune-cold phenotypes[5].

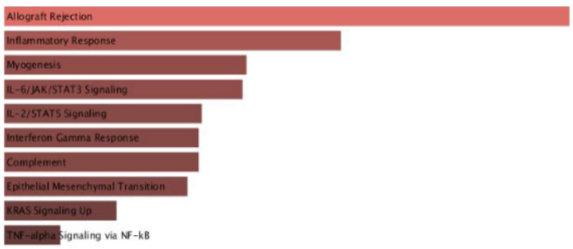

Figure 6: Significantly Downregulated Genes in MSigDB Hallmark 2020 Pathways in *Nsd1* knockout organoids, sorted by P-Value.

In addition, 32-weeks *Nsd1* KO organoids displayed an upregulation of Extracellular Matrix (ECM) Organization genes (eg. Apical Junction) which are responsible for monitoring cell movement (Figure 7). This upregulation is influential in controlling the movement and spread of cancer cells from the primary tumor in the beginnings of metastasis[15] . Our results align with previous studies investigating the relationship between *Nsd1* and cancer, which imply that the *Nsd1* gene may contribute to ECM modeling associated with tumor metastasis.

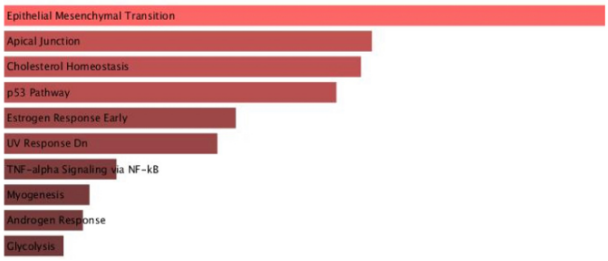

Figure 7: Significantly Upregulated Genes in MSigDB Hallmark 2020 Pathways in Nsd1 knockout organoids, sorted by P-Value.

We also observed a significant upregulation in the epithelial mesenchymal transition (EMT) pathway genes (Figures 7 and 8), which allows epithelial cells to undergo changes and transform into a mesenchymal cell phenotype. Mesenchymal cells are multipotent stem cells with a dynamic cytoskeleton and which exhibit cellular plasticity, and are characterized by extensive migratory capabilities. Studies[8, 9, 10] have shown that EMT is involved in tumor progression and metastasis, as well as initial angiogenesis (the formation of new blood vessels around a tumor, enhancing tumor growth) sprouting. In our *Nsd1* KO organoids, genes which generally promote EMT such as Snai1, Cdh2, Col11a1, Mmp16, Zeb1, Zeb2 and others were significantly downregulated, while genes which inhibit EMT including Cldn4, Cldn3, Epcam, Patj, Ocln, and more were significantly upregulated (Figure 10). Thus, it can be speculated that *Nsd1* may be regulating genes that promote EMT, so that when it is heavily mutated or knocked out, tumor cells have less EMT and metastasize less rapidly.

A graph of select significant genes and their Log2 fold changes is shown below (Figure 8). Negative Log2 Fold Changes signify significant downregulation, while positive changes signify significant upregulation.

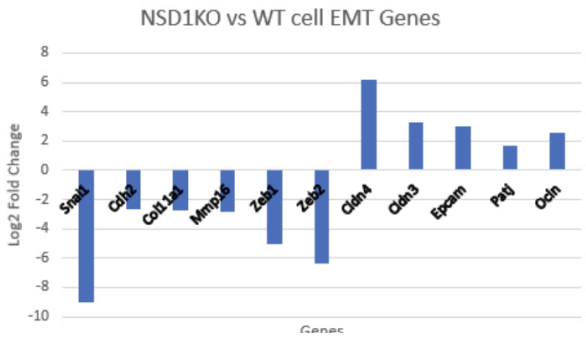

Figure 8: Log2 Fold Changes of EMT genes comparing Nsd1 KO vs. WT organoids.

IV. Discussion

Based on the results of our analysis, we can speculate that there is a correlation between loss of *Nsd1* and changes in EMT pathway as well as a downregulation of inflammation response. Our results align with previous studies[11, 12, 13] that suggested a relationship between the EMT pathway and inflammation in carcinogenesis. Cancer cells which have undergone EMT have altered cell-cell junction complexes (complexes allowing for cell communication and associated with preventing overgrowth) and become more aggressive, stimulating the production of proinflammatory factors such as IL-1 and TNF, by cancer cells. Consistent with our findings, studies show that a decrease in EMT is associated with a decrease in inflammation[16].

Similarly, we noticed a correlation between EMT pathway expression and ECM Organization pathways. Shared genes in the ITGB family, Laminin Subunit Alpha genes, and more were present in both EMT pathways and Extracellular Matrix Organization pathways.

ITGB1, Integrin Subunit Beta 1, and ITGAM, Integrin Subunit Alpha M, are protein-coding genes which interact with the ECM. In particular, ITGB forms a heterodimeric complex (a protein composed of two different polypeptide chains) with an alpha subunit of integrins, allowing integrins to interact with ligands in the ECM. This allows it to regulate various cellular processes including migration, proliferation, and survival. Mutations and dysregulations of ITGB have been implicated in various aspects of cancer development and metastasis. In *Nsd1* KO organoids, ITGB1 is significantly upregulated compared to in wild-type cells.

Another gene family including LAMB3, LAMC2, LAMB1, and LAMA5 encodes for different subunits of the laminin protein family. Laminins are proteins with sugar molecules attached (glycoproteins) that play a crucial role in the formation of the basement membranes in the extracellular matrix. Laminin-encoding genes are involved in both the EMT pathways and the ECM Organization pathways, and are a component in determining cell growth, communication, and movement. Reduced expression of LAMA5 has been found in some types of cancer, and restoring its expression has been shown to inhibit tumor growth and metastasis. LAMA5, LAMB1, LAMB3, and LAMC2 were among significantly upregulated genes in organoids where *Nsd1* was knocked out.

Clinically, several studies have linked both lower metastatic potential and increased chemotherapy responsiveness to genes involved in the EMT pathway. However, the exact extent of EMT's role in these observations requires further investigation.

Our data also revealed the upregulation and downregulation of many genes that are not part of the above pathways, including upregulations in pathways associated with T-cells, NK cells and others in *Nsd1* KO cells, and downregulations in other pathways. However, The significance of these pathways in cancer pathophysiology is still not clear and warrants further studies.

In summary, it was observed that *Nsd1* KO organoids at 32 weeks had a lower level of cancer aggressiveness compared to wild-type cells at the same time point. The downregulation of genes associated with inflammation pathways in *Nsd1* KO organoids suggests that they trigger a weaker immune response than wild-type cells, which the body reacts strongly to. In contrast, upregulation of genes associated with ECM organization in *Nsd1* KO organoids indicates an effect on the mobility of cancer cells during the early stages of metastasis. Furthermore, *Nsd1* may promote epithelial-mesenchymal transition (EMT) (the process where epithelial cells gain migratory properties and become mesenchymal stem cells), allowing cells higher migratory capability. In *Nsd1KO* organoids, we found that genes that promote EMT were significantly downregulated, while those that inhibit EMT were upregulated.

The decrease in inflammation and inhibition of EMT and ECM Organization Pathways in *Nsd1* knockout mouse HNSCC organoids are consistent with clinical reports of a better prognosis in human HNSCC when *Nsd1* is heavily mutated. These findings have significant implications for cancer management and treatment, as *Nsd1* presents a promising target for biomarker development and therapeutic intervention, to recognize genetic signs and intervene before cancer development[12]. However, further research is needed to fully understand the mechanisms by which *Nsd1* functions in cancer progression, and to determine the optimal strategies for targeting this gene in cancer therapy.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011;61:69–90.
2. Weinhold B. Epigenetics: the science of change. Environ Health Perspect. 2006 Mar;114(3):A160-7. doi: 10.1289/ehp.114-a160. PMID: 16507447; PMCID: PMC1392256.
3. Nargess Farhangdoost, Cynthia Horth, Bo Hu, Eric Bareke, Xiao Chen, Yinglu Li, Mariel Coradin, Benjamin A. Garcia, Chao Lu, Jacek Majewski, Chromatin dysregulation associated with Nsd1 mutation in head and neck squamous cell carcinoma, Cell Reports, Volume 34, Issue 8, 2021, 108769, ISSN 2211-1247,
4. Pan, C., Izreig, S., Yarbrough, W. G., & Issaeva, N. (2019). Nsd1 mutations by HPV status in head and neck cancer: differences in survival and response to DNA-damaging agents. Cancers Head Neck, 4, 3. PMID: 31321084; PMCID: PMC6613249.
5. Brennan K, Shin JH, Tay JK, Prunello M, Gentles AJ, Sunwoo JB, Gevaert O. NSD1 inactivation defines an immune cold, DNA hypomethylated subtype in squamous cell carcinoma. Sci Rep. 2017 Dec 6;7(1):17064. doi: 10.1038/s41598-017-17298-x. PMID: 29213088; PMCID: PMC5719078.
6. Oshi, M., Le, L., Angarita, F.A., Tokumaru, Y., Yan, L., Matsuyama, R., Endo, I., & Takabe, K. (2021). Association of allograft rejection response score with biological cancer aggressiveness and with better survival in triple-negative breast cancer (TNBC). Journal of Clinical Oncology, 39(15_suppl), 561. https://doi.org/10.1200/JCO.2021.39.15_suppl.561
7. Krossa I, Strub T, Aplin AE, Ballotti R, Bertolotto C. Lysine Methyltransferase Nsd1 and Cancers: Any Role in Melanoma? Cancers (Basel). 2022 Oct;14(19):4865. doi: 10.3390/cancers14194865. PMID: 36230787; PMCID: PMC9563040.
8. Huang, Y., Hong, W. & Wei, X. The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. *J Hematol Oncol* 15, 129 (2022). https://doi.org/10.1186/s13045-022-01347-8
9. Ribatti D, Tamma R, Annese T. Epithelial-Mesenchymal Transition in Cancer: A Historical Overview. Transl Oncol. 2020 Jun;13(6):100773. doi: 10.1016/j.tranon.2020.100773. Epub 2020 Apr 22. PMID: 32334405; PMCID: PMC7182759.
10. Masao Saitoh, Involvement of partial EMT in cancer progression, *The Journal of Biochemistry*, Volume 164, Issue 4, October 2018, Pages 257–264, https://doi.org/10.1093/jb/mvy047
11. Suarez-Carmona M, Lesage J, Cataldo D, Gilles C. EMT and inflammation: inseparable actors of cancer progression. Mol Oncol. 2017 Jul;11(7):805-823. doi: 10.1002/1878-0261.12095. Epub 2017 Jun 26. PMID: 28599100; PMCID: PMC5496491.
12. Ricciardi, M., Zanotto, M., Malpeli, G. *et al.* Epithelial-to-mesenchymal transition (EMT) induced by inflammatory priming elicits mesenchymal stromal cell-like immune-modulatory properties in cancer cells. *Br J Cancer* 112, 1067–1075 (2015). https://doi.org/10.1038/bjc.2015.29
13. Md Shariful Islam, Md Reaz Morshed, Golap Babu, Md Asaduzzaman Khan. The role of inflammations and EMT in carcinogenesis, Advances in Cancer Biology - Metastasis Volume 5, 2022,100055,ISSN 2667-3940, https://doi.org/10.1016/j.adcanc.2022.100055.
14. Yang C, Wang K, Liang Q, Tian TT, Zhong Z. Role of Nsd1 as potential therapeutic target in tumor. Pharmacol Res. 2021 Nov;173:105888. doi: 10.1016/j.phrs.2021.105888. Epub 2021 Sep 16. PMID: 34536546.
15. Popova NV, Jücker M. The Functional Role of Extracellular Matrix Proteins in Cancer. Cancers (Basel). 2022 Jan 4;14(1):238. doi: 10.3390/cancers14010238. PMID: 35008401; PMCID: PMC8750014.
16. Suarez-Carmona M, Lesage J, Cataldo D, Gilles C. EMT and inflammation: inseparable actors of cancer progression. Mol Oncol. 2017 Jul;11(7):805-823. doi: 10.1002/1878-0261.12095. Epub 2017 Jun 26. PMID: 28599100; PMCID: PMC5496491.