Title: Multi-omic Dissection of Lymph Node Metastasis in Oral Squamous Cell Carcinoma

Ayan Chanda[1], Daniel Young[1,6], Steven C. Nakoneshny[2], T. Wayne Matthews[2,4], Shamir Chandarana[2,4], Robert D. Hart[2,4], Joseph C. Dort[2,4,5], Antoine Dufour[1,6], Martin D. Hyrcza[2,3], Pinaki Bose[1,2,3,5].

[1]Department of Biochemistry and Molecular Biology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.

[2]Ohlson Research Initiative, Arnie Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.

[3]Department of Pathology and Laboratory Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.

[4]Department of Surgery, Section of Otolaryngology – Head & Neck Surgery, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.

[5]Department of Oncology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.

[6]Department of Physiology and Pharmacology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.

Keywords

Oral squamous cell carcinoma (OSCC), Lymph node metastasis (LNM), Multi-omics, Metastatic drivers, Biomarker discovery, Non-coding alterations.

Abstract

Background: Lymph node metastasis (LNM) is a major predictor of poor outcomes in oral squamous cell carcinoma (OSCC). The molecular events enabling this spread to remain poorly understood, and single-omics approaches have not fully explained why some tumors metastasize.

Methods: We analyzed 60 primary OSCC tumors with matched adjacent normal tissues from the Marathon of Hope Cancer Centres Network (MOHCCN) Gold Cohort. Each sample underwent whole genome sequencing (WGS), whole transcriptome sequencing (WTSeq), ATAC-seq for chromatin accessibility, proteomics, and metabolomics. We integrated these data using multi-omic factor analysis (MOFA2), an unsupervised framework that identifies latent factors capturing shared biological variation across molecular layers. Non-coding somatic mutations were also profiled and intersected with chromatin accessibility data.

Results: MOFA2 revealed latent factors that differentiate LNM-positive from LNM-negative tumors, uncovering complex regulatory programs not visible within individual data types. Recurrent non-coding mutations were enriched in LNM-positive tumors, suggesting potential roles in transcriptional regulation. Ongoing work is examining how these mutations may influence chromatin accessibility and gene expression.

Conclusion: This integrative multi-omic framework exposes hidden biological drivers of metastasis in OSCC, identifies potential biomarkers for risk stratification, and supports new hypotheses for

therapeutic targeting. Future work will extend these findings with spatial transcriptomics and validation in external cohorts, underscoring the value of systems-level approaches in cancer research.

Plain Language Summary

When cancer starts in the oral cavity (oral squamous cell carcinoma, or OSCC), one of the most important signs that it might be aggressive is whether it spreads to nearby lymph nodes. Tumors that spread usually lead to significantly worse outcomes, but we still do not fully understand the molecular underpinnings of this phenomenon.

To learn more, we studied tumor samples donated by 60 patients with OSCC and compared them to nearby normal tissue. For each tumor, we used several powerful tests:

- DNA sequencing to look for changes in the genetic code,
- RNA sequencing to see which genes are turned on or off,
- Chromatin accessibility testing (ATAC-seq) to learn how the DNA is "packaged" and whether important regions are open or closed,
- Detailed protein and metabolism analyses to understand the tumor's activity at other levels.

We combined all this information using advanced computer methods that can reveal hidden patterns across these different types of data. These patterns clearly separated tumors that had spread to lymph nodes from those that had not. We also found changes in parts of the DNA that don't directly code for genes but may still control how genes behave. Our next step is to see whether these DNA changes affect how the tumor's DNA is organized and how it turns genes on or off.

By bringing together all these layers of information, our goal is to find reliable molecules or pathways (biomarkers) that can show which oral cancers are more likely to spread. This could help doctors better predict risk and, in the future, guide treatment decisions. We will also test these findings in larger groups of patients and use new techniques, like spatial transcriptomics, which show where different cells and molecules sit inside a tumor.