ORIGINAL RESEARCH PAPER
Volume - 14 | Issue - 09 | September - 2025 | PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

ROLE OF COMPUTATIONAL GENOMICS IN MICROBIOME-BASED DIAGNOSIS OF ORAL SQUAMOUS CELL CARCINOMA (OSCC)

Oral Pathology

Dr. Parul Dixit* Junior Resident, Government Dental College, Thiruvananthapuram *Corresponding Author

Dr. R Rajeev Associate Professor, Government Dental College, Thiruvananthapuram

Dr. S.K Padmakumar HOD & Professor, Government Dental College, Thiruvananthapuram

ABSTRACT

Oral squamous cell carcinoma (OSCC) remains a global public health concern, with rising incidence and limited early detection strategies. Recent advances in metagenomics and computational genomics have revealed that the oral microbiome plays a pivotal role in OSCC initiation and progression. This review discusses the integration of high-throughput sequencing and bioinformatics tools to decode microbiome alterations associated with OSCC. Emphasis is laid on the identification of pathogenic microbial taxa, functional genomic shifts, host-microbe interactions, and predictive biomarker discovery. The discussion also explores machine learning approaches and multi-omics integration strategies aimed at improving early diagnosis, personalized risk assessment, and therapeutic targeting in OSCC.

KEYWORDS

Oral squamous cell carcinoma; computational genomics; oral microbiome; metagenomics; biomarker discovery; microbial dysbiosis; bioinformatics; machine learning.

INTRODUCTION

Oral squamous cell carcinoma (OSCC), the most common malignancy of the oral cavity, accounts for over 90% of oral cancers globally. Despite advances in surgery and radiotherapy, the 5-year survival rate remains at 50–60%, largely due to delayed diagnosis and frequent recurrence[1]. Conventional diagnostic methods, including histopathological biopsy, offer limited insight into the biological changes preceding malignant transformation. Recently, the role of the oral microbiome in carcinogenesis has emerged as a key area of interest. Composed of bacteria, viruses, and fungi, the oral microbiota interacts with host tissues and immune cells, contributing to both oral health and disease[2]. Dysbiosis, or the imbalance of microbial communities, has been implicated in inflammatory responses, epithelial barrier disruption, and oncogenic signalling in the oral mucosa[3].

The application of next-generation sequencing (NGS), including 16S rRNA gene sequencing and whole-metagenome shotgun sequencing (WMGS), has revolutionized our ability to study the oral microbiome in unprecedented detail. Computational genomics—encompassing the use of bioinformatics algorithms and pipelines—plays a central role in analyzing these massive datasets, identifying key microbial signatures, and linking them to host gene expression and clinical outcomes[4].

DISCUSSION

1. Microbiome Dysbiosis and Molecular Carcinogenesis in OSCC

Microbial dysbiosis plays a critical role in OSCC by modulating host immune responses, promoting chronic inflammation, and reprogramming host cellular pathways toward oncogenesis. Periodontal pathogens such as Porphyromonas gingivalis and Fusobacterium nucleatum contribute to the disruption of oral epithelial homeostasis. P. gingivalis activates NF-κB and STAT3 signaling, both of which promote epithelial–mesenchymal transition (EMT), a key process in tumor invasion and metastasis[7]. Additionally, F. nucleatum increases pro-inflammatory cytokines like IL-6 and TNF-α, disrupts epithelial integrity by cleaving E-cadherin, and enhances β-catenin signaling, thereby supporting cellular proliferation and migration[8–9].

2. Metagenomic Profiling Through Computational Genomics

In OSCC patients, metagenomic studies have consistently demonstrated a shift in microbial composition marked by an increased abundance of pathobionts like P. gingivalis, T. denticola, and F. nucleatum, and a concurrent depletion of beneficial commensals like Streptococcus sanguinis[12]. Functionally, microbial pathways associated with polyamine biosynthesis, lipopolysaccharide (LPS) production, and butyrate metabolism are enriched in OSCC samples[13]. These molecules influence the tumor microenvironment by enhancing oxidative stress, altering immune signaling, and modulating host epigenetic landscapes. For example, elevated levels of cadaverine and putrescine—polyamines produced by dysbiotic bacteria—can promote DNA strand breaks and oncogene activation[14]. Importantly, strain-level profiling has enabled researchers to link specific virulence factors (e.g., fimbriae, gingipains, hemolysins) to disease severity. This is critical in the context of OSCC, where bacterial virulence, not just presence, may determine disease progression. Emerging pipelines also integrate pangenomic analysis, allowing for identification of mobile genetic elements, antibiotic resistance genes, and prophages in OSCC-associated microbiota[15].

3. Integrative Host–Microbe Transcriptomics

To understand how microbial communities interact with host tissues in OSCC, integrative transcriptomics approaches are gaining prominence. Bulk and single-cell RNA sequencing (scRNA-Seq) can capture host gene expression changes in response to microbial colonization. When integrated with microbial abundance data, these tools reveal a dynamic molecular crosstalk at the tumor–microbe interface[16].

Studies using DESeq2 and edgeR have demonstrated significant changes in host gene expression upon microbial invasion. P. gingivalis, for example, downregulates tumor suppressor genes such as TP53, while upregulating genes involved in angiogenesis (VEGFA), matrix remodeling (MMP9), and EMT (TWIST1, ZEB1)[17]. These changes are facilitated by microbial metabolites like short-chain fatty acids (SCFAs) and polyamines that act as epigenetic modulators, altering histone acetylation and DNA methylation[18]. Moreover, MixOmics and iHMP (Integrative Human Microbiome Project) platforms enable the correlation of microbial signatures with host transcriptomic profiles. Through canonical correlation analysis and partial least squares regression, these tools identify significant host–microbiota interactions that may drive carcinogenesis[19].

Single-cell transcriptomics has further unraveled the spatial and cellular complexity of immune responses in OSCC. Tools like CIBERSORTx and CellPhoneDB deconvolve immune cell populations and infer ligand–receptor interactions influenced by microbial stimuli[20]. This has revealed increased infiltration of immunosuppressive regulatory T cells (Tregs), M2-polarized tumor-associated macrophages (TAMs), and exhausted CD8+ T cells in OSCC tissues harboring dysbiotic microbial communities[21]. These immune changes facilitate immune escape, tumor progression, and resistance to therapy.

Furthermore, novel multi-modal pipelines now integrate single-cell ATAC-Seq, chromatin accessibility, and microbiome data, enabling epigenetic mapping of host–microbe interactions. These analyses suggest that microbial influence extends beyond transcriptional modulation to chromatin remodeling, especially at promoter regions of

oncogenes and immune checkpoints[22].

4. Machine Learning For Microbial Biomarker Discovery

Machine learning (ML) algorithms have been deployed to analyze high-dimensional microbiome datasets, aiding in the identification of diagnostic biomarkers for OSCC. Techniques such as Random Forests, Support Vector Machines (SVMs), and LASSO regression have demonstrated high classification accuracy when applied to oral and salivary microbiota profiles[23].

For example, ML-based models trained on salivary microbiome data have achieved >90% sensitivity and specificity in distinguishing OSCC patients from healthy individuals[24]. Feature selection algorithms consistently identify F. nucleatum, P. gingivalis, and altered butyrate-producing taxa as key discriminatory features. Integration of these models with host transcriptomic or metabolomic data enhances predictive performance, facilitating precision diagnostics.

5. Metabolomics And Microbial Interactomes

Metabolomic profiling of microbial communities reveals key metabolites—such as SCFAs, hydrogen sulfide (H_2S), and polyamines—that influence host cell proliferation, DNA integrity, and immune regulation[25]. Computational models like MICOM and AGORA simulate microbial metabolic networks under different environmental contexts, predicting interspecies interactions and host outcomes[26]. For instance, a shift from butyrate-producing to propionate-producing species is associated with pro-inflammatory and pro-tumorigenic signaling in the oral mucosa.

6. Spatial Genomics And Niche-specific Microbial Signatures

Recent advances in spatial transcriptomics and in situ sequencing techniques have enabled three-dimensional mapping of microbial communities within tumor tissues. Tools like Seurat, SPATA, and STUtility facilitate integration of spatially resolved transcriptomic data with microbial DNA signals[27]. This has uncovered niche-specific microbial enrichment—especially at tumor invasive fronts and necrotic cores—where bacteria co-localize with increased expression of angiogenic and immunosuppressive markers such as IL-8, VEGF, and Cd163[28].

7. Clinical Translation And Limitations

Despite the growing promise of microbiome-based diagnostics in OSCC, several translational hurdles persist. Inter-individual variability, lack of standardised reference microbiomes, differences in DNA extraction protocols, and sequencing platforms lead to poor reproducibility[29]. Moreover, the majority of studies are cross-sectional, limiting their prognostic utility. Longitudinal cohort studies are needed to establish causal relationships and dynamic microbial changes across premalignant to malignant transitions.

Computational barriers also exist. High-performance computing infrastructure is required to process large multi-omics datasets, and bioinformatics expertise is not universally available. Initiatives like the NIH iHMP, the Earth Microbiome Project, and OpenBiome are working to standardize data pipelines and increase accessibility to curated datasets[30].

Integration of real-time sequencing technologies such as Oxford Nanopore and AI-powered analytics promises a future of point-of-care microbial diagnostics. Ethical considerations—especially concerning data sharing, consent, and algorithmic bias—must be addressed to ensure equitable implementation.

CONCLUSION

Computational genomics has emerged as a cornerstone in decoding the role of the microbiome in OSCC pathogenesis and diagnosis. Through high-resolution taxonomic profiling, functional pathway analysis, and host–microbiome integration, researchers are uncovering microbial contributions to carcinogenesis with increasing clarity. Bioinformatics tools and machine learning models have enabled the identification of microbial biomarkers with potential for non-invasive diagnosis and prognosis. As multi-omics approaches mature and clinical validation expands, microbiome-informed diagnostics may soon become a standard adjunct in OSCC management.

REFERENCES

1. Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. Oral Oncol. 2009;45(4-5):309–316.
2. Deo PN, Deshmukh R. Oral microbiome: Unveiling the fundamentals. J Oral Maxillofac Pathol. 2019;23(1):122.
3. Perera M, Al-Hebshi NN, Speicher DJ, Perera I, Johnson NW. Emerging role of bacteria in oral carcinogenesis: A review with special reference to Fusobacterium nucleatum. J Oral Microbiol. 2016;8(1):32762.
4. Knight R, Vrbanac A, Taylor BC, et al. Best practices for analysing microbiomes. Nat Rev Microbiol. 2018;16(7):410–422.
5. Atanasova KR, Yilmaz Ö. Looking in the Porphyromonas gingivalis cabinet of curiosities: the microbium, the host and cancer association. Mol Oral Microbiol. 2014;29(2):55–66.
6. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. Cell Host Microbe. 2013;14(2):195–206.
7. Han YW. Fusobacterium nucleatum: a commensal-turned pathogen. Curr Opin Microbiol. 2015;23:141–147.
8. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol. 2017;35(9):833–844.
9. Beghini F, McIver LJ, Blanco-Míguez A, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. eLife. 2021;10:e65088.
10. Yang CY, Yeh YM, Yu HY, Chin CY, Hsu CW, Liu H, Huang PJ. Oral microbiota community dynamics associated with oral squamous cell carcinoma staging. Front Microbiol. 2022;13:806318.
11. Li Y, Tan X, Zhao X, Xu Y, Liu J. Oral microbiota features in oral cancer and precancer: A systematic review. Front Oncol. 2021;11:616452.
12. Gholizadeh P, Eslami H, Yousefi M, et al. Role of oral microbiome on oral cancers, a review. Biomed Pharmacother. 2016;84:552–558.
13. Escapa IF, Chen T, Huang Y, Gajare P, Dewhirst FE, Lemon KP. New insights into human nostril microbiome from the expanded Human Oral Microbiome Database (eHOMD): a resource for microbiome studies of health and disease. mSystems. 2018;3(6):e00187–18.
14. Zhou W, Sailani MR, Contrepois K, et al. Longitudinal multi-omics of host–microbe dynamics in prediabetes. Nature. 2019;569(7758):663–671.
15. Groeger S, Meyle J. Oral mucosal epithelial cells. Front Immunol. 2019;10:208.
16. Louis P, Flint HJ. Formation of propionate and butyrate by the human colonic microbiota. Environ Microbiol. 2017;19(1):29–41.
17. Rohart F, Gautier B, Singh A, Lê Cao KA. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput Biol. 2017;13(11):e1005752.
18. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453–457.
19. Zhang Y, Kurupati R, Liu L, et al. Enhancing CD8+ T cell response to viral infection with DNA vaccines. J Clin Invest. 2013;123(10):4373–4383.
20. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr Protoc Mol Biol. 2015;109:21.29.1–9.
21. Pasolli E, Truong DT, Malik F, Waldron L, Segata N. Machine learning meta-analysis of large metagenomic datasets: Tools and biological insights. PLoS Comput Biol. 2016;12(7):e1004977.
22. Ren Z, Jin L, Zhang Y, et al. Multi-omics study reveals a new biomarker panel for oral squamous cell carcinoma. Front Oncol. 2020;10:1113.
23. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. Nat Rev Microbiol. 2014;12(10):661–672.
24. Diener C, Gibbons SM, Resendis-Antonio O. MICOM: Metagenome-scale modeling to infer metabolic interactions in the gut microbiota. mSystems. 2020;5(1):e00606–19.
25. Moncada R, Barkley D, Wagner F, et al. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. Nat Biotechnol. 2020;38(3):333–342.
26. Gopinath D, Dey P, Subbannayya T, et al. Spatial and molecular profiling of tumor microenvironment in oral squamous cell carcinoma. JCI Insight. 2021;6(20):e144906.
27. Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. Microbiome. 2015;3:31.
28. Proctor LM, Creasy HH, Fettweis JM, et al. The Integrative Human Microbiome Project. Nature. 2019;569(7758):641–648.