Article

Applications for Circulating Cell-Free DNA in Oral Squamous Cell Carcinoma: A Non-Invasive Approach for Detecting Structural Variants, Fusions, and Oncoviruses

Mahua Bhattacharya [1,†], Dan Yaniv [2,†], Dylan P. D'Souza [1,7] Eyal Yosefof [2], Sharon Tzelnick [3], Rajesh Detroja [1,8], Tal Wax [4], Adva Levy-Barda [5,6], Gideon Baum [1], Aviram Mizrachi [2], Gideon Bachar [2]* and Milana Frenkel Morgenstern [1,7]*

[1] Azrieli Faculty of Medicine, Bar Ilan University, Safed, Israel.
[2] Rabin Medical Research Centre, Tel Aviv, Israel.
[3] Division of Head and Neck Surgery, Princess Margaret Cancer Center, University of Toronto, Toronto, ON, Canada.
[4] Department of Pathology, Rabin Medical Centre, Petah Tikva, Israel.
[5] Biobank, Department of Pathology, Rabin Medical Centre; Petah Tikva, Israel.
[6] Department of Digital Medical Technologies, Holon Institute of Technology, Holon, Israel.
[7] **Current address:** Scojen Institute of Synthetic Biology, Reichman University, Herzliya, Israel.
[8] **Current address:** Princess Margaret Cancer Centre - University Health Network, Toronto, Canada
† Authors contributed equally.
* Correspondence: author: milana.morgenstern@runi.ac.il and to gideonb@clalit.org.il

Summary: A comprehensive study analyzed circulating cell-free DNA (cfDNA) from 30 oral squamous cell carcinoma (OSCC) patients, 199 public OSCC, and 192 normal samples to explore its potential in early-stage diagnostics and prognosis. The findings suggest that cfDNA concentration correlates with tumor stage, malignancy, and survival prognosis in OSCC patients.
Deep genome sequencing of cfDNA revealed several genomic alterations:

- *CNVs and SNVs:* Amplifications and deletions were observed at loci 1q, 2q, 3p, 3q, and 8q22. Notably, amplifications of TP53 and PIK3CA, along with other genes related to keratinization, were identified in relapsed OSCC cases. Single nucleotide variations were observed in genes highly mutated in OSCC.
- *Fusion Gene Identification:* A novel fusion gene, TRMO-TRNT1, was detected in seven high-grade tumor samples. The parental genes, TRMO and TRNT1, are involved in tRNA modification and DNA repair, respectively.
- *Viral Integrations*: Integrations of human papillomavirus, simian virus, and enterovirus were found in the OSCC samples, potentially indicating viral involvement in OSCC pathogenesis.

These results underscore the utility of cfDNA as a non-invasive biomarker for OSCC, offering insights into tumor genomics and aiding in early detection and prognosis assessment.

Abstract: Circulating cell-free DNA (cfDNA) has been widely used as a prognostic marker for different cancers. In this study, we used 30 cfDNA samples from oral squamous cell carcinoma (OSCC), 199 public OSCC and 192 normal samples to study various correlation factors that could improve the early-stage diagnostics and/or prognosis of OSCC. We found that OSCC patient cfDNA concentration can serve as an indicator of tumor stage, malignancy, and survival prognosis. Deep genome sequencing of cfDNA revealed genomic alterations, such as copy number variations, fusion gene identification, and viral integrations. A copy number variation analysis suggested correlation with amplification and deletion in chromosomes at loci 1q, 2q, 3p, 3q, and chromosome 8 at loci q22. Moreover, at these loci, amplification of *TP53, PIK3CA,* and other genes related to keratinization in OSCC patients was observed. In addition, we identified a novel abundant fusion gene, *TRMO-TRNT1 'chimera',* in seven high-grade tumor samples. The parental genes of this chimera, *TRMO* and *TRNT1*, are known to play roles in tRNA modification and DNA repair, respectively. We have identified structural variants in our OSCC cohort. Some of these SNVs like *KMT2C, MUC3A, MUC6* have been identified as common SNVs in different HNSCC populations. Finally, we detected integrations of human papillomavirus, simian virus, and enterovirus in the OSCC samples, which may point to the origins of OSCC. Therefore, our study indicates that liquid

biopsy may thus serve as a sensitive tool to study OSCC patient genomic alterations by exploring cfDNA circulating in the plasma, employing an easy-to-use blood test.

Introduction

Oral squamous cell carcinoma (OSCC) considered a part of head and neck squamous cell carcinoma (HNSCC) [1], affects the oral and buccal regions, including the gums, the floor of the mouth, and the tongue, palate, and lips [2]. OSCC is the 16th most common form of cancer and the 15th highest contributor to morbidity rates due to cancer [3]. Oral cancer is mostly associated with lifestyle, being linked to smoking and drinking, although some authors suggest that pre-disposition to human papilloma virus (HPV) is also linked to OSCC[3–6]. While early diagnosis can help prevent metastasis and progression, OSCC often goes undiagnosed until a late stage when symptoms are observed, making radiotherapy and chemotherapy less effective [7].

The discovery of non-invasive methods, like liquid biopsy, has led to progress in cancer diagnostics[8], specifically improving the diagnosis of lung, blood, colorectal, and cervical cancers, among others. At the same time, liquid biopsy analysis of cell-free DNA (cfDNA) in plasma, saliva and urine has helped determine the survival prognosis [8–12] of patients. CfDNA fragments appear in the plasma because of cellular apoptosis or necrosis. The amount of released cfDNA correlates with tumor burden and metastasis [13–18]. These approaches offer rapid, repeatable sampling with minimal risk to the patient. As technology advances, liquid biopsy is increasingly being explored for early detection, treatment monitoring, and recurrence prediction across a broader spectrum of malignancies [8–12]. The cfDNA analyses have been already conducted on HNSCC patients, specifically, on OSCC patients[19–21]. Human papilloma virus (HPV) is one of the major causes of HNSCC, with a significantly mutated *HPV* gene being a major cause of HNSCC[22–24]. The presence of HPV and other pathogens as revealed by cfDNA analyses of OSCC patients has made diagnosis and prognosis at early stages possible 1,14,25–31.

Previous studies reported the presence of a second virus, herpes simplex virus (HSV), in HNSCC and other squamous cell carcinomas, like cervical cancer, and correlated the presence of HPV and HSV with oncogenesis of these cancers [31]. Since OSCC is one of the most common head and neck cancers, we were interested in analyzing cfDNA from patients belonging to an Israeli cohort to identify genomic alterations associated with the disease, such as copy number variation (CNV), single nucleotide variation (SNV), integrated viral sequences, and fusion genes, along with correlation with tumor stage for better disease prognosis [23].

In this study, we analysed cfDNA from 30 patients diagnosed with oral squamous cell carcinoma (OSCC) and identified a statistically significant positive Pearson correlation (at 5% false discovery rate) between cfDNA concentration and tumour stage, the emergence of secondary malignancies, and lesion size. Additionally, we observed that elevated cfDNA levels were associated with increased recurrence rates and reduced survival probability in OSCC patients. Our genomic analyses revealed frequent copy number variations (CNVs) in cfDNA at loci containing key cancer-related genes, including *TP53*, *TP63*, *FGFR1*, and *MLH1*. We have identified structural variants like insertion, deletions and frameshift mutations in our cohort. Common SNVs identified in our samples are KMT2C, MUC3A and MUC6, which are previously reported in our cohort. Furthermore, the presence of oncogenic viral sequences—such as human papillomavirus (HPV), herpes simplex virus (HSV), simian virus (SV), and enterovirus (EV), was detected in several patient samples, suggesting a potential role for viral integration in the oncogenic progression of OSCC. Notably, we identified a novel fusion gene, *TRMO-TRNT1*, representing a chimera of two genes involved in tRNA processing. *TRMO* (located on chromosome 9q22) encodes a tRNA methyltransferase, while *TRNT1* (on chromosome 3p26) is responsible for the addition of CCA nucleotides to tRNA molecules. According to GTEx data, both genes are ubiquitously expressed across human tissues. Although *TRMO* has previously been linked to thyroid cancer, its role in OSCC has yet to be elucidated. *TRNT1*, while not currently associated with cancer, is known to harbor mutations that cause sideroblastic anaemia. To validate our findings, we analysed 199 publicly available

RNA-seq datasets from OSCC patients and 192 from healthy controls. Approximately 50% of the chimeric transcripts identified in OSCC liquid biopsies were also detected in public OSCC biopsy datasets. Notably, ten chimeras exhibited expression levels as high as 88 RPKM, underscoring their potential relevance. Collectively, our findings contribute to a more comprehensive understanding of the molecular landscape of OSCC. The identification of recurrent cfDNA alterations, viral integrations, and novel chimeric transcripts may support the development of minimally invasive biomarkers for early diagnosis, prognostication, and personalized therapeutic strategies in OSCC.

Materials And Methods

Methodology for Sample Size Collection

The sample size in our study was determined based on the availability of clinically well-characterized OSCC patients and matched healthy controls within the approved timeframe and ethical guidelines. A total of 30 OSCC patients were enrolled, representing diverse tumour stages and anatomical sub-sites, along with 25 healthy individuals serving as controls. All participants provided informed consent, and the study was conducted under the approval of the relevant institutional ethics committee (Helsinki committee approval - OSCC #0813-16-RMC). The inclusion criteria ensured representation across primary and secondary malignancies, and sample collection was standardized to minimize technical variability. This sample size allowed for preliminary statistical analyses and correlation assessments between cfDNA levels and clinical parameters, to expand in future cohorts for higher statistical power.

Sample collection

Blood samples of 30 OSCC patients were drawn before surgery and collected in cfDNA preservation tubes at the Rabin Medical Center. Blood samples were centrifuged at **300 × g** for **20 minutes** at room temperature (RT) to separate the plasma and PBMCs. The plasma samples were then collected, aliquoted into Eppendorf tubes, and stored at **-80°C** until use.

cfDNA extraction

Aliquots (1 mL) of plasma from 30 patients were processed for cfDNA extraction using the Qiagen QIAamp ccfDNA/RNA Kit (Cat. No. 55184), in accordance with the manufacturer's protocol and following informed patient consent. All procedures were conducted under the approval of the institutional ethics committee.

Inclusion Criteria:

- Patients diagnosed with histopathologically confirmed oral squamous cell carcinoma (OSCC)
- Age ≥ 18 years
- Availability of complete clinical and demographic data
- Ability to provide informed consent
- No prior systemic cancer therapy (chemotherapy or radiotherapy) at the time of sample collection

Exclusion Criteria:

- Patients with other concurrent malignancies
- History of autoimmune or chronic inflammatory diseases
- Recent major surgical procedures or trauma within the past month
- Inadequate sample quality or volume for cfDNA analysis

Appropriate buffers were added at the specified volume per 1 mL of serum, vortexed, and incubated as instructed. Centrifugation was performed at 12,000g for 3 minutes to pellet the precipitate. The supernatant was then transferred to a new tube, and ice-cold isopropanol was added

in a 1:1 ratio, briefly vortexed, and centrifuged at 5,000g for 1 minute. The washing step was conducted at 5,000g for 5 minutes. cfDNA was eluted in 50 µL of elution buffer.

Extracted cfDNA concentrations were measured using the Qubit dsDNA High Sensitivity Assay (Thermo Fisher Scientific, Waltham, MA) with a Qubit 2.0 fluorometer. Additionally, a Bioanalyzer 2100 DNA High Sensitivity Assay was performed to determine the cfDNA fragment size distribution.

Copy number analyses

Extracted cfDNA samples from all patients were sequenced using Illumina high-throughput whole genome sequencing (at least 50 million reads per sample, NextSeq 550). Paired-end sequencing was performed. For copy number analysis, the iChorCNA35 and Gistic2.036 pipelines were used with the default parameters.

Chimera identification and validation

Using the *in-house* chimera detection pipeline ChiTaH 37, novel fusion genes were identified in sequenced cfDNA samples. To validate the ChiTaH results, primers were designed upstream and downstream of the gene-gene junction sequence of a fusion gene (See supplementary Table2) PCR was performed using Hylab Hy-Taq Ready Mix (2X) (Cat no. EZ3006/7/636) (For PCR conditions refer supplementary Table 3). PCR products were purified using a Macherey-Nagel NucleoSpin Gel and PCR Clean-up Kit (Cat no: 12303368), followed by the Sanger sequencing using Macrogen.

Gene-Interaction network analysis:

GeneMania38 was used to find the gene regulators of TRNT1 and TRMO genes. Only co-expressed and physical interaction partners were used for the study.

Identification of coding potential

The coding potential of a chimera was determined using the CPAT 39and CNIT40 webserver tools, employing the default parameters.

Virus detection in cfDNA

The VirusFinder2.0 pipeline was employed to detect the presence of oncogenic viruses in circulating cfDNA samples, using its default parameters [41]. This tool integrates sequence alignment and viral-host fusion detection to accurately identify viral integration events, enabling the characterization of viral contributions to tumorigenesis. Its application in this study supports a non-invasive approach to uncovering virus-associated cancers and complements genomic analyses aimed at identifying molecular drivers of disease.

SNV analysis:

SNV analysis was performed using three tools. Varscan42 using default parameters for analysis. Only exonic SNVs were considered for the study. The SNVs from VarScan were annotated using ANNOVAR43. DeepVariant44 and Mutect245. SNVs that were identified by all three tools were selected for the study.

Gene ontology analysis

Gene ontology analyses were performed using PantherDB 46 software.

Statistical analysis

Statistical analysis was performed using GraphPad Prism. A Shapiro-Wilk test was used for various correlation analyses. The t-test was performed to analyze the significant difference between the cfDNA concentration of healthy and OSCC patients. Survival analysis was done using a

Kaplan-Meyer curve, included in the Prism GraphPad software. *p-value* was considered at ≤0.05 for statistical significance at 5% FDR.

Results

It is known that high concentrations of cfDNA were observed in tissue injuries, cancer, and various autoimmune diseases. The demography of all 30 OSCC patients and 25 healthy individuals were considered in our studies (see Supplementary Tables 1A). The demographic parameters of all 30 OSCC patients with clinical history are mentioned. (see Supplementary Tables 1B). The cfDNA samples of OSCC and healthy individuals were extracted and subjected to Bioanalyzer™ evaluations (Figure S1, Supplementary Material). Firstly, we studied the difference in cfDNA concentration of OSCC patients vs. healthy individuals and identified that cfDNA in OSCC patients were significantly higher than the cfDNA concentrations of healthy individuals (Figure 1A) (*p-value <0.05,* t-Test). Thereby suggesting that OSCC patients have a significantly higher cfDNA concentration than healthy individuals. Next, we sought to understand the correlation of increased cfDNA levels with various tumorigenic parameters, like tumor stage, a second malignancy, patient survival and more. In this study, cancers of oral cavity sub-sites samples were considered. These included twenty tongues, eight alveolar ridges, one floor of the mouth, and one retromolar trigone. Eleven patients showed mandibular involvement at presentation, meaning the bone of the mandible was invaded by the tumor at diagnosis, nine had undergone segmental mandibulectomy and two had experienced marginal mandibulectomy procedures. To check if increased cfDNA levels had direct correlation with tumor stage, we performed the Pearson correlation analyses. We observed that even though the average cfDNA concentrations in patients with tumor stage 3 or 4 were high, as compared to patients with tumor stages 1 and 2, the overall distributions were similar, other than a few cases with stage 3 and 4 tumors that presented very high concentrations of cfDNA in their blood samples (Figure 1B). These results indicate that cfDNA concentrations specify the presence of stage 3 or 4 OSCC tumors (5% FDR.

Fig 1A

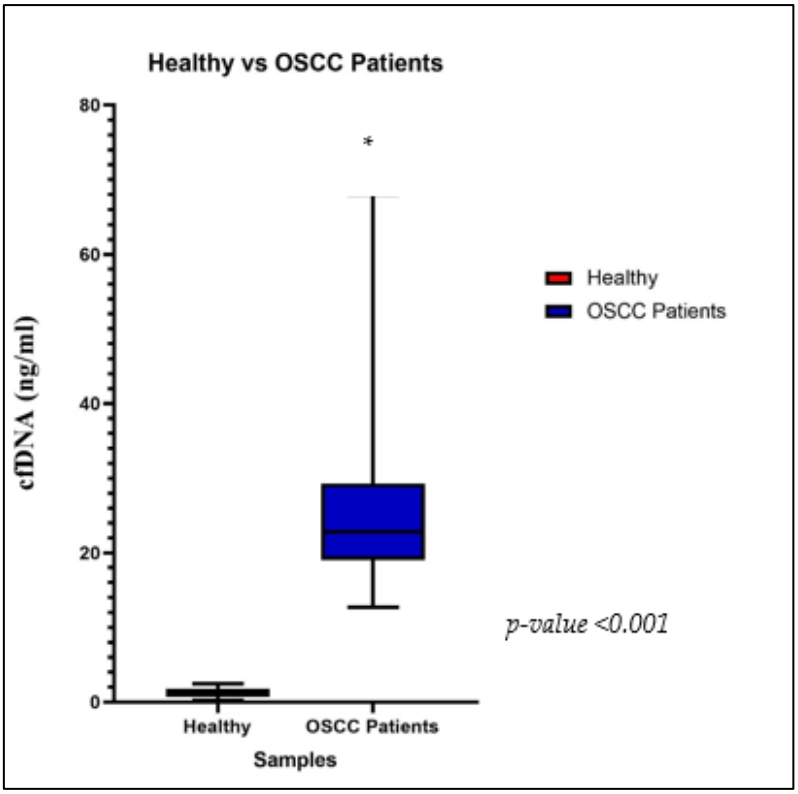


Fig 1B

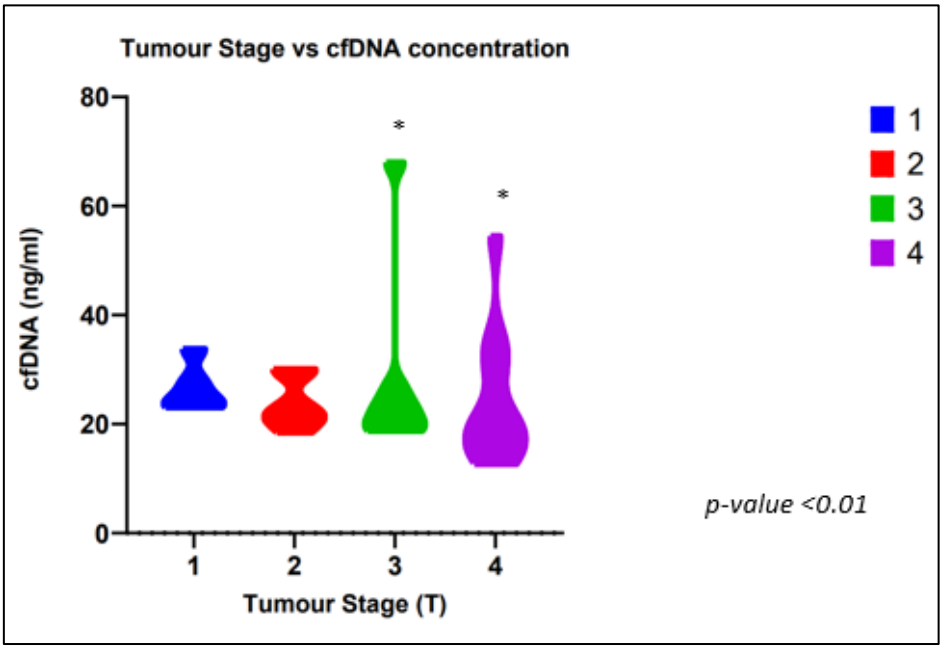


Fig 1C

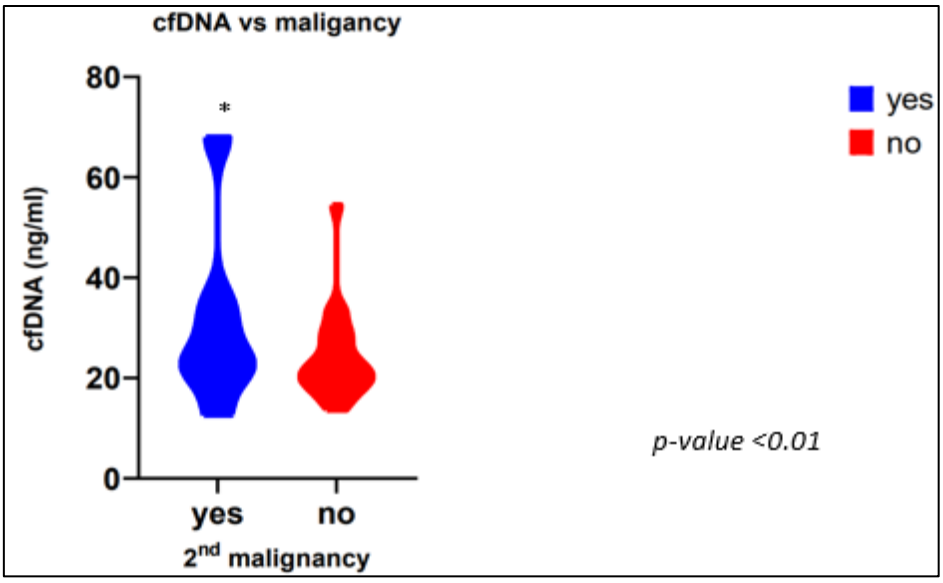


Fig 1D

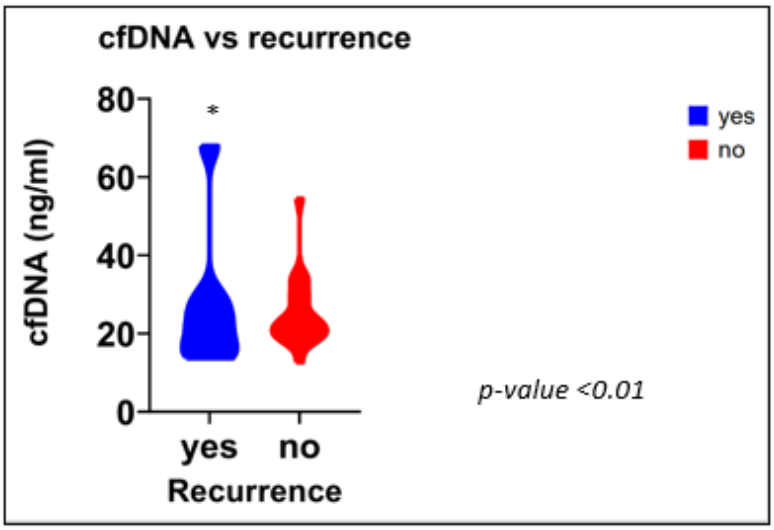


Fig 1E

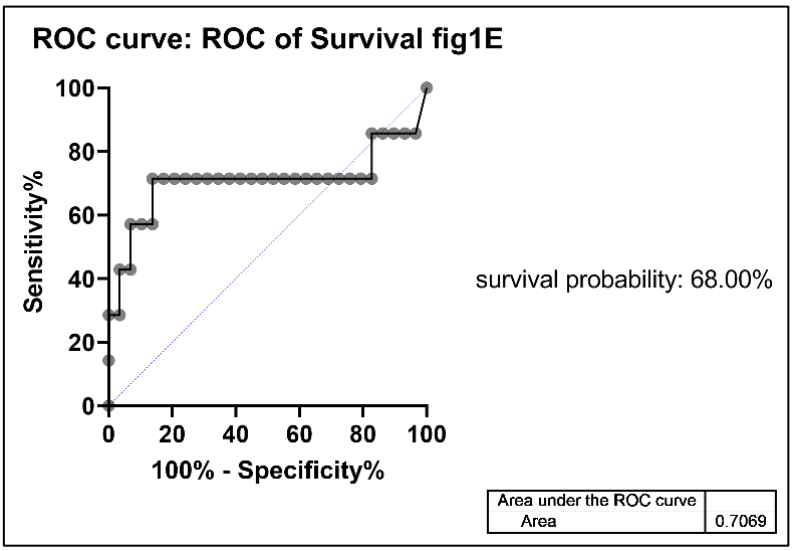


Fig 1F

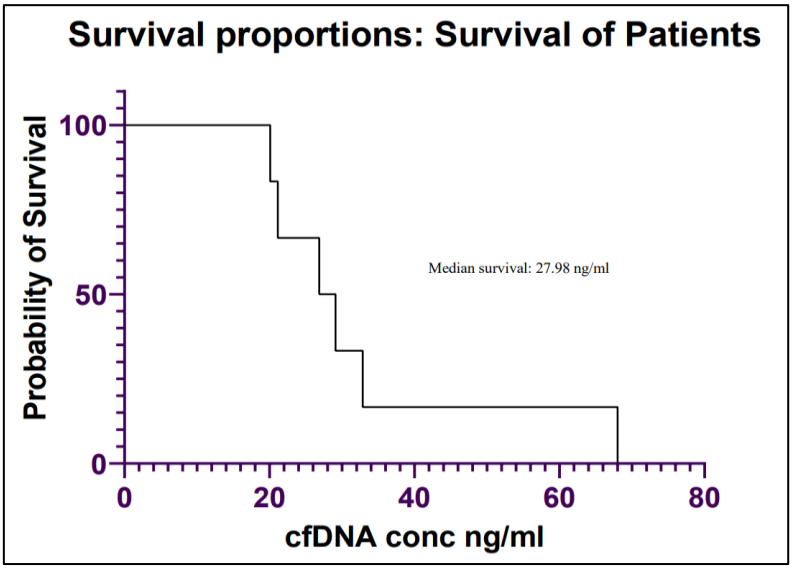


Figure 1. Correlation of cfDNA concentration with different tumor profiles. A. cfDNA concentration of healthy vs OSCC patients (*p-value ≤0.05)* using t-test. B. cfDNA concentration compared with tumor stage

using Shapiro-Wilk normality. C. cfDNA vs. secondary malignancy. D. cfDNA vs. recurrence. E. The probability of survival was determined using a survival curve, with the threshold of survival being 27 ng/ml. F. Survival analysis using the Kaplan-Meyer method showed 68% probability of survival within the cohort.

We also observed similar results with secondary vs. primary malignancies and observed that patients with secondary malignancies had high cfDNA concentration at *p-value <0.05* (Figure 1C). Furthermore, the increased cfDNA concentration was significantly associated with a higher risk of disease recurrence in OSCC patients (Figure 1D). Then, using the Kaplan-Meyer survival analysis method on data describing the OSCC recurrence and death vs. cfDNA concentration, we found that patients with cfDNA concentrations of 27 ng/ml or less had a higher survival probability (68%; Figures 1E and 1F). These results indicate that cfDNA concentrations above a threshold of 27 ng/ml can be used to recognize the tumor stages 3 and 4, which, furthermore, can be associated with decreased survival. This study thus helped in identifying the probability of survival and underlining the tumor stages using the cfDNA concentration as a biomarker. In future, this method could be used as an even more sensitive biomarker for the early detection of cancer.

Oncogenic viruses were identified when analyzing cfDNA

In the present study, we detected partial contigs of HPV and other oncogenic viruses. Specifically, nine of 30 OSCC patients showed the presence of viral DNA integration into cfDNA (Figure 2A). Moreover, we observed sequences from HSV strains 1, 3 and 5, SV, HPV and EV (Figure 2A). HSV integrations were identified in six OSCC patients in the present study, three patients had HSV-1, and three patients had HSV-5 integrations. The relation of EV with cancer has not been completely established, however, some evidence shows a link to malignancy in gliomas and to progression of cervical cancers. SV has been identified in the serum of lung cancer patients. However, SV and EV have not yet been recognized as causing OSCC. These results indicate that our study revealed the presence of more than two viruses in OSCC patient sera (Figures 2A and 2B). However, the status of these viruses cannot yet be statistically correlated with the tumor stage in our cohort. Nonetheless, the presence of viruses may serve some oncogenic role in OSCC, and more samples should be used to identify their characteristics in OSCC patients.

Fig 2A

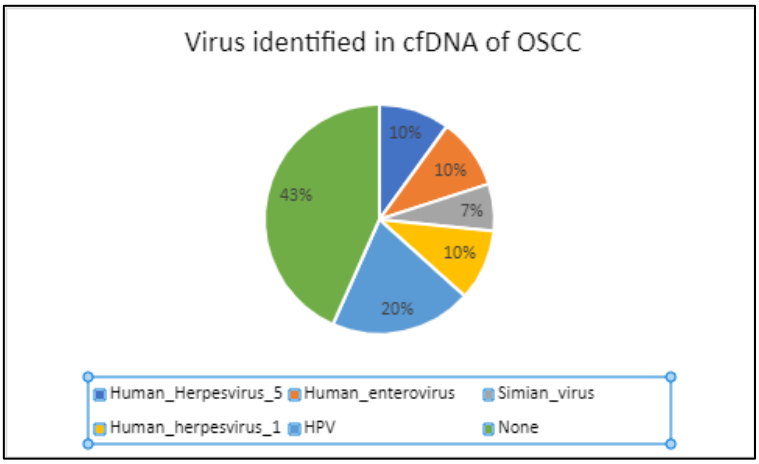


Fig 2B

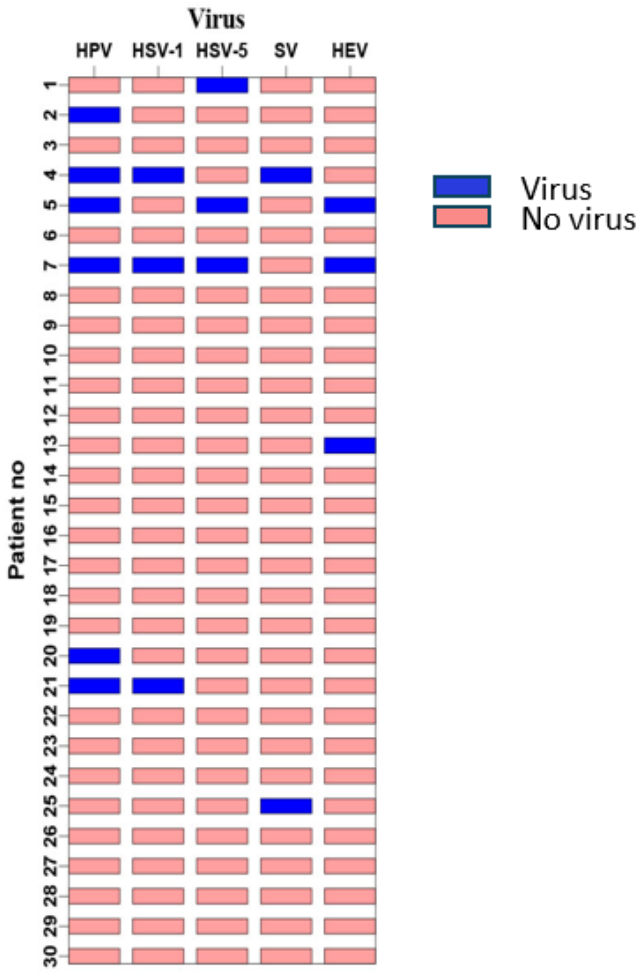


Figure 2. Identification of viral DNA in cfDNA samples. A. Pie chart showing the percentage of samples containing viral DNA. B. Presence of different viral DNA contigs in the samples. Some patients had DNA from more than one virus in their sample.

A novel chimera identified in cfDNA

Fusion genes produced by chromosomal translocations were common in lymphomas and leukemias, yet are extremely rare in HNSCC 23,47. Using our *in-house* chimera analysis tool ChiTaH, we identified the novel fusion *TRMO-TRNT1* in the cfDNA samples of two patients (sequence., Figure S2, Supplementary Material). TRMO (tRNA methyl-O-transferase) and TRNT1 (tRNA nucleotide transferase I) were highly expressed in all tissue types. To validate this chimera at both RNA and DNA levels, we designed primers and performed PCR on total RNA and DNA from OSCC biopsies and used patient 3 and patient 5 as a positive control (Figure 3). Subsequently, we conducted PCR on all the 30 cfDNA samples and observed a 108 bp fused sequence in seven of them (Figure S3A, Supplementary Material). The band containing this fusion was excised from the agarose gel, purified, and assessed by Sanger sequencing, which confirmed the fusion junction of the two parental genes, *TRMO* and *TRN1* (Figure S3B, Supplementary Material). The fusion was identified in seven patients using PCR against the junction sequence and the Sanger sequencing result confirming the same. Showing the tumor advanced malignancy and a proliferation status. The altered expression of *TRMO* has been observed in patients with thyroid carcinoma 33, while altered *TRNT1* expression was seen in colorectal cancer patients 35. These results suggest that fusion genes could be used as biomarkers and explored in the search for additional OSCC patients who could be revealed by means of the deep sequencing of cfDNA samples. Finally, we checked the coding potential of the observed chimera to produce a functional protein. For this, we utilized three tools commonly used for analyzing the coding/non-coding potential of a given DNA

sequence. The scores obtained by CPAT and CNIT programs suggested that the fusion has a <0.3% coding probability. Therefore, the fusion does not have coding potential (Figures S4A and S4B) and may serve as a biomarker of the OSCC progression through a highly abundant non-coding cfDNA fragment. Next, we checked for the common interactors of *TRNT1-TRMO* genes (Figure 3B). Three common interactors identified are *THADA, COMMD2* and *CTBP2.* In addition, we downloaded 199 OSCC and 192 normal RNA-seq datasets from public databases (Supplementary Data). We found that 50% of the chimeras identified in the OSCC plasma samples were also observed in the public biopsies of OSCC patients. Furthermore, 10 chimeras had abundances as high as 88 RPKM. Our findings may help further characterize novel abundant chimeras as biomarkers in OSCC in early diagnostics and survival prognostics.

Fig 3A

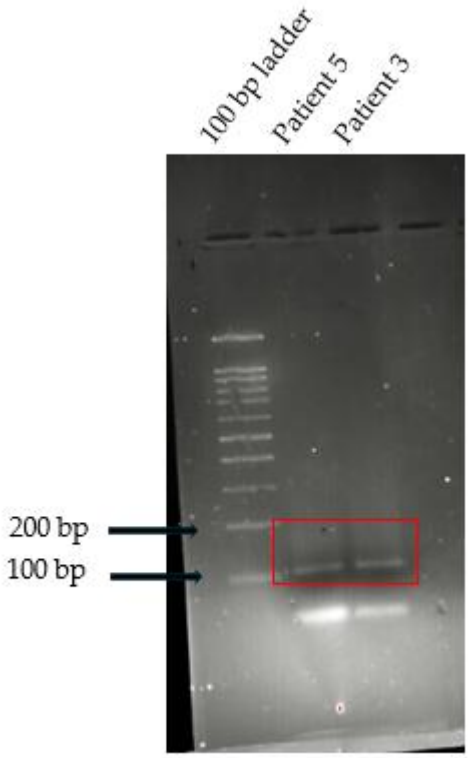


Fig 3B

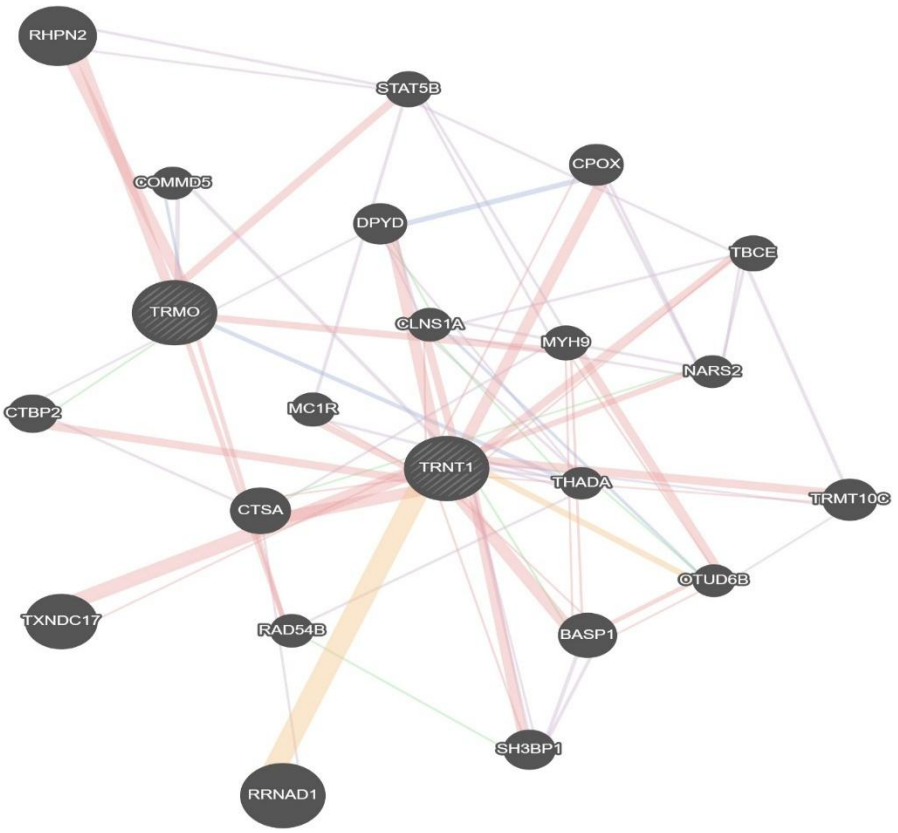


Figure 3. Identification of a novel fusion. A. Agarose gel of PCR products from two confirmed patient samples with the fusion highlighted by a red box. B. Gene-gene interaction network analysis. The orange line represents co-expressed genes, purple line represents physical interactions. *CTBP2, COMMD2, THADA* are the common interactors of the parental genes.

Copy number variation analysis of OSCC patient cfDNA

Copy Number Variation (CNV) is commonly observed in many cancers. Various studies have identified CNV and somatic mutations using cfDNA14,48–51. Our aim here was to identify significant changes in copy numbers in OSCC patients using cfDNA genomic data. We performed CNV analysis on the sequenced cfDNA of OSCC patients using Gistic2.0, which assesses low pass copy number variations. The Q value was set at 0.01 for level significance. We observed significant amplifications in chromosomes 3q22.2, 7q35, and 15q26. Additional amplifications were also seen in chromosomes 8p21.3, 10q26,2, 2p25 and 1q21.3 (Figure 4A). Some of our findings correlate with TCGA HNSCC data for CNV and significantly mutated genes. Next, we performed functional enrichment analysis using PantherDB 46 and observed that commonly amplified genes in the present study were predicted to play a significant role in GPCR pathways, opioid pathways, immune system activation pathways, apoptosis, and angiogenesis activation pathways (Figures 4C). Particularly, we identified deleterious regions in patient cfDNA samples. In our cohort, we observed that chromosomes 1q23, 2q22, 3p21, 3q26, 7q31, 8p21, 9q34, 13q13 and 14q21 had significant deletions (Figure 4B). Significantly mutated genes, like *TP53*, *TP63*, *FGFR1*, or *MLH1*, have been shown to present deletions in those regions and were also identified in the present study 23,52,53. We also found deleted regions of *THADA* in our cohort. Amplification of regions of *CTBP2* genes has been observed in the present study. Both *THADA* and *CTBP2* are common interactors of fusion genes *TRNT1-TRMO* (Figure 3B). Our gene ontology (GO) analysis of amplification regions found the deleterious genes in our samples assigned to the affected genes as being mostly involved in immune activation pathways, apoptosis signaling pathways, particularly, the *CCKR* gene, cell cycle pathways and various metabolic pathways (Figure 4D). These results indicate that cfDNA analyses may reveal significant CNV in the OSCC samples and may be used for the OSCC early diagnosis and prognosis.

Fig 4A

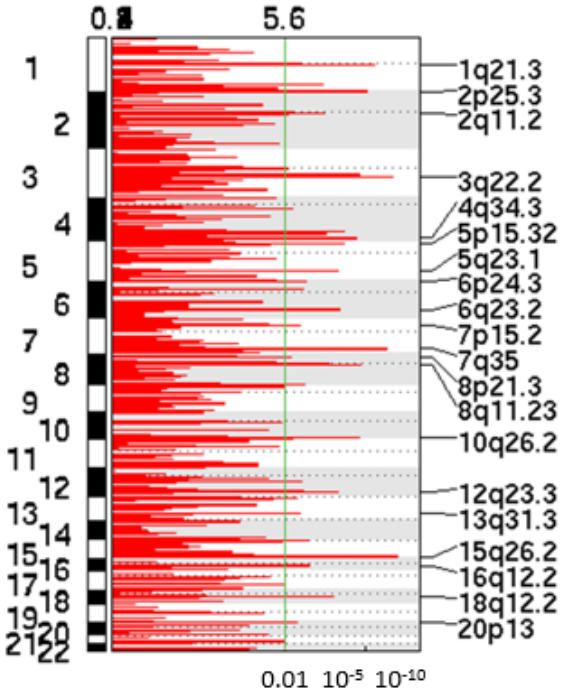


Fig 4B

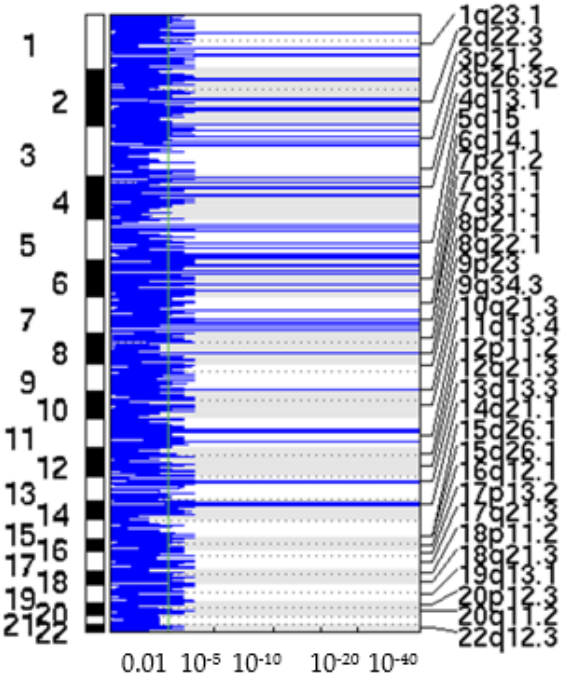


Figure 4 C

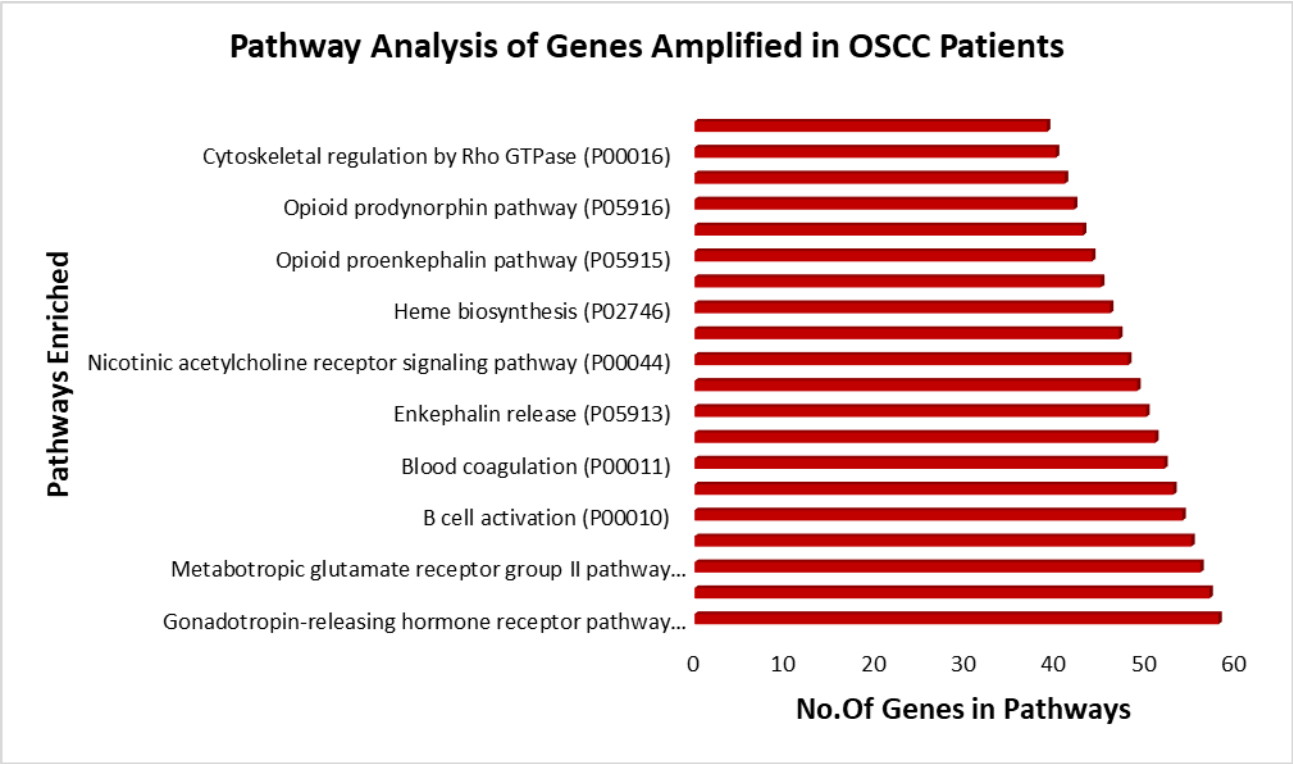


Figure 4D

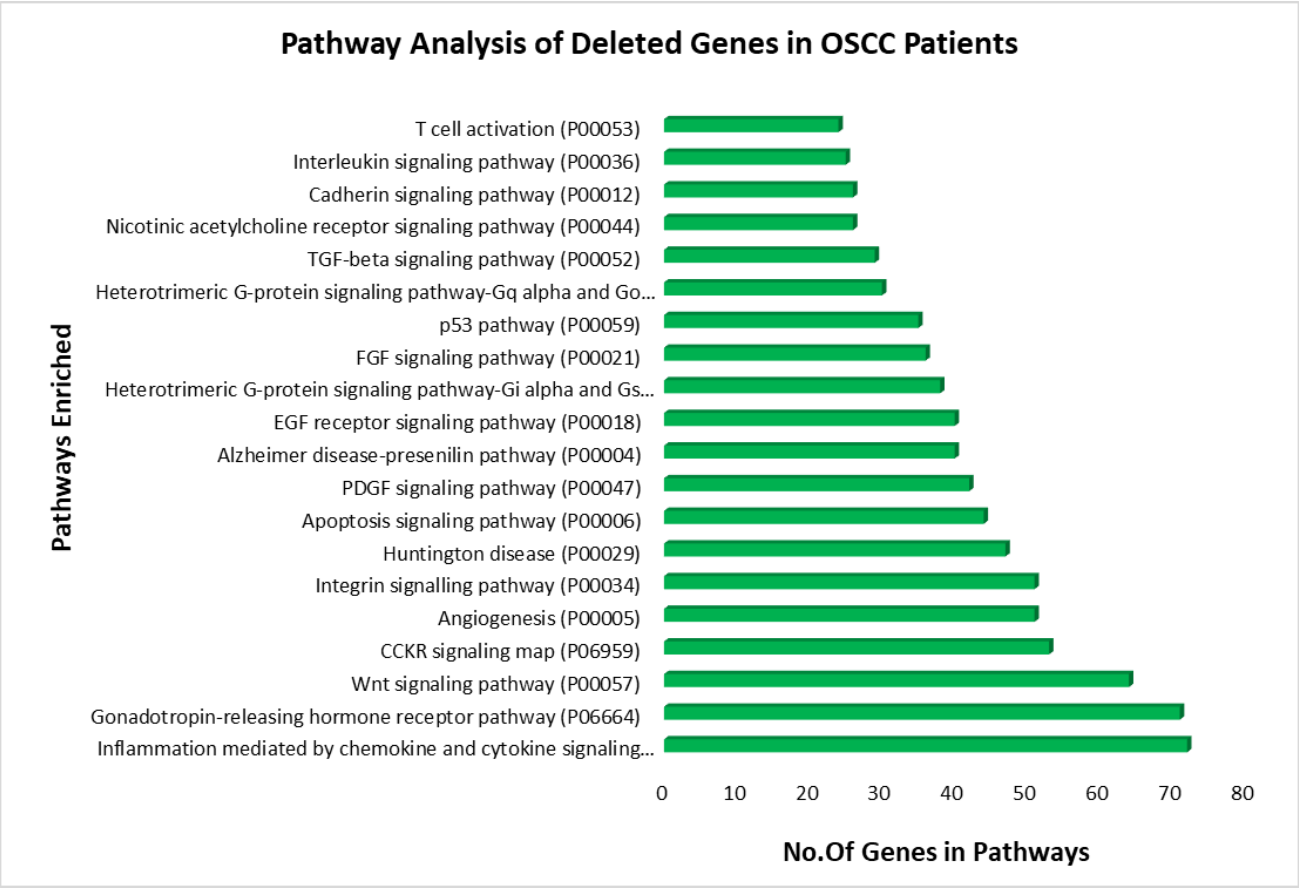


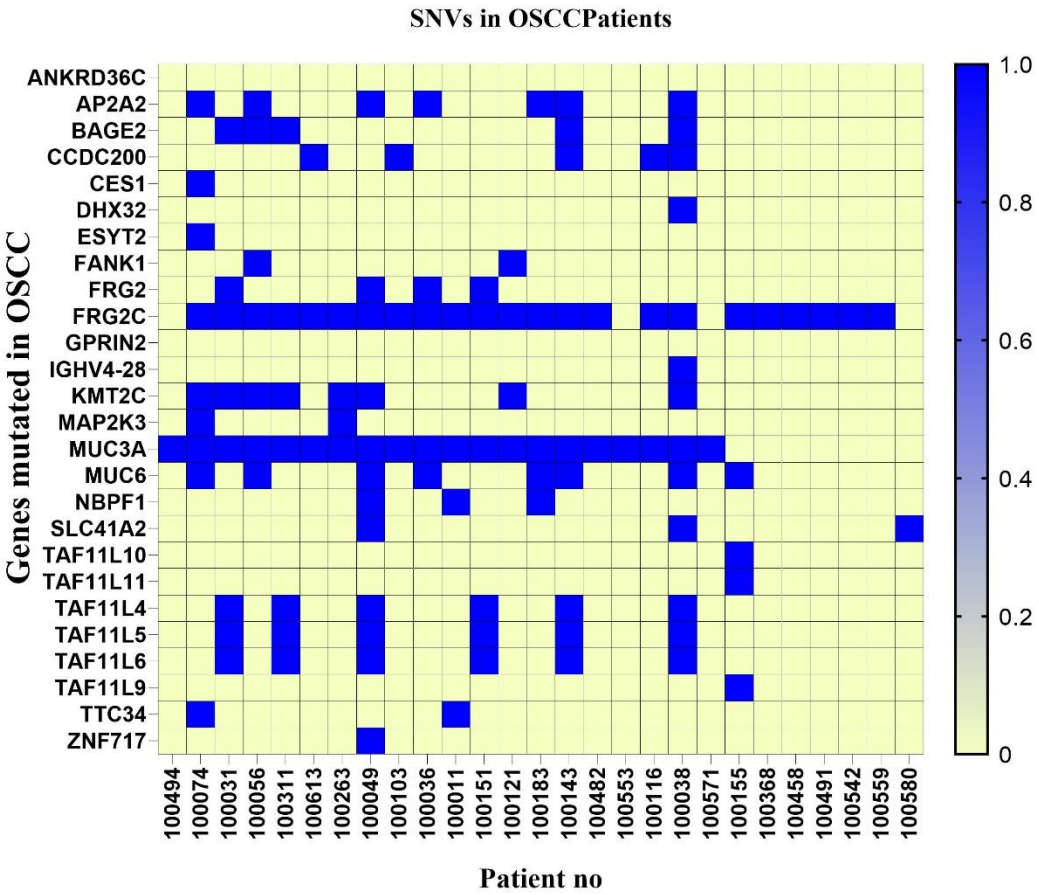
Figure 4. Copy number variation analyses of cfDNA sequences: CNV analysis was done on 30 patient cfDNA samples using Gistic2.0. An amplification plot suggests (A) amplification in the regions 8p21.3, 10q26,2, 2p25 and 1q21.3 and (B) deletions in chromosomes 1q23, 2q22, 3p21, 3q26, 7q31, 8p21, 9q34,

13q13 and 14q21. C, D. Gene ontology analysis of amplified (C) and deleted regions (D) using PantherDB. Genes related to various signaling pathways and the apoptosis pathway were enriched.

SNV Analysis of OSCC Patients' cfDNA:

SNV analysis suggests structural variants and is a key to identifying mutations that cause cancer54. We used three tools to perform the SNV analysis of the cfDNA of OSCC patients. We identified common SNVs present in our samples, such as *MUC6* (27%), *MUC3A* (70%), *FRG2C* (80%), *KMT2C* (26%), *AP2A2* (23%), *BAGE (17%),* and *CCDC200 (20%)* in our cohort. *TAF11L4, TAF11L5,* and *TAF11L6* (23% each) (Figure 5). In the present study, we considered non-synonymous mutations, frameshifts, and in-Del mutations in exonic regions. Most mutations observed are C>T, C>A, G>T, G>A. Some deletions in *MUC3A* and *FRGC2* have been observed. These observations are in concordance with other studies reported for HNSCC. These mutations can be used to diagnose OSCC.

Approximately 4% of SNVs found had deletions in *KMT2C* and *MUC3A*, 1% was frameshift mutations was observed in *MUC6* and 20% were non-synonymous substitution mutations observed majorly in *KMT2C, MUC6, MUC3A, ZNF717* and *SLC41A2*(Table 8). These mutations have been identified as significant mutations across various studies in HNSCC and OSCC. The mutations in these genes have been identified to play oncogenic roles in HNSCC.



s

Figure 5. SNV analysis of OSCC patients. The heatmap shows SNVs identified in our OSCC cohort. Blue indicates SNVs in the sample, and yellow indicates no SNVs identified. On the Y-axis are genes that are mutated in our cohort. *MUC3A, MUC6, FRGC2, TAF11,* and *KMT2C* have high mutations in patients' samples.

TABLE 1: SNVs in OSCC samples	

Mutations	Observed frequencies
Deletions	4%
Non-synonymous	20%
Framshift	1%
Insertion	0.1%
Transversion	1.2%

Table1: shows frequencies of non-synonymous mutations in OSCC patients' cfDNA.

Discussion

cfDNA has been explored as a prognostic and diagnostic biomarker in many cancers 12,16,17,19,30,55. Correlation of cfDNA levels with various tumors has been established in glioblastoma and neuroblastoma. High cfDNA concentrations were observed in samples with cancer, as compared to cfDNA from healthy samples 12,56. Many studies have found that the concentration of cfDNA was correlated with tumor burden 12,31,56. In the present study, we observed the cfDNA concentration of OSCC patients was significantly higher than cfDNA concentration of healthy patients. This result suggests that cfDNA concentration of OSCC patients could be used for early and late diagnostics. Then, we explored the correlation between the cfDNA concentration from 30 OSCC patients and tumor stages. High cfDNA levels were observed in patients with stage 3 and 4 tumors, while comparatively lower cfDNA concentrations were observed in patients with stage 1 and 2 tumors. We also observed a direct correlation between higher cfDNA concentration at presentation and the possibility of future recurrence and a secondary malignancy at 5% FDR. We also observed that the threshold of maximum cfDNA concentration for prognostication was <27 ng/ml. The Kaplan-Meyer survival analysis showed that the probability of survival to be 68% for patients with OSCC at cfDNA concentrations above 27 ng/ml. Therefore, our study suggests that cfDNA could be used as a non-invasive method to detect OSCC at an early stage. Still, the analysis of more samples could have resulted in a better level of significance for the correlation between the cfDNA concentration and the tumor parameters.

Various genomic alterations in the cfDNA of OSCC patients have also been addressed. One such alteration was observed as the presence of viral DNA integrations. Deep sequencing of the cfDNA of OSCC patients revealed the presence of HPV, SV, HSV and EV integrations. While a high proportion of patients (48%) did not show the presence of any virus, a few patients presented more than one virus. Interestingly, HPV has been studied as a cause of HNSCC, cervical cancer and other gynecological and urological cancers 4,14,20, 31, 57,58. We identified HPV in six out of 30 samples. HSV strains 1 and 5 were identified in six of 30 samples. HSV was previously reported in samples from HNSCC and OSCC patients and its association with the progression of cancer was claimed 59–61 Enterovirus has been associated with malignant glioma62. SV was reported in samples of osteosarcoma and prostate cancer 63–65. However, these viruses have not yet been associated with OSCC. Future studies could establish the association of these viruses with OSCC and determine if they are involved in oncogenic progression.

Using the *in-house* ChiTaH pipeline37, we identified the novel fusion *TRMO-TRNT1*. PCR and Sanger sequencing validated the presence of this fusion in seven samples from patients with late-stage tumors. Mutation in *TRMO* has been associated with thyroid carcinoma 33. *TRNT1* has not been associated with any cancer thus far, although its mutation is responsible for causing congenital disease and sideroblastic anemia 34 The fusion product does not have any coding potential, according to CPAT 39 and CNIT 40. Gene regulatory studies revealed three potential regulators of both TRNT1 and TRMO: THADA, CTBP2, and COMMD2. Future research could explore the mechanism of action of this fusion gene in OSCC development.

CNV was considered. CNV in tissues, as well as in cfDNA, has been studied and was shown to be associated with many cancers, including OSCC. 48. In the present study, we found regions of chromosomes 3q22.2, 7q35, and 15q26 to be amplified. Amplifications were also seen in

chromosomes 8p21.3, 10q26,2, 2p25 and 1q21.3. Gene Ontology pathway analysis of genes at these loci revealed them to be associated with angiogenesis, immune activation pathways, and GPCR, interleukin, VEGF, FGF, and vasopressin activation pathways. The deleterious regions at chromosomes 1q23, 2q22, 3p21, 3q26, 7q31, 8p21, 9q34, 13q13 and 14q21 affect genes associated with apoptosis, *CCKR*, cell cycle, and immune activation and cell signaling pathways. Mutations in these genes and structural variations associated with these loci have been studied in TCGA and by other groups. Our study agrees with TCGA and previous studies 23,53,66,67. We also identified deleted regions of *THADA* in our cohort. Amplification of regions of the *CTBP2* gene was observed in the present study. Both *THADA* and *CTBP2* are common regulators of the fusion gene *TRNT1-TRMO*. In the future, it will be interesting to investigate whether structural variations in *THADA, COMMD2*, and *CTBP2* play a significant role in OSCC progression, and whether the fusion gene *TRNT1-TRMO* is involved in CNV. Additionally, targeted sequencing of these genes or CNV analysis of these chromosomal loci could be used as potential biomarkers for OSCC. Additionally, from the 199 public OSCC RNA-seq datasets, we found that 50% of the chimeras identified in OSCC liquid biopsies were also observed in public biopsies. Furthermore, 10 chimeras exhibited abundances as high as 88 RPKM and may serve as potential biomarkers for OSCC. These findings could help explore their association with cancer and enhance applications in early diagnostics and survival prognostics.

SNV analysis helps in understanding key mutations leading to oncogenesis23,47,55,68. They help in identifying potential risk mutations that can lead to carcinogenesis. SNVs can be a result of various factors, including environmental factors, smoking status, chewing betel quid and viral integration status69–72. SNV analysis was performed in the present study to identify significant mutations in genes. Our study suggested common mutations in genes such as *MUC6* (27%)*, MUC3A* (70%)*, FRG2C* (80%)*, KMT2C* (26%)*, AP2A2* (23%)*, BAGE (17%),* and *CCDC200 (20%)* in our cohort. Mutational profiles of *TAF11L4, TAF11L5,* and *TAF11L6* (23% each) have not been reported in HNSCC. Previous studies reported the Mucin genes were altered in OSCC. Most mutated Mucin genes are *MUC1, MUC4, MUC3, MUC6*[70]. *MUC1* and *MUC4* have been identified as prognostic biomarkers for OSCC, which leads to oral mucosa carcinogenesis 74–76. *MUC6* have been recently identified as a potential biomarker in Asian OSCC patients. The mutations in *MUC6* have been associated with malignancy and tumorigenesis of oral carcinoma73,74,77. Recently, *MUC3A* has also been associated with the progression of oral cancer. Mutations in *MUC3A* have been linked to oncogenesis 73,74,78. Another study showed that *ZNF717* and *MUC3A* are mutated in tongue squamous cell carcinoma 78. In the present study, we identified SNV of *ZNF717* in 1 sample. *FRG2C* is a signature gene mutated in ESCC (esophageal squamous cell carcinoma) and proliferative verrucous leukoplakia (PVL)79,80. Another gene highly mutated in the present study cohort, *KMT2C*, was observed to be rarely mutated in the South Asian patients with HNSCC 79,81. *BAGE* was identified as a clinically significant immunogenic antigen in HNSCC patients 82. *AP2A2,* an Adapting protein family molecule, has not been reported to be oncogenic. However, it has been linked to inhibiting the activity of Erlotinib, an anti-cancer drug for lung cancer 83.

All these mutational signatures of genes can be drug targets for precision medicines. Also, these genes can be used to study further, the relevance of mutation in OSCC progression and thereby used as a biomarker for non-invasive cancer diagnosis and prognosis.

Conclusion

Our study highlights the potential of circulating cfDNA as a non-invasive biomarker for the diagnosis, prognosis, and therapeutic monitoring of OSCC. Elevated cfDNA levels correlated significantly with tumor stage and patient outcomes, suggesting its value in early detection and recurrence risk assessment. We also uncovered a spectrum of genomic alterations in cfDNA, including viral integrations, chimeric transcripts, copy number variations (CNVs), and single nucleotide variants (SNVs), many of which have known or emerging relevance to OSCC progression. Notably, we identified and validated a novel fusion gene, TRMO-TRNT1, associated with late-stage tumors,

and observed recurring mutations in genes such as MUC3A, MUC6, FRG2C, and KMT2C, with implications for oncogenesis and potential as therapeutic targets. CNV analysis revealed altered regions linked to immune response and oncogenic pathways, supporting the use of these alterations as molecular biomarkers. The consistency of several chimeric transcripts with public datasets further underscores their potential clinical relevance. The limitation of the present study is the low number of samples studied. Altogether, our findings establish a robust foundation for integrating cfDNA-based assays into OSCC diagnostics and precision medicine frameworks, while pointing toward novel molecular targets for future therapeutic exploration. Continued research with larger cohorts and functional validation will be essential to fully realize the translational potential of these biomarkers.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/doi/s1, Figure S1: **Bioanalyzer graph of cfDNA size of OSCC patients**; Figure S2: **Fusion sequence**; Figure S3: **Fusion validation in other OSCC samples**; Figure S4**Coding potential of the fusion sequence**; Table S1A: **Demographic information of healthy vs OSCC patients** ; Table S1B: **Details of OSCC patients,** Table S2: **Fusion primers,** Table S3: **PCR conditions for fusions** File S1: Somatic Variant Calling, Filtering, and Annotation. File S2: OSCC_copy umber variations. File S3: Fusions in 30 OSCC samples. File S4: Fusions in 192 OSCC patients from public dataset. References [x,x] are cited in the supplementary materials..

Author Contributions: Conceptualisation: MB, DY, GB, GBc, and MFM ; Methodology: MB, GB, and MFM ; Software: DPD, and RD ; Validation: MB, GB, DYP and RJ ; formal analysis: MB, DY, GBc and MFM ; resources: MFM, GBc, ST, TW, EY, and ALB ; investigation: MB, DY, and MFM ; data curation: DPD and RJ ; writing-original draft: MB, and DY ; Writing-review and editing: MB, DY, DPD, GB, GBc, and MFM ; supervision: MFM and GBc ; Project administration: MFM ; Funding acquisition: MFM and GBc. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Bar-Ilan University core funding and Funding from Rabin Medical Centre

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Rabin Medical Centre (protocol code OSCC #0813-16-RMCand date of approval: 17th October 2019)."

Informed Consent Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Rabin Medical Centre (protocol code OSCC #0813-16-RMCand date of approval: 17th October 2019)"

Data Availability Statement: All data are available upon publication

Conflicts of Interest: The Authors state no conflict of interest

References

[1] Sinevici, N.; O'sullivan, J. Oral Cancer: Deregulated Molecular Events and Their Use as Biomarkers. *Oral Oncol*, **2016**, *61*, 12–18. https://doi.org/10.1016/J.ORALONCOLOGY.2016.07.013.

[2] Montero, P. H.; Patel, S. G. Cancer of the Oral Cavity. *Surgical Oncology Clinics of North America*. W.B. Saunders July 1, 2015, pp 491–508. https://doi.org/10.1016/j.soc.2015.03.006.

[3] Inchingolo, F.; Santacroce, L.; Ballini, A.; Topi, S.; Dipalma, G.; Haxhirexha, K.; Bottalico, L.; Charitos, I. A. Oral Cancer: A Historical Review. *International Journal of Environmental Research and Public Health*. MDPI AG May 1, 2020. https://doi.org/10.3390/ijerph17093168.

[4] Katirachi, S. K.; Grønlund, M. P.; Jakobsen, K. K.; Grønhøj, C.; von Buchwald, C. The Prevalence of HPV in Oral Cavity Squamous Cell Carcinoma. *Viruses*. MDPI February 1, 2023. https://doi.org/10.3390/v15020451.

[5] Hernandez, B. Y.; Lynch, C. F.; Chan, O. T. M.; Goodman, M. T.; Unger, E. R.; Steinau, M.; Thompson, T. D.; Gillison, M.; Lyu, C.; Saraiya, M. Human Papillomavirus DNA Detection, P16 INK4a , and Oral Cavity Cancer in a U.S. Population. *Oral Oncol*, **2019**, *91*, 92–96. https://doi.org/10.1016/j.oraloncology.2019.03.001.

[6] *Oral Human Papilloma Virus Infection: An Overview of Clinical-Laboratory Diagnosis and Treatment*.

[7] Nakagaki, T.; Tamura, M.; Kobashi, K.; Omori, A.; Koyama, R.; Idogawa, M.; Ogi, K.; Hiratsuka, H.; Tokino, T.; Sasaki, Y. Targeted Next-Generation Sequencing of 50 Cancer-Related Genes in Japanese Patients with Oral Squamous Cell Carcinoma. *Tumor Biology*, **2018**, *40* (9). https://doi.org/10.1177/1010428318800180.

[8] Sumanasuriya, S.; Lambros, M. B.; de Bono, J. S. Application of Liquid Biopsies in Cancer Targeted Therapy. *Clin Pharmacol Ther*, **2017**, *102* (5), 745–747. https://doi.org/10.1002/CPT.764.

[9] Kang, Z.; Stevanovic, S.; Hinrichs, C. S.; Cao, L. Circulating Cell-Free DNA for Metastatic Cervical Cancer Detection, Genotyping, and Monitoring. *Clinical Cancer Research*, **2017**, *23* (22), 6856–6862. https://doi.org/10.1158/1078-0432.CCR-17-1553.

[10] Cordeiro Mitchell, C. N.; Murdock, T.; Fader, A. N.; Stone, R. L. Advanced Ovarian Cancer Treated in Pregnancy and Detected by Cell-Free DNA Aneuploidy Screening. *Gynecol Oncol Rep*, **2018**, *24* (March), 48–50. https://doi.org/10.1016/j.gore.2018.03.006.

[11] Wang, X.; Wang, L.; Su, Y.; Yue, Z.; Xing, T.; Zhao, W.; Zhao, Q.; Duan, C.; Huang, C.; Zhang, D.; et al. Plasma Cell-Free DNA Quantification Is Highly Correlated to Tumor Burden in Children with Neuroblastoma. *Cancer Med*, **2018**, *7* (7), 3022–3030. https://doi.org/10.1002/cam4.1586.

[12] Fawzy, A.; Sweify, K. M.; El-Fayoumy, H. M.; Nofal, N. Quantitative Analysis of Plasma Cell-Free DNA and Its DNA Integrity in Patients with Metastatic Prostate Cancer Using ALU Sequence. *J Egypt Natl Canc Inst*, **2016**, *28* (4), 235–242. https://doi.org/10.1016/j.jnci.2016.08.003.

[13] Wang, Y.; Springer, S.; Mulvey, C. L.; Silliman, N.; Schaefer, J.; Sausen, M.; James, N.; Rettig, E. M.; Guo, T.; Pickering, C. R.; et al. Detection of Somatic Mutations and HPV in the Saliva and Plasma of Patients with Head and Neck Squamous Cell Carcinomas. *Sci Transl Med*, **2015**, *7* (293). https://doi.org/10.1126/SCITRANSLMED.AAA8507.

[14] Grisanti, S.; Almici, C.; Consoli, F.; Buglione, M.; Verardi, R.; Bolzoni-Villaret, A.; Bianchetti, A.; Ciccarese, C.; Mangoni, M.; Ferrari, L.; et al. Circulating Tumor Cells in Patients with Recurrent or Metastatic Head and Neck Carcinoma: Prognostic and Predictive Significance. *PLoS One*, **2014**, *9* (8). https://doi.org/10.1371/JOURNAL.PONE.0103918.

[15] Krebs, M. G.; Metcalf, R. L.; Carter, L.; Brady, G.; Blackhall, F. H.; Dive, C. Molecular Analysis of Circulating Tumour Cells - Biology and Biomarkers. *Nat Rev Clin Oncol*, **2014**, *11* (3), 129–144. https://doi.org/10.1038/NRCLINONC.2013.253.

[16] Di Meo, A.; Bartlett, J.; Cheng, Y.; Pasic, M. D.; Yousef, G. M. Liquid Biopsy: A Step Forward towards Precision Medicine in Urologic Malignancies. *Mol Cancer*, **2017**, *16* (1). https://doi.org/10.1186/S12943-017-0644-5.

[17] Yu, M.; Stott, S.; Toner, M.; Maheswaran, S.; Haber, D. A. Circulating Tumor Cells: Approaches to Isolation and Characterization. *Journal of Cell Biology*, **2011**, *192* (3), 373–382. https://doi.org/10.1083/JCB.201010021.

[18] Alpaugh, R. K.; Zhou, L. L.; Dicker, D. T.; Matthew, E.; El-Deiry, W. S. Circulating Tumor Cells: Silent Predictors of Metastasis. *F1000Res*, **2017**, *6*. https://doi.org/10.12688/F1000RESEARCH.11313.1.

[19] Mazurek, A. M.; Rutkowski, T.; Fiszer-Kierzkowska, A.; Małusecka, E.; Składowski, K. Assessment of the Total CfDNA and HPV16/18 Detection in Plasma Samples of Head and Neck Squamous Cell Carcinoma Patients. *Oral Oncol*, **2016**, *54*, 36–41. https://doi.org/10.1016/J.ORALONCOLOGY.2015.12.002.

[20] Xiao, X.; Liu, S.; Deng, H.; Song, Y.; Zhang, L.; Song, Z. Advances in the Oral Microbiota and Rapid Detection of Oral Infectious Diseases. *Frontiers in Microbiology*. Frontiers Media S.A. February 6, 2023. https://doi.org/10.3389/fmicb.2023.1121737.

[21] Lousada-Fernandez, F.; Rapado-Gonzalez, O.; Lopez-Cedrun, J.-L.; Lopez-Lopez, R.; Muinelo-Romay, L.; Suarez-Cunqueiro, M. Liquid Biopsy in Oral Cancer. *Int J Mol Sci*, **2018**, *19* (6), 1704. https://doi.org/10.3390/ijms19061704.

[22] Bodelon, C.; Untereiner, M. E.; Machiela, M. J.; Vinokurova, S.; Wentzensen, N. Genomic Characterization of Viral Integration Sites in HPV-Related Cancers. *Int J Cancer*, **2016**, *139* (9), 2001–2011. https://doi.org/10.1002/ijc.30243.

[23] Lawrence, M. S.; Sougnez, C.; Lichtenstein, L.; Cibulskis, K.; Lander, E.; Gabriel, S. B.; Getz, G.; Ally, A.; Balasundaram, M.; Birol, I.; et al. Comprehensive Genomic Characterization of Head and Neck Squamous Cell Carcinomas. *Nature*, **2015**, *517* (7536), 576–582. https://doi.org/10.1038/nature14129.

[24] Morgan, I. M.; Dinardo, L. J.; Windle, B. Integration of Human Papillomavirus Genomes in Head and Neck Cancer: Is It Time to Consider a Paradigm Shift? *Viruses*, **2017**, *9* (8). https://doi.org/10.3390/v9080208.

[25] Saproo, S.; Sarkar, S. S.; Gautam, V.; Konyak, C. W.; Dass, G.; Karmakar, A.; Sharma, M.; Ahuja, G.; Gupta, A.; Tayal, J.; et al. Salivary Protein Kinase C Alpha and Novel MicroRNAs as Diagnostic and Therapeutic Resistance Markers for Oral Squamous Cell Carcinoma in Indian Cohorts. *Front Mol Biosci*, **2023**, *9*. https://doi.org/10.3389/fmolb.2022.1106963.

[26] Arantes, L. M. R. B.; De Carvalho, A. C.; Melendez, M. E.; Carvalho, A. L. Serum, Plasma and Saliva Biomarkers for Head and Neck Cancer. *Expert Review of Molecular Diagnostics*. Taylor and Francis Ltd 2017, pp 85–112. https://doi.org/10.1080/14737159.2017.1404906.

[27] Lin, L. H.; Chang, K. W.; Kao, S. Y.; Cheng, H. W.; Liu, C. J. Increased Plasma Circulating Cell-Free DNA Could Be a Potential Marker for Oral Cancer. *Int J Mol Sci*, **2018**, *19* (11). https://doi.org/10.3390/ijms19113303.

[28] Chattopadhyay, I.; Verma, M.; Panda, M. Role of Oral Microbiome Signatures in Diagnosis and Prognosis of Oral Cancer. *Technol Cancer Res Treat*, **2019**, *18*. https://doi.org/10.1177/1533033819867354.

[29] Patel, S.; Shah, K.; Mirza, S.; Shah, K.; Rawal, R. Circulating Tumor Stem like Cells in Oral Squamous Cell Carcinoma: An Unresolved Paradox. *Oral Oncol*, **2016**, *62*, 139–146. https://doi.org/10.1016/J.ORALONCOLOGY.2016.10.019.

[30] Hamana, K.; Uzawa, K.; Ogawara, K.; Shiiba, M.; Bukawa, H.; Yokoe, H.; Tanzawa, H. Monitoring of Circulating Tumour-Associated DNA as a Prognostic Tool for Oral Squamous Cell Carcinoma. *Br J Cancer*, **2005**, *92* (12), 2181–2184. https://doi.org/10.1038/SJ.BJC.6602635.

[31] Adalsteinsson, V. A.; Ha, G.; Freeman, S. S.; Choudhury, A. D.; Stover, D. G.; Parsons, H. A.; Gydush, G.; Reed, S. C.; Rotem, D.; Rhoades, J.; et al. Scalable Whole-Exome Sequencing of Cell-Free DNA Reveals High Concordance with Metastatic Tumors. *Nature Communications 2017 8:1*, **2017**, *8* (1), 1–13. https://doi.org/10.1038/s41467-017-00965-y.

[32] Mermel, C. H.; Schumacher, S. E.; Hill, B.; Meyerson, M. L.; Beroukhim, R.; Getz, G. GISTIC2.0 Facilitates Sensitive and Confident Localization of the Targets of Focal Somatic Copy-Number Alteration in Human Cancers. *Genome Biol*, **2011**, *12* (4), 1–14. https://doi.org/10.1186/GB-2011-12-4-R41/FIGURES/7.

[33] Detroja, R.; Gorohovski, A.; Giwa, O.; Baum, G.; Frenkel-Morgenstern, M. ChiTaH: A Fast and Accurate Tool for Identifying Known Human Chimeric Sequences from High-Throughput Sequencing Data. *NAR Genom Bioinform*, **2021**, *3* (4). https://doi.org/10.1093/NARGAB/LQAB112.

[34] Warde-Farley, D.; Donaldson, S. L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C. T.; et al. The GeneMANIA Prediction Server: Biological Network Integration for Gene Prioritization and Predicting Gene Function. https://doi.org/10.1093/nar/gkq537.

[35] Wang, L.; Park, H. J.; Dasari, S.; Wang, S.; Kocher, J. P.; Li, W. CPAT: Coding-Potential Assessment Tool Using an Alignment-Free Logistic Regression Model. *Nucleic Acids Res*, **2013**, *41* (6). https://doi.org/10.1093/nar/gkt006.

[36] Guo, J. C.; Fang, S. S.; Wu, Y.; Zhang, J. H.; Chen, Y.; Liu, J.; Wu, B.; Wu, J. R.; Li, E. M.; Xu, L. Y.; et al. CNIT: A Fast and Accurate Web Tool for Identifying Protein-Coding and Long Non-Coding Transcripts Based on Intrinsic Sequence Composition. *Nucleic Acids Res*, **2019**, *47* (W1), W516–W522. https://doi.org/10.1093/nar/gkz400.

[37] Koboldt, D. C.; Chen, K.; Wylie, T.; Larson, D. E.; Mclellan, M. D.; Mardis, E. R.; Weinstock, G. M.; Wilson, R. K.; Ding, L. VarScan: Variant Detection in Massively Parallel Sequencing of Individual and Pooled Samples. *BIOINFORMATICS APPLICATIONS NOTE*, **2009**, *25* (17), 2283–2285. https://doi.org/10.1093/bioinformatics/btp373.

[38] Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data. https://doi.org/10.1093/nar/gkq603.

[39] Poplin, R.; Chang, P. C.; Alexander, D.; Schwartz, S.; Colthurst, T.; Ku, A.; Newburger, D.; Dijamco, J.; Nguyen, N.; Afshar, P. T.; et al. A Universal SNP and Small-Indel Variant Caller Using Deep Neural Networks. *Nature Biotechnology 2018 36:10*, **2018**, *36* (10), 983–987. https://doi.org/10.1038/nbt.4235.

[40] Benjamin, D.; Sato, T.; Cibulskis, K.; Getz, G.; Stewart, C.; Lichtenstein, L. Calling Somatic SNVs and Indels with Mutect2. https://doi.org/10.1101/861054.

[41] Mi, H.; Ebert, D.; Muruganujan, A.; Mills, C.; Albou, L. P.; Mushayamaha, T.; Thomas, P. D. PANTHER Version 16: A Revised Family Classification, Tree-Based Classification Tool, Enhancer Regions and Extensive API. *Nucleic Acids Res*, **2021**, *49* (D1), D394–D403. https://doi.org/10.1093/NAR/GKAA1106.

[42] Stransky, N.; Egloff, A. M.; Tward, A. D.; Kostic, A. D.; Cibulskis, K.; Sivachenko, A.; Kryukov, G. V.; Lawrence, M. S.; Sougnez, C.; McKenna, A.; et al. The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. *Science (1979)*, **2011**, *333* (6046), 1157–1160. https://doi.org/10.1126/science.1208130.

[43] Kulkarni, O.; Sugier, P. E.; Guibon, J.; Boland-Augé, A.; Lonjou, C.; Bacq-Daian, D.; Olaso, R.; Rubino, C.; Souchard, V.; Rachedi, F.; et al. Gene Network and Biological Pathways Associated with Susceptibility to Differentiated Thyroid Carcinoma. *Sci Rep*, **2021**, *11* (1), 8932. https://doi.org/10.1038/S41598-021-88253-0.

[44] Wedatilake, Y.; Niazi, R.; Fassone, E.; Powell, C. A.; Pearce, S.; Plagnol, V.; Saldanha, J. W.; Kleta, R.; Chong, W. K.; Footitt, E.; et al. TRNT1 Deficiency: Clinical, Biochemical and Molecular Genetic Features. *Orphanet J Rare Dis*, **2016**, *11* (1). https://doi.org/10.1186/S13023-016-0477-0.

[45] Shlien, A.; Malkin, D. Copy Number Variations and Cancer. *Genome Medicine*. June 16, 2009. https://doi.org/10.1186/gm62.

[46] Zhang, N.; Wang, M.; Zhang, P.; Huang, T. Classification of Cancers Based on Copy Number Variation Landscapes. *Biochim Biophys Acta Gen Subj*, **2016**, *1860* (11), 2750–2755. https://doi.org/10.1016/J.BBAGEN.2016.06.003.

[47] Tan, E. S.; Knepper, T. C.; Wang, X.; Permuth, J. B.; Wang, L.; Fleming, J. B.; Xie, H. Copy Number Alterations as Novel Biomarkers and Therapeutic Targets in Colorectal Cancer. *Cancers (Basel)*, **2022**, *14* (9). https://doi.org/10.3390/CANCERS14092223.

[48] Palande, V.; Detroja, R.; Gorohovski, A.; Glass, R.; Flueh, C.; Kanner, A. A.; Laviv, Y.; Har Nof, S.; Levy-Barda, A.; Yust-Katz, S.; et al. *Detection of Point Mutations and Gene Fusions from Circulating Cell-Free DNA of Glioma Patients*.

[49] Wang, C.-I.; Kao, H.-K.; Chen, T.-W.; Huang, Y.; Cheng, H.-W.; Yi, J.-S.; Hung, S.-Y.; Wu, C.-S.; Lee, Y.-S.; Chang, K.-P. Characterization of Copy Number Variations in Oral Cavity Squamous Cell Carcinoma Reveals a Novel Role for MLLT3 in Cell Invasiveness. *Oncologist*, **2019**, *24* (12), e1388–e1400. https://doi.org/10.1634/theoncologist.2019-0063.

[50] Rhie, A.; Park, W. S.; Choi, M. K.; Kim, J. H.; Ryu, J.; Ryu, C. H.; Kim, J. Il; Jung, Y. S. Genomic Copy Number Variations Characterize the Prognosis of Both P16-Positive and P16-Negative Oropharyngeal Squamous Cell Carcinoma after Curative Resection. *Medicine (United States)*, **2015**, *94* (50). https://doi.org/10.1097/MD.0000000000002187.

[51] Ye, Y.; Wang, J.; Liang, F.; Song, P.; Yan, X.; Wu, S.; Huang, X.; Han, P. Identification of Key Genes for HNSCC from Public Databases Using Bioinformatics Analysis. *Cancer Cell Int*, **2021**, *21* (1). https://doi.org/10.1186/s12935-021-02254-7.

[52] Perdomo, S.; Avogbe, P. H.; Foll, M.; Abedi-Ardekani, B.; Lescher Facciolla, V.; Anantharaman, D.; Chopard, P.; Le Calvez-Kelm, F.; Vilensky, M.; Polesel, J.; et al. Circulating Tumor DNA Detection in Head and Neck Cancer: Evaluation of Two Different Detection Approaches. *Oncotarget*, **2017**, *8* (42), 72621–72632. https://doi.org/10.18632/ONCOTARGET.20004.

[53] Palande, V.; Siegal, T.; Detroja, R.; Gorohovski, A.; Glass, R.; Flueh, C.; Kanner, A. A.; Laviv, Y.; Har-Nof, S.; Levy-Barda, A.; et al. Detection of Gene Mutations and Gene-Gene Fusions in Circulating Cell-Free DNA of Glioblastoma Patients: An Avenue for Clinically Relevant Diagnostic Analysis. *Mol Oncol*, **2022**, *16*, 2098. https://doi.org/10.1002/1878-0261.13157.

[54] Sastre-Garau, X.; Diop, M.; Martin, F.; Dolivet, G.; Marchal, F.; Charra-Brunaud, C.; Peiffert, D.; Leufflen, L.; Dembele, B.; Demange, J.; et al. A NGS-Based Blood Test for the Diagnosis of Invasive HPV-Associated Carcinomas with Extensive Viral Genomic Characterization. *Clinical Cancer Research*, **2021**, *27* (19), 5307–5316. https://doi.org/10.1158/1078-0432.CCR-21-0293.

[55] Costa, R. L.; Boroni, M.; Soares, M. A. Distinct Co-Expression Networks Using Multi-Omic Data Reveal Novel Interventional Targets in HPV-Positive and Negative Head-and-Neck Squamous Cell Cancer. *Sci Rep*, **2018**, *8* (1). https://doi.org/10.1038/s41598-018-33498-5.

[56] Brown, S. H.; States, V. A. R.; Afghan, A. K.; Satyanarayana, G. Herpes Simplex Virus-Infected Squamous Cell Carcinoma: A Case Report. *BMC Infect Dis*, **2022**, *22* (1). https://doi.org/10.1186/S12879-021-06995-8.

[57] Mosmann, J. P.; Talavera, A. D.; Criscuolo, M. I.; Venezuela, R. F.; Kiguen, A. X.; Panico, R.; Ferreyra De Prato, R.; López De Blanc, S. A.; ré, V.; Cuffini, C. G. Sexually Transmitted Infections in Oral Cavity Lesions: Human Papillomavirus, Chlamydia Trachomatis, and Herpes Simplex Virus. *J Oral Microbiol*, **2019**, *11* (1). https://doi.org/10.1080/20002297.2019.1632129.

[58] Wołaçewicz, M.; Becht, R.; Grywalska, E.; Niedźwiedzka-Rystwej, P. Herpesviruses in Head and Neck Cancers. *Viruses*, **2020**, *12* (2). https://doi.org/10.3390/V12020172.

[59] Zhang, X.; Wang, H.; Sun, Y.; Qi, M.; Li, W.; Zhang, Z.; Zhang, X. E.; Cui, Z. Enterovirus A71 Oncolysis of Malignant Gliomas. *Molecular Therapy*, **2020**, *28* (6), 1533–1546. https://doi.org/10.1016/j.ymthe.2020.04.005.

[60] Mazzoni, E.; Benassi, M. S.; Corallini, A.; Barbanti-Brodano, G.; Taronna, A.; Picci, P.; Guerra, G.; D'Agostino, A.; Trevisiol, L.; Nocini, P. F.; et al. Significant Association between Human Osteosarcoma and Simian Virus 40. *Cancer*, **2015**, *121* (5), 708–715. https://doi.org/10.1002/CNCR.29137.

[61] Shah, K. V.; Galloway, D. A.; Knowles, W. A.; Viscidi, R. P. Simian Virus 40 (SV40) and Human Cancer: A Review of the Serological Data. *Rev Med Virol*, **2004**, *14* (4), 231–239. https://doi.org/10.1002/RMV.432.

[62] Shah, K. V.; Weismann, M.; Murphy, G. P. Occurrence of Simian Virus 40 (SV40)-reacting Antibodies in Sera of Some Patients with Prostatic Cancer. *J Surg Oncol*, **1972**, *4* (2), 89–93. https://doi.org/10.1002/JSO.2930040203.

[63] Wu, F.; Du, Y.; Hou, X.; Cheng, W. A Prognostic Model for Oral Squamous Cell Carcinoma Using 7 Genes Related to Tumor Mutational Burden. *BMC Oral Health*, **2022**, *22* (1). https://doi.org/10.1186/s12903-022-02193-3.

[64] Wang, C.-I.; Kao, H.-K.; Chen, T.-W.; Huang, Y.; Cheng, H.-W.; Yi, J.-S.; Hung, S.-Y.; Wu, C.-S.; Lee, Y.-S.; Chang, K.-P. Characterization of Copy Number Variations in Oral Cavity Squamous Cell Carcinoma Reveals a Novel Role for MLLT3 in Cell Invasiveness. *Oncologist*, **2019**, *24* (12), e1388–e1400. https://doi.org/10.1634/theoncologist.2019-0063.

[65] Solomon, B.; Young, R. J.; Rischin, D. Head and Neck Squamous Cell Carcinoma: Genomics and Emerging Biomarkers for Immunomodulatory Cancer Treatments. *Semin Cancer Biol*, **2018**, *52* (October 2017), 228–240. https://doi.org/10.1016/j.semcancer.2018.01.008.

[66] Ghasemi, F.; Prokopec, S. D.; MacNeil, D.; Mundi, N.; Gameiro, S. F.; Howlett, C.; Stecho, W.; Plantinga, P.; Pinto, N.; Ruicci, K. M.; et al. Mutational Analysis of Head and Neck Squamous Cell Carcinoma Stratified by Smoking Status. *JCI Insight*, **2019**, *4* (1), e123443. https://doi.org/10.1172/JCI.INSIGHT.123443.

[67] Panda, B.; Krishnan, N.; Gupta, S.; Palve, V.; Varghese, L.; Pattnaik, S.; Jain, P.; Khyriem, C.; Hariharan, A.; Dhas, K.; et al. Integrated Analysis of Oral Tongue Squamous Cell Carcinoma Identifies Key Variants and Pathways Linked to Risk Habits, HPV, Clinical Parameters and Tumor Recurrence. *F1000Res*, **2015**, *4*, 1215. https://doi.org/10.12688/F1000RESEARCH.7302.1.

[68] Maitra, A.; Biswas, N. K.; Amin, K.; Kowtal, P.; Kumar, S.; Das, S.; Sarin, R.; Majumder, P. P.; Bagchi, I.; Bairagya, B. B.; et al. Mutational Landscape of Gingivo-Buccal Oral Squamous Cell Carcinoma Reveals New Recurrently-Mutated Genes and Molecular Subgroups. *Nature Communications 2013 4:1*, **2013**, *4* (1), 1–10. https://doi.org/10.1038/ncomms3873.

[69] Su, S. C.; Chang, L. C.; Lin, C. W.; Chen, M. K.; Yu, C. P.; Chung, W. H.; Yang, S. F. Mutational Signatures and Mutagenic Impacts Associated with Betel Quid Chewing in Oral Squamous Cell Carcinoma. *Hum Genet*, **2019**, *138* (11–12), 1379–1389. https://doi.org/10.1007/S00439-019-02083-9/METRICS.

[70] Mahomed, F. Recent Advances in Mucin Immunohistochemistry in Salivary Gland Tumors and Head and Neck Squamous Cell Carcinoma. *Oral Oncol*, **2011**, *47* (9), 797–803. https://doi.org/10.1016/J.ORALONCOLOGY.2011.06.008.

[71] Lu, H.; Liang, D.; Zhu, Y.; Xu, W.; Zhou, K.; Liu, L.; Liu, S.; Yang, W. Prognostic and Clinicopathological Significance of MUC Expression in Head and Neck Cancer: A Systematic Review and Meta-Analysis. *Oncotarget*, **2017**, *8* (56), 96359. https://doi.org/10.18632/ONCOTARGET.19648.

[72] John, S.; Sharma, K.; Singh, P.; Chandra, S.; Singh, G.; Gupta, S. Quantitative Proteogenomic Characterization in MUC1 and MUC4 in Oral Squamous Cell Carcinoma, Oral Potentially Malignant Disorders, and Normal Oral Mucosa in Carcinogenesis. *European Archives of Oto-Rhino-Laryngology*, **2024**, 1–10. https://doi.org/10.1007/S00405-024-09056-3/METRICS.

[73] Xia, P.; Choi, A. H.; Deng, Z.; Yang, Y.; Zhao, J.; Wang, Y.; Hardwidge, P. R.; Zhu, G. Cell Membrane-Anchored MUC4 Promotes Tumorigenicity in Epithelial Carcinomas. *Oncotarget*, **2017**, *8* (8), 14147–14157. https://doi.org/10.18632/ONCOTARGET.13122.

[74] Hua, C. H.; Chuang, C. Y.; Chien, Y. C.; Su, C. W.; Chen, S. C.; Liu, L. C.; Yang, S. F.; Yu, Y. L. Analysis of MUC6 Polymorphisms on the Clinicopathologic Characteristics of Asian Patients with Oral Squamous Cell Carcinoma. *J Cell Mol Med*, **2023**, *27* (17), 2594–2602. https://doi.org/10.1111/JCMM.17886.

[75] Zheng, X.; Sun, Y.; Li, Y.; Ma, J.; Lv, Y.; Hu, Y.; Zhou, Y.; Zhang, J. A Novel Tongue Squamous Cell Carcinoma Cell Line Escapes from Immune Recognition Due to Genetic Alterations in HLA Class I Complex. *Cells*, **2023**, *12* (1), 35. https://doi.org/10.3390/CELLS12010035/S1.

[76] Farah, C. S.; Shearston, K.; Melton, P. E.; Fox, S. A. Genome-Wide Characterization of the Mutational Landscape of Proliferative Verrucous Leukoplakia. *Oral Surg Oral Med Oral Pathol Oral Radiol*, **2024**, *138* (1), 99–111. https://doi.org/10.1016/J.OOOO.2024.04.005.

[77] Zhang, R.; Li, C.; Wan, Z.; Qin, J.; Li, Y.; Wang, Z.; Zheng, Q.; Kang, X.; Chen, X.; Li, Y.; et al. Comparative Genomic Analysis of Esophageal Squamous Cell Carcinoma among Different Geographic Regions. *Front Oncol*, **2023**, *12*, 999424. https://doi.org/10.3389/FONC.2022.999424/BIBTEX.

[78] Ghias, K.; Rehmani, S. S.; Razzak, S. A.; Madhani, S.; Azim, M. K.; Ahmed, R.; Khan, M. J. Mutational Landscape of Head and Neck Squamous Cell Carcinomas in a South Asian Population. *Genet Mol Biol*, **2019**, *42* (3), 526–542. https://doi.org/10.1590/1678-4685-GMB-2018-0005.

[79] Expression of cancer-testis antigens as possible targets for antigen-specific immunotherapy in head and neck squamous cell carcinoma. Cancer Biology & Therapy, 5(9), 1218–1225 | 10.4161/cbt.5.9.3174 https://sci-hub.se/https://www.tandfonline.com/doi/abs/10.4161/cbt.5.9.3174 (accessed Jan 30, 2025).

[80] Shin, J.; Nile, A.; Oh, J. W. Role of Adaptin Protein Complexes in Intracellular Trafficking and Their Impact on Diseases. *Bioengineered*, **2021**, *12* (1), 8259–8278. https://doi.org/10.1080/21655979.2021.1982846.