# IDENTIFICATION OF PREDICTIVE AND PROGNOSTIC BIOMARKERS OF HEAD AND NECK CANCER IN MIZO POPULATION

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

**ZOTHANZAMI**

**MZU REGN NO.: 1506904**

**PH.D. REGN NO.: MZU/Ph.D./1274 OF 08.08.2018**


**DEPARTMENT OF BIOTECHNOLOGY**

**SCHOOL OF LIFE SCIENCE**

**OCTOBER 2024**

# IDENTIFICATION OF PREDICTIVE AND PROGNOSTIC BIOMARKERS OF HEAD AND NECK CANCER IN MIZO POPULATION

BY

ZOTHANZAMI

Department of Biotechnology

Name of the Supervisor: Prof. N Senthil Kumar

Submitted

In partial fulfillment of the requirement of the Degree of Doctor of Philosophy in Biotechnology of Mizoram University, Aizawl.

# CERTIFICATE

This is to certify that the thesis entitled "**Identification of Predictive and Prognostic Biomarkers of Head and Neck Cancer in Mizo Population**" submitted to the Mizoram University; in partial fulfillment for the degree of Doctor of Philosophy in Biotechnology is a record of research work carried out by Zothanzami under my personal supervision and guidance.

No part of this thesis has been reproduced elsewhere for any degree.

Dated:

(Dr. N Senthil Kumar)
Professor
Dept. of Biotechnology
Mizoram University
Aizawl, Mizoram

**DECLARATION**

**Mizoram University**

**October 2024**

I, **Zothanzami**, hereby declare that the subject matter of this thesis is the record of work done by me, that the contents of this thesis did not form basis of the award of any previous degree to me or to the best of my knowledge to anybody else, and that the thesis has not been submitted by me for any research degree in any other University/Institute.

This is being submitted to the Mizoram University for the degree of Doctor of Philosophy in Biotechnology.

(Candidate)

(Dr. John Zothanzama)
(Head of Department)

(Dr. N Senthil Kumar)
(Supervisor)

## ACKNOWLEDGEMENT

To Almighty God, for good health, the strength and resilience to persevere through challenges.

To my supervisor, Prof. Senthil. You've been more than just a mentor, you've been a scientific sherpa, guiding me through the treacherous mountains of research. Thanks for your patience and your unwavering support even when my data went south faster than a rogue centrifuge.

I want to sincerely thank all the patients, for their willingness and trusting us with their samples and data. I am truly grateful for your role in this important journey.

To the skilled doctors and nurses and staffs (ENT department, CHA & MSCI) and technician Ms. T. Lalhriatpuii who helped in collecting samples.

To my amazing lab mates (Mr. David Zorinsanga, Mr. Souvik Ghatak, Mrs. Payel Ghatak, Dr. Lalawmpuii Pachuau, Mrs. Christine Vanlalbiakdiki Sailo, Mrs. Lallianmawii Pachuau, Mr. Ranjan Jyoti Sarma, Mr. Andrew Vanlallawma, Ms. Lalfakzuali, Mr. Lalrohlua Pachuau) for all the support, ignoring questionable lab attire choices and for reminding me that science is not all about serious faces and lab coats (at least not all the time!).

To my friends and family, you were my sanity check when spreadsheets started looking like hieroglyphics. Thanks for believing in me even when I doubted myself, for celebrating the small victories as much as the big ones.

And finally, to the one true constant throughout this entire journey, the fuel that kept me awake through countless nights and powered my brain cells: coffee, you beautiful, bitter bean juice, I owe you one.

**CONTENT**

# List of tables

## List of figures

# INTRODUCTION

## Head and Neck Cancer region

Head and neck cancers encompass a group of malignancies arising in various head and neck regions. These include the oral cavity (C00 – C06), nasopharynx (C11), oropharynx (C09 - C10), hypopharynx (C13), larynx (C32), salivary glands (C08), paranasal sinuses and nasal cavity (C30 & C31). Head and Neck Squamous Cell Carcinoma (HNSCC), a cancer originating in the epithelial cells, is the predominant type, accounting for roughly 90% of head and neck cancers (Mastronikolis et al., 2024). HNSCC includes major anatomical sites - the oral cavity, larynx, hypopharynx, nasopharynx, oropharynx and paranasal sinuses and nasal cavity. Cancers originating in the salivary glands, melanomas, sarcomas of the head and neck regions make up the remaining 10% of head and neck cancer (National Cancer Institute). Figure 1 depicts the anatomical regions of head and neck cancer.

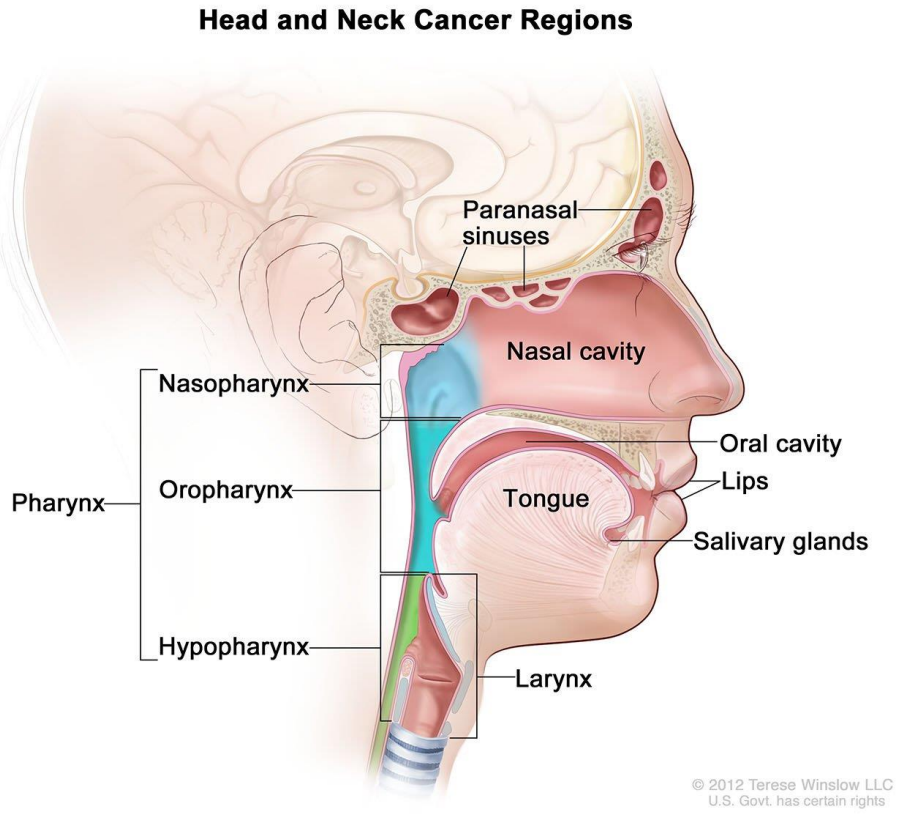
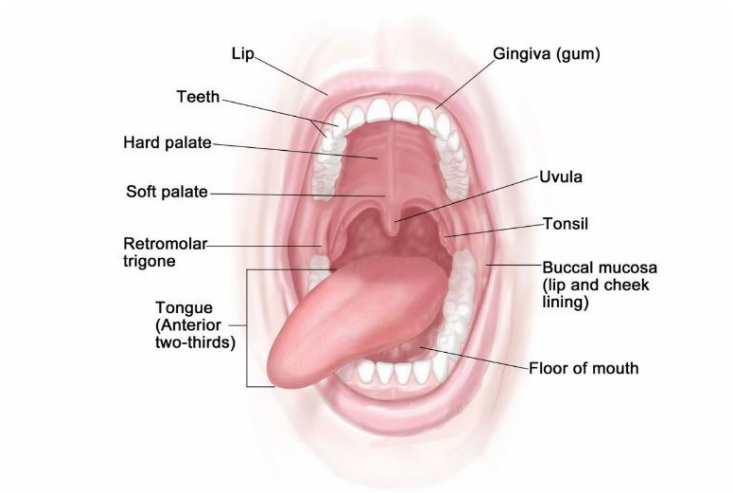


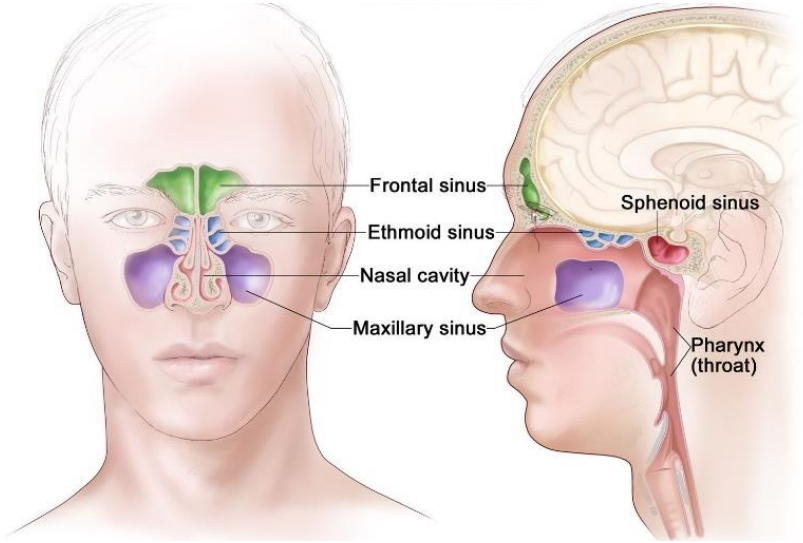

Figure 1. Illustrates head and neck cancer regions (Credit: ©Terese Winslow)

Oral cavity includes the lips, front two thirds of the tongue, the buccal mucosa, hard palate, gingiva (gums), floor of the mouth and retromolar trigone [Figure 2(A)]. The paranasal sinuses are air-filled hollow spaces surrounding the nose which includes frontal sinuses, maxillary sinuses, ethmoid sinuses and sphenoid sinuses. The nasal cavity covers the area above the roof of the mouth and connects behind the nose [Figure 2(B)]. Nasopharynx is located in the upper part of the pharynx, just behind the nose (Figure 1). The anterior wall is bordered by the nasal cavity's posterior choanae. The mucosa covering the superior pharyngeal constrictor encloses the posterior wall. The tori tubarius, eustachian tube orifices, and Rosenmuller fossae defined the lateral wall. It is restricted superiorly by the sphenoid, and inferiorly by the soft palate.

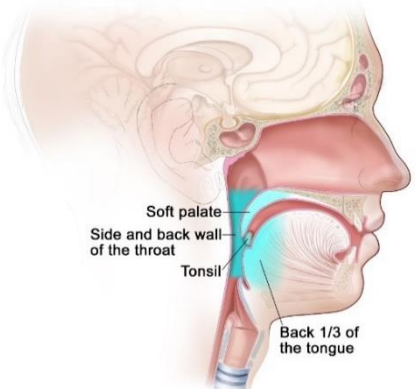
The middle part of the pharynx is the oropharynx. It includes the soft palate, tonsils, one-third of the tongue, the lateral and posterior walls of the throat [Figure 2(C)]. The hypopharynx is located at the bottom of the pharynx, covering mainly three regions- the posterior wall (back of the hypopharynx), pyriform sinuses/fossa (paired recesses located on the side of the wall) and the postcricoid region (located behind the cricoid cartilage) [Figure 2 (D)]. Adjacent to the hypopharynx is the larynx, commonly known as the voice box. It has three main parts – the upper part called the supraglottis (above the vocal cords), the middle part called the glottis and the lower part known as the subglottis (located between the trachea and the vocal cords) [Figure 2 (E)]. Table 1 shows the different subsites of squamous cell carcinoma of head and neck regions with ICD-10 codes (World Health Organization, 2019).



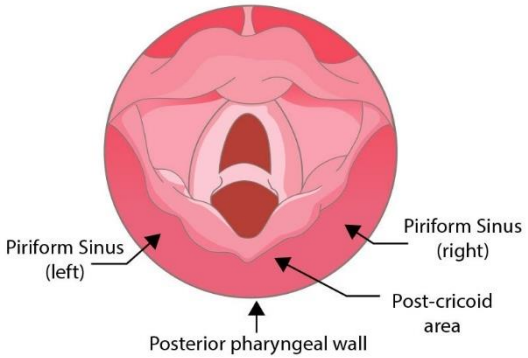

2 (A)



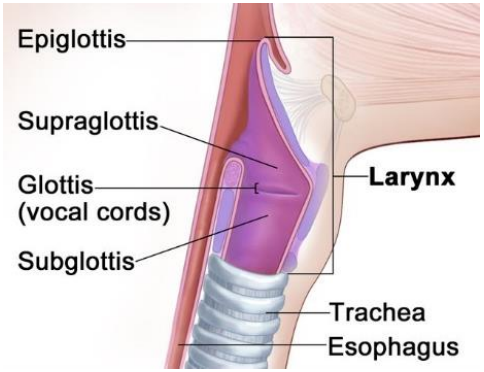

2 (B)




2 (C)




2 (D)




2 (E)

Figure 2. Illustrates anatomy of (A) Oral cavity (B) Nasal cavity and paranasal sinuses (C) Oropharynx (D) Hypopharynx (E) Larynx (Credit: ©Terese Winslow)

Table 1. Different anatomical sites of HNSCC with ICD-10 codes included in the study

| ICD-10 | | Site |
|---|---|---|
| C00 | | Malignant neoplasm of lip |
| | C00.0 | External upper lip |
| | C00.1 | External lower lip |
| | C00.2 | External lip, unspecified |
| | C00.3 | Internal upper lip |
| | C00.4 | Internal lower lip |
| | C00.5 | Internal lip, unspecified |
| | C00.6 | Commissure of lip |
| | C00.8 | Overlapping lesion of lip |
| | C00.9 | Lip, unspecified |
| C01 | | Malignant neoplasm of base of tongue |
| C02 | | Malignant neoplasm of other and unspecified parts of tongue |
| | C02.0 | Dorsal surface of tongue |
| | C02.1 | Border of tongue |
| | C02.2 | Ventral surface of tongue |
| | C02.3 | Anterior two-thirds of tongue |
| | C02.8 | Overlapping lesion of tongue |
| | C02.9 | Tongue, unspecified |
| C03 | | Malignant neoplasm of gum |
| | C03.0 | Upper gum |
| | C03.1 | Lower gum |
| | C03.9 | Gum, unspecified |
| C04 | | Malignant neoplasm of floor of mouth |
| | C04.0 | Anterior floor of mouth |
| | C04.1 | Lateral floor of mouth |
| | C04.8 | Overlapping lesion of floor of mouth |
| | C04.9 | Floor of mouth, unspecified |
| C05 | | Malignant neoplasm of palate |
| | C05.0 | Hard palate |
| | C05.1 | Soft palate |
| | C05.2 | Uvula |
| | C05.8 | Overlapping lesion of palate |
| | C05.9 | Palate, unspecified |
| C06 | | Malignant neoplasm of other and unspecified parts of mouth |
| | C06.0 | Cheek mucosa |
| | C06.1 | Vestibule of mouth |
| | C06.2 | Retromolar area |
| | C06.8 | Overlapping lesion of other and unspecified parts of mouth |
| | C06.9 | Mouth, unspecified |
| C09 | | Malignant neoplasm of tonsil |
| | C09.0 | Tonsillar fossa |
| | C09.1 | Tonsillar pillar (anterior & posterior) |

| | | |
|---|---|---|
| | C09.8 | Overlapping lesion of tonsil |
| | C09.9 | Tonsil, unspecified |
| C10 | | Malignant neoplasm of oropharynx |
| | C10.0 | Vallecula |
| | C10.1 | Anterior surface of epiglottis |
| | C10.2 | Lateral wall of oropharynx |
| | C10.3 | Posterior wall of oropharynx |
| | C10.8 | Overlapping lesion of oropharynx |
| | C10.9 | Oropharynx, unspecified |
| C11 | | Malignant neoplasm of nasopharynx |
| | C11.0 | Superior wall of nasopharynx |
| | C11.1 | Posterior wall of nasopharynx |
| | C11.2 | Lateral wall of nasopharynx |
| | C11.3 | Anterior wall of nasopharynx |
| | C11.8 | Overlapping lesion of nasopharynx |
| | C11.9 | Nasopharynx, unspecified |
| C12 | | Malignant neoplasm of pyriform sinus |
| C13 | | Malignant neoplasm of hypopharynx |
| | C13.0 | Postcricoid region |
| | C13.1 | Aryepiglottic fold, hypopharyngeal aspect |
| | C13.2 | Posterior wall of hypopharynx |
| | C13.8 | Overlapping lesion of hypopharynx |
| | C13.9 | Hypopharynx, unspecified |
| C14 | | Malignant neoplasm of other and ill-defined sites in the lip, oral cavity and pharynx |
| | C14.0 | Pharynx, unspecified |
| | C14.8 | Overlapping lesion of lip, oral cavity and pharynx |
| C30 | | Malignant neoplasm of nasal cavity and middle ear |
| | C30.0 | Nasal cavity |
| C31 | | Malignant neoplasm of accessory sinuses |
| | C31.0 | Maxillary sinus |
| | C31.1 | Ethmoidal sinus |
| | C31.2 | Frontal sinus |
| | C31.3 | Sphenoid sinus |
| | C31.8 | Overlapping lesion of accessory sinuses |
| | C31.9 | Accessory sinus, unspecified |

**Statistics of Head and Neck Squamous Cell Carcinoma**

HNSCC accounts for more than 0.89 million cases and 0.45 million deaths worldwide (Bray et al., 2024). In India, there are more than 0.14 million cases and 0.13 million deaths (Bray et al., 2024). Head and neck cancers are among the leading cancer sites for both men and women in Mizoram, Northeast India (ICMR NCDIR, 2020). According to Population Based Cancer Registries, NCDIR, Mizoram (Aizawl District)

ranks 5th in male (AAR of 45.6 per 100,000) and 8th in female (AAR of 22.7 per 100,000) among Head and Neck Cancer in India (ICMR NCDIR, 2020). Over an 18-year period from 2003 to 2020, Head and Neck Cancer was the second most common cancer among men in Mizoram after stomach cancer, with an overall Age Standardized Incidence Rate (ASIR) of 31.6 per 100,000. Among women, it ranked as the 6th most common cancer, with an ASIR of 9.7 per 100,000. Regarding mortality, Head and Neck Cancer was the 4th leading cause of cancer deaths among men, with an overall Age Standardized Mortality Rate (ASMR) of 15.9 per 100,000. Among women, it was the 6th leading cause of cancer deaths with an ASMR of 4 per 100,000 (Zomawia et al., 2023). The trend in ASIR of Head and Neck Cancer in men increases significantly over the past 18 years with an Annual Percent Change (APC) of 1.1% and in women with an APC of 0.3%. Likewise, the ASMR increases significantly in men with APC of 5.8% and 5.4% in women (Zomawia et al., 2023).

**Head and Neck Cancer staging**

Head and neck cancer was staged using the TNM (Tumour, Node, Metastasis) system established by the American Joint Committee on Cancer (AJCC 8th edition) (Tables 2 & 3) (Amin et al., 2017). The TNM staging system classifies the anatomical extent of cancer using three attributes: tumour (T), lymph node' (N), and metastasis (M). 'T' describes the size and extent of the primary tumour, 'N' indicates the absence or presence and extent of regional lymph nodes and 'M' reflects the absence or presence of distant metastasis (Huang et al., 2017). The staging system is essential for providing the best possible estimate of disease extent, which is crucial for selecting appropriate treatments, predicting outcomes, designing research, and guiding cancer management. The lymph nodes in the head and neck are subdivided into subsites and grouped into seven levels – Submental (sublevel IA), Submandibular (sublevel IB), Upper Jugular (sublevels IIA and IIB), Middle Jugular (level III), Lower Jugular (level IV), Posterior Triangle (sublevels VA and VA), Anterior Compartment (Level VI) and Superior Mediastinal (level VII). Extra-nodal extension (ENE) is used to denote the extension of metastatic tumour within and beyond the confines of the lymph node, designated as ENE (+) or ENE (-). The most common sites of distant metastasis are lungs and bones and very less often brain and hepatic metastasis.

Table 2. TNM Classification for each subsite of HNSCC

| | **Oral Cavity** | **Hypopharynx** | **Oropharynx** | **Larynx** | | | **Salivary Gland** | **Nasal Cavity and Paranasal Sinuses** | | **Nasopharynx** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | **Supraglottis** | **Glottis** | **Subglottis** | | **Maxillary sinus** | **Nasal Cavity and Ethmoid Sinus** | |
| **Tx** | Primary tumour cannot be assessed | | | | | | | | | |
| **T0** | | | | | | | No evidence of primary tumour | | | No tumour identified but EBV positive |
| **Tis** | Carcinoma *in situ* | | | | | | | | | |
| **T1** | Tumour ≤ 2 cm or ≤ 5 mm DOI (DOI is Depth of Invasion and not tumour thickness | Tumour is limited to one subsite of hypopharynx and/or ≤ 2 cm | Tumour ≤ 2 cm | Tumour limited to one subsite of supraglottis with normal vocal cord mobility | Tumour limited to the vocal cord(s) (may involve anterior or posterior commissure) with normal mobility | Tumour limited to the subglottis | Tumour ≤ 2 cm without extra parenchymal extension (evidence of invasion of soft tissues) | Tumour limited to maxillary sinus mucosa with no erosion or destruction of bone | Tumour restricted to any one subsite, with or without bony invasion | Tumour confined to nasopharynx, or extension to oropharynx and/or nasal cavity without parapharyngeal involvement |
| **T1a** | | | | | Tumour limited to one vocal cord | | | | | |
| **T1b** | | | | | Tumour involves | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | both vocal cords | | | | | |
| **T2** | Tumour ≤ 2 cm, DOI > 5 mm and ≤ 10 mm or Tumour > 2 cm but ≤ 4 cm and ≤ 10 mm DOI | Tumour invades more than one subsite of hypopharynx or an adjacent site, or > 2 cm but < 4 cm without fixation of hemilarynx | Tumour > 2 cm but ≤ 4 cm | Tumour invades mucosa of more than one adjacent subsite of supraglottis or glottis or region outside the supraglottis (e.g., mucosa of base of tongue, vallecula, medial wall of pyriform sinus) without fixation of the larynx | Tumour extends to supraglottis and/or subglottis, and/or with impaired vocal cord mobility | Tumour extends to vocal cord(s) with normal or impaired mobility | Tumour > 2 cm but ≤ 4 cm without extra parenchymal extension | Tumour causing bone erosión or destruction including extensión into the hard palate and/or middle nasal meatus, except extension to posterior wall of m axillary sinus and pterygoid plates | Tumour invading two subsites in a single región or extending to involve an adjacent region within the nasoethmoidal complex, with or without bony invasion | Tumour with extension to parapharyngeal space, and/or adjacent soft tissue involvement (medial pterygoid, lateral pterygoid, prevertebral muscles) |
| **T3** | Tumour > 4 cm or any Tumour> | Tumour > 4 cm or with fixation of hemilarynx or | Tumour > 4 cm or extension to lingual | Tumour limited to larynx with vocal cord fixation | Tumour limited to the larynx with vocal cord fixation | Tumour limited to larynx with vocal cord fixation | Tumour > 4 cm or tumour having extra | Tumour invades any of the following: bone of the | Tumour extends to invade the medial wall or floor of | Tumour with infiltration of bony structures at skull base, cervical |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 10 mm DOI | extension to oesophagus | surface of epiglottis | and/or invades any of the following: postcricoid area, preepiglottic space, paraglottic space, and/or inner cortex of thyroid cartilage | and/or invasion of paraglottic space and/or inner cortex of the thyroid cartilage | and/or invasion of paraglottic space and/or inner cortex of the thyroid cartilage | parenchymal extension | posterior wall of maxillary sinus, subcutaneous tissues. floor or medial wall of orbit, pterygoid fossa, ethmoid sinuses | the orbit, maxillary sinus, palate, or cribriform píate | vertebra, pterygoid structures, and/or paranasal sinuses |
| **T4** | Moderately advanced or very advanced local disease | | | | | | | | | Tumour with intracranial extension, involvement of cranial nerves, hypopharynx, orbit, parotid gland, and/ or extensive soft tissue infiltration beyond the lateral surface of the lateral |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | pterygoid muscle |
| **T4a** | Tumour invades adjacent structures only (e.g., through cortical bone of the mandible or maxilla, or involves the maxillary sinus or skin of the face) | Tumour invades thyroid/cricoid cartilage, hyoid bone, thyroid gland, or central compartment soft tissue | Tumour invades the larynx, extrinsic muscle of tongue, medial pterygoid, hard palate, or mandible | Tumour invades through the outer cortex of the thyroid cartilage and/or invades tissues beyond the larynx (e.g., trachea, soft tissues of neck including deep extrinsic muscle of the tongue, strap muscles, thyroid, or oesophagus) | Tumour invades through the outer cortex of the thyroid cartilage and/or invades tissues beyond the larynx (e.g., trachea, cricoid cartilage, soft tissues of neck including deep extrinsic muscle of the tongue, strap muscles, thyroid, or oesophagus) | Tumour invades cricoid or thyroid cartilage and/or invades tissues beyond the larynx (e.g., trachea, soft tissues of neck including deep extrinsic muscle of the tongue, strap muscles, thyroid, or oesophagus) | Tumour invades skin, mandible, ear canal, and/or facial nerve | Tumour invades anterior orbital contents, skin of cheek, pterygoid plates, infratemporal fossa, cribriform píate, sphenoid or frontal sinuses | Tumour invades any of the following: anterior orbital contents, skin of nose or cheek, minimal extension to anterior cranial fossa, pterygoid plates, sphenoid or frontal sinuses | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **T4b** | Tumour invades masticator space, pterygoid plates, or skull base and/or encases the internal carotid artery | Tumour invades prevertebral fascia, encases carotid artery, or involves mediastinal structures | Tumour invades lateral pterygoid muscle, pterygoid plates, lateral nasopharynx, or skull base or encases carotid artery | Tumour invades prevertebral space, encases carotid artery, or invades mediastinal structures | Tumour invades skull base and/or pterygoid plates and/or encases carotid artery | Tumour invades any of the following: orbital apex, dura, brain, middle cranial fossa, cranial nerves other than maxillary division of trigeminal nerve (V2), nasopharynx, or clivus | Tumour invades any o f the following: orbital apex, dura, brain, middle cranial fossa, cranial nerves other than (V2), nasopharynx, or clivus | |
| **NX** | Regional lymph nodes cannot be assessed | | | | | | | |
| **N0** | No regional lymph node metastasis | | | | | | | |
| **N1** | Metastasis in a single ipsilateral node ≤ 3 cm and ENE (-) | | | | | | | Unilateral metastasis in cervical lymph node(s) and/ or unilateral or bilateral metastasis in retropharyngeal |

| | | |
|---|---|---|
| | | lymph node(s), ≤ 6 cm, above the caudal border of cricoid cartilage. |
| **N2** | Metastasis in a single ipsilateral node > 3 cm but < 6 cm and ENE (-) | Bilateral metastasis in cervical lymph node(s), ≤ 6 cm, above the caudal border of cricoid cartilage |
| **N2a** | Metastasis in a single ipsilateral node > 3 cm but < 6 cm and ENE (-) | |
| **N2b** | Metastasis in a multiple ipsilateral nodes < 6 cm and ENE (-) | |
| **N2c** | Metastasis in bilateral or contralateral lymph nodes, < 6 cm and ENE (-) | |
| **N3** | Metastasis in a lymph node > 6 cm and ENE (-) | Unilateral or bilateral metastasis in cervical lymph node(s), > 6 cm and/ or extension below the caudal border of cricoid cartilage |

| | | |
|---|---|---|
| **N3a** | Metastasis in a lymph node > 6 cm and ENE (-) | |
| **N3b** | Metastasis in any node(s) and ENE (+) | |
| **M0** | No distant metastasis | |
| **M1** | Distant metastasis | |

T – Tumour, N – Nodes, M – Metastasis, ENE - Extra-nodal extension, EBV – Epstein Barr Virus

Table 3. Stages corresponding to the TNM classification for each subsite of HNSCC

| | **Oral cavity** | | | **Hypopharynx & Oropharynx, Larynx, Salivary Glands, Nasal Cavity and Paranasal Sinuses** | | | **Nasopharynx** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Stage** | **T** | **N** | **M** | **T** | **N** | **M** | **T** | **N** | **M** |
| **0** | | | | Tis | N0 | M0 | Tis | N0 | M0 |
| **I** | T1 | N0 | M0 | T1 | N0 | M0 | T1 | N0 | M0 |
| **II** | T2 | N0 | M0 | T2 | N0 | M0 | T1, T0 | N1 | M0 |
| | | | | | | | T2 | N0 | M0 |
| | | | | | | | T2 | N1 | M0 |
| **III** | T3 | N0 | M0 | T3 | N0 | M0 | T1, T0 | N2 | M0 |
| | T1, T2, T3 | N1 | M0 | T1, T2, T3 | N1 | M0 | T2 | N2 | M0 |
| | | | | | | | T3 | N0 | M0 |
| | | | | | | | T3 | N1 | M0 |
| | | | | | | | T3 | N2 | M0 |
| **IVA** | T4a | N0, N1 | M0 | T4a | N0, N1 | M0 | T4 | N0 | M0 |
| | T1, T2, T3, T4a | N2 | M0 | T1, T2, T3, T4a | N2 | M0 | T4 | N1 | M0 |
| | | | | | | | T4 | N2 | M0 |
| | | | | | | | Any T | N3 | M0 |
| **IVB** | Any T | N3 | M0 | Any T | N3 | M0 | Any T | Any N | M1 |
| | T4b | Any N | M0 | T4b | Any N | M0 | | | |
| **IVC** | Any T | Any N | M1 | Any T | Any N | M1 | | | |

**Rationale of the study**

HNSCC is a highly heterogenous cancer with varied clinical presentations and outcomes and is greatly influenced by the molecular landscape of each tumour site. Despite advances in treatment modalities for many cancers, early staged head and neck cancer are treated with definitive radiotherapy, while concurrent chemoradiotherapy (CCRT) is the mainstay of locally advanced head and neck cancer (Adelstein et al., 2017). The prognosis for HNSCC remains poor, with survival rates ranging from 70 to 80% for early-stage disease, below 70% for advanced-stage HNSCC, and less than 40% for metastatic cases (Barsouk et al., 2023).

Identifying biomarkers that can predict outcomes is crucial to optimize treatment strategies and improve patient survival. Resistance to chemotherapy and radiotherapy, as well as the occurrence of tumour relapses and recurrent or occult metastasis, significantly decreases the survival rates of late-stage HNSCC patients (Picon & Guddati, 2020; López-Verdín et al., 2018). Differences in lifestyle habits and demographic factors among the Mizo population may result in distinct molecular and clinical factors of HNSCC.

This study aims to highlight the epidemiology of HNSCC by investigating the association of various tumour sites established risk factors, family history of cancer (FHC) and lifestyle habits such as smoked food, zozial, tuibur, and locally made alcohol which are unique to the Mizo population. Additionally, it seeks to evaluate survival outcomes in HNSCC patients within Mizoram, providing insights into the treatment modalities used and analysing their two-year outcomes. The study also aims to identify variables that may influence overall survival (OS) and progression-free survival (PFS) in patients with HNSCC. Furthermore, this study aims to quantify the differentially expressed proteins in serum that my serve as prognostic or predictive biomarkers that can predict the treatment response and outcome of the patients. Lastly, the study aims to characterize the molecular profile of hypopharyngeal cancer in the population.

## Review of Literature

Consumption of alcohol, smoked food, tobacco and betelnut chewing has been a custom among Mizo people for decades. Tobacco may be consumed by smoking, sipping (sahdah), chewable gutkha products and liquid form known as 'tuibur'. Tuibur is consumed by smoking with a pipe or placed in mouth and spit out. Along with branded cigarettes, locally made without filter cigarettes called as 'Zozial' are also available within the state. Tobacco Specific Nitrosamines (TSNAs) and Polycyclic Aromatic Hydrocarbons (PAHs) are well studied carcinogens found in cigarettes as well as smokeless tobacco products (Jethwa et al., 2017). High concentrations of N - nitrosonornicotine (NNN) and heavy metals have been observed in Tuibur (Lalrammawia et al., 2022). Smoking of meat and vegetables is a traditional method of preservation which is still in practice till date. Consumption of smoked food are also a source of Polycyclic Aromatic Hydrocarbons (PAHs) (Gomaa et al., 1993).

Areca nuts, also known as betel nuts, are the fruits of a tropical palm tree called Areca catechu. These nuts are widely consumed by wrapping in Areca leaf (Piper areca) along with slaked lime, a preparation known as Kuhva. Areca nuts contain tannins, polyphenols, and alkaloids which are carcinogenic due to their ability to generate reactive oxygen species in the presence of slaked lime (Sharan et al., 2012). HNSCC is strongly associated with the use of tobacco, areca nuts, and alcohol abuse (Jethwa et al., 2017; Alsahafi et al., 2019). Acetaldehyde, a metabolic byproduct of alcohol, is highly reactive and a potent carcinogen forming DNA adducts and causing alterations in DNA methylation (Brooks et al., 2014). Alcohol is responsible for 21.6% of laryngeal cancers, 30.5% of pharyngeal cancers and 26.4% of lip and oral cavity cancers (Marziliano et al., 2020).

As a heterogeneous cancer, each subsite differs in terms of risk behaviours, disease presentations, population-wide prevalence, and treatment approaches (Johnson et al., 2023). According to the National Comprehensive Cancer Network (NCCN) guidelines, early-stage patients are treated with a single-modality approach, either surgery (S) or radiotherapy (RT). In contrast, patients with advanced-stage disease, a multi-modality approach i.e. combination of chemotherapy (CT), RT, and surgery (S) is recommended. The chemotherapy (CT) regimen typically includes cisplatin, 5-

fluorouracil, docetaxel, paclitaxel, and/or carboplatin which can be administered either alone or in combination (Adelstein et al., 2017). While induction chemotherapy (IC) and concurrent chemoradiotherapy (CCRT) have been shown to improve response rates, no statistically significant differences have been observed in OS (Zhang et al., 2015; Lee et al., 2020). The most effective combination for IC has not yet been established, despite RT and CCRT being the primary treatment modalities for HNSCC (Adelstein et al., 2017; Lee et al., 2020). Low survival rates in HNSCC have been linked to cigarette smoking, betel nut chewing, and advanced T and N staging (Lee et al., 2020; Irawan et al., 2022; Su et al., 2016). Additionally, a study by Milrud and colleagues found that HNSCC is associated with elevated leukocyte and neutrophil counts, which are linked to patient survival (Millrud et al., 2012). Numerous studies have also linked leukocytosis and neutrophilia to the outcomes of HNSCC after evaluating various therapeutic strategies (Schernberg et al., 2018; Chen et al., 2014; Roh et al., 2019; Gouw et al., 2018).

According to National Cancer Institute, a biomarker is defined as "A biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease. A biomarker may be used to see how well the body responds to a treatment for a disease or condition." (http://www.cancer.gov). Blood based tumour biomarkers, substances produced either by tumours or as part of the body's response to tumour development and progression, have proven to be highly valuable in cancer screening and early diagnosis, for predicting prognosis, detecting recurrence and monitoring the effectiveness of treatments (Zhou et al., 2024). Over the decades, exploration of serum protein biomarkers has advanced the therapeutic strategies in cancer treatment. Table 4 summarized the candidate tumour protein biomarkers that have been found in serum samples for head and neck cancer using various approaches.

Table 4. Serum based candidate tumour protein biomarkers

| Protein Name | Highlights | Treatment | Methods | Significance | Reference |
|---|---|---|---|---|---|
| Interleukin - 2 | Low IL-2 showed better survival rate | | ELISA | | Gross et al., 2016 |
| Interleukin - 8 | Increase level in metastatic and local regional recurrence | Radiotherapy | ELISA | t-test, p-value <0.01 | Gokhale et al., 2005 |
| | Increased level post treatment had poorer survival | CCRT | Luminex multiplex assay | Cox proportional hazards models (RR, 1.6; 95% CI, 1.2-2.2; $P = 0.05$) | Allen et al., 2007 |
| Interleukin - 6 | Increased level post treatment had poorer survival | CCRT | Luminex multiplex assay | Cox proportional hazards models (RR, 3.8 (2.0-7.4), P=0.004) | Allen et al., 2007 |
| | Higher pre-treated serum IL-6 levels had worse 5-year overall survival and Disease Specific Survival | CCRT | ELISA | Log-rank test (p-value<0.05) | Chang et al., 2013 |
| | Higher level associated with shorter OS and PFS | | ELISA | Cox Proportional Hazard (OS, HR=1.953, P=0.040 PFS, HR=1.885, P=0.049 | Hao et al., 2013 |
| | High pretreated level predicts worse OS and DFS | CCRT | ELISA | Cox Proportional Hazard (OS, HR=7.61 (1.82–31.80), P=0.005, DFS, HR= 3.39 (1.22–9.39), P=0.02 | De Schutter et al., 2005 |
| | Higher IL6 was associated with higher second primary cancer | | Chemiluminescent immunometric assays | Cox Proportional Hazard (p-value<0.05) | Meyer et al., 2010 |

| | | | | | |
|---|---|---|---|---|---|
| ARG1, CCL4, CCL17, CCL19, CCL20, CCL23, CXCL5, CXCL13, CASP-8, CD5, CD244, FASL, IL6, IL7, IL10, IL13, LAMP3, MMP12, MCP-4, TNF, TNFSF14, | Differentially expressed proteins after treatment when compared to baseline expression (prior treatment) | CCRT | Multiplex immunoassay analysis | Two-way ANOVA (p-value<0.05) | Astradsson et al., 2022 |
| CDK4 | Elevated expression among non-survivors compared to survivors | | Western Blot | t-test, p-value <0.001 | Banerjee et al., 2017 |
| C - reactive protein | Low CRP showed better survival than high CRP, cut off 11.3mg/L | Surgery/CCRT/RT | Routine Test | Cox Proportional Hazard (p-value < 0.0001) | Zhang et al., 2022 |
| C-reactive protein (CRP) and Tumour Necrosis Factor-α (TNF-α) | Low levels of CRP & TNF-α corresponds to high Survival Rate | | ELISA | Cox Proportional Hazard (p-value<0.05) | Andersson et al., 2014 |
| midkine | Higher expression of midkine = poor outcome | Chemotherapy | ELISA | RR = 3.77, p-value = 0.027 | Yamashita et al., 2016 |

| | | | | | |
|---|---|---|---|---|---|
| VEGF-A | Serum VEGF-A was elevated in non-responders compared to responder patients | CCRT | ELISA | Mann–Whitney (p-value<0.05) | Srivastava et al., 2014 |
| TWEAK | Low TWEAK level had higher risk of recurrence of tumour | RT/CCRT | ELISA | Cox Proportional Hazard (HR=1.8, p-value = 0.001) | Terra et al., 2015 |
| CD109 | Preoperative CD109 level was significantly associated with node metastasis. Low level = low overall survival. | Surgery | ELISA | Log-rank test (p-value = 0.046) | Hagiwara et al., 2021 |
| Neutrophil extracellular traps (NETs) | G-CSF stimulates NETs producing Neutrophils correlates with the progression of the disease | | Immunofluorescent Staining | p-value < 0.01 | Decker et al., 2019 |
| Apolipoprotein A-I | Reduced level of ApoA-1 after therapy = shorter PFS | Immunotherapy | Routine test | Cox Proportional Hazard (HR, 2.27; 95% CI, 1.11–4.61; $p = 0.034$) | Xiao et al., 2023 |
| Lactate dehydrogenase | High pretreated LDH level = poorer OS, DMFS, and DFS | | Meta-analysis | HR (OS) = 1.79<br>HR (DMFS) = 1.85<br>HR (DFS) = 1.63 | Zhang et al., 2016 |

Hypopharyngeal cancer is one of the most prevalent sites of head and neck cancer in Mizoram. Among all PBCR states in India, Mizoram has the second highest increase in age-adjusted incidence rate (APC) for hypopharyngeal cancer. Hypopharyngeal cancer is one of the least characterized subtypes of HNSCC, with limited data available on its mutational profile. Data from the TCGA, based on only 9 samples, identified *TP53* as the most frequently mutated gene (5/9), followed by *BRCA2* and *MUC16*, each mutated in 3/9 samples. Among the top ten most mutated genes in the TCGA dataset, *PIK3CA*, *FAT4*, *EGFR*, *TENT5C*, *LRP1B*, *KMT2D*, and *NUMA1* were each present in 2 of the 9 samples. In a study of ten hypopharyngeal cancer tissues from the Chinese population, whole-exome sequencing (WES) identified 8,113 mutations across 5,326 genes (Yao et al., 2023). The gene *KMT2C* was mutated in all the samples, while other frequently mutated genes *MEGF8*, *ITPR1*, *DYSF*, *DNAH10*, *CUL7*, *MYH14*, *LRP1* and *ASTN1* were found in 6 out of 10 samples. In another Chinese cohort of 23 hypopharyngeal cancer tissues paired with adjacent normal samples, WES identified 15 somatic mutations in *TP53, REC8, PRB4, EI24, NSD1, CDKN2A, KLK3, ALDH2, BICD1, CDK2AP1, PIK3CA, PEG3, CNGA4, SULF1* and *LATS1* (Wu et al., 2017). Additionally, copy number variations were reported in the genes *ATF1, CDKN2A* and *CDKN2B*. In a study with 23 hypopharyngeal and 25 laryngeal cancer cases, *TP53, FAT1, NOTCH1, KMT2C* and *CDKN2A* were identified as the most frequently altered genes (Machnicki et al., 2022). Additionally, they compared data from 67 hypopharyngeal cancers with 595 HNSCC samples from other subsites using Maftools and found *CASP8* and *HRAS* to be significantly different, as these genes were rarely mutated in hypopharyngeal cancer. Genetic alterations in a selected panel of oncogenes were screened using Polymerase Chain Reaction revealing that amplification of the 11q13 region (including *CCND1, FGF3, FGF4* and *EMS1* genes) was the most frequently reported alteration, followed by mutations in *ERBB1* and *MYC* oncogenes. Additionally, loss of heterozygosity (LOH) was observed in *TP53* and *NAT2* genes (Rodrigo et al., 2002). Apart from these findings, the genetic profile of hypopharyngeal cancer remains largely underexplored.

## OBJECTIVES

1. To identify the epidemiological risk factors associated with the Head and Neck cancer and to correlate with the progression of the disease in the Mizo population.
2. To identify potential predictive and prognostic protein biomarker(s) in Head and Neck cancer.
3. To identify the genomic alterations involved in Hypopharyngeal Cancer.

# METHODOLOGY

## Ethical Clearance

Ethical Clearance for this study was obtained (No.B.12018/1/13-CH(A)/IEC/69) from Institutional Ethics Committee (IEC), Civil Hospital, Aizawl, Mizoram.

## Sample and data collection

### Epidemiological study

Biopsy diagnosed HNSCC at Civil Hospital Aizawl between 2017 and 2019 were included. The following sites in the study included oral cavity (ICD-10 codes C00.0 - C06.2), nasopharynx (ICD codes 11.0–11.9), oropharynx (ICD codes C09.0 - C10.9), hypopharynx (ICD codes C12 - C13.9) and larynx (ICD codes C32.0 - C32.2). Data were collected from 100 HNSCC patients in the form of questionnaires comprising of lifestyle habits like consumption of alcohol, smoked food, tobacco in the form of smoking (cigarette), dipping (sahdah), chewing/ingested gutkha products and FHC. Questionnaires from 200 age-matched healthy controls were also recorded.

As per the National Cancer Institute, USA, smoking was quantified in pack years (http://www.cancer.gov/). This was calculated by dividing the number of cigarettes smoked per day by 10 (since a standard pack typically contains 10 cigarettes in the region) and then multiplying by the number of years the individual has smoked. Participants/patients were then classified into three groups: non-smokers, smokers with below and above average pack years. Alcohol consumption level was measured by multiplying number of days an individual drank in a week with the duration (in years) of alcohol consumption. Participants were categorized into non-drinkers, below and above average alcohol consumption.

Family history of cancer was documented to determine if the participants/patients had any known blood-related family members with cancer at any site. Participants/patients were categorized into three classes: Family History of Cancers (FHC), First-Degree Family History of Cancer (First-Degree FHC) and Second-Degree Family History of Cancer (Second-Degree FHC). First-Degree FHC includes parents and siblings while Second-Degree FHC includes uncles/aunties,

cousins and grandparents, FHC includes those with either First or Second-Degree FHC.

**Survival Analysis**

A retrospective cohort study was conducted to perform a survival analysis on patients with HNSCC diagnosed between 2017 and 2020 at the Mizoram State Cancer Institute (MSCI) in Mizoram, Northeast India. Data were extracted from medical records at MSCI and patients were followed up for two years. Out of 850 patients diagnosed with head and neck cancer during this period, 210 were selected based on specific inclusion and exclusion criteria (Figure 3). Only HNSCC patients were included in the study. Patients having squamous cell carcinoma primarily from oral cavity, nasopharynx, oropharynx, hypopharynx or larynx were included. All the patients had M0 (Metastasis) at the time of diagnosis. Only patients belonging to Mizo tribe and residing within Mizoram were selected. Patients diagnosed in Mizoram, but not receiving treatment or not registered in the studied institute were excluded from the study. Also, patients registered in the institute but given referrals to other institutes in other states were excluded. Patients who were lost to follow up or who left before treatment initiated were also excluded from the study. Patients registered in the institute but either refused to or were unfit to receive treatment were excluded. This study was reported in accordance with the Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) guidelines for observational studies (von Elm et al., 2007).

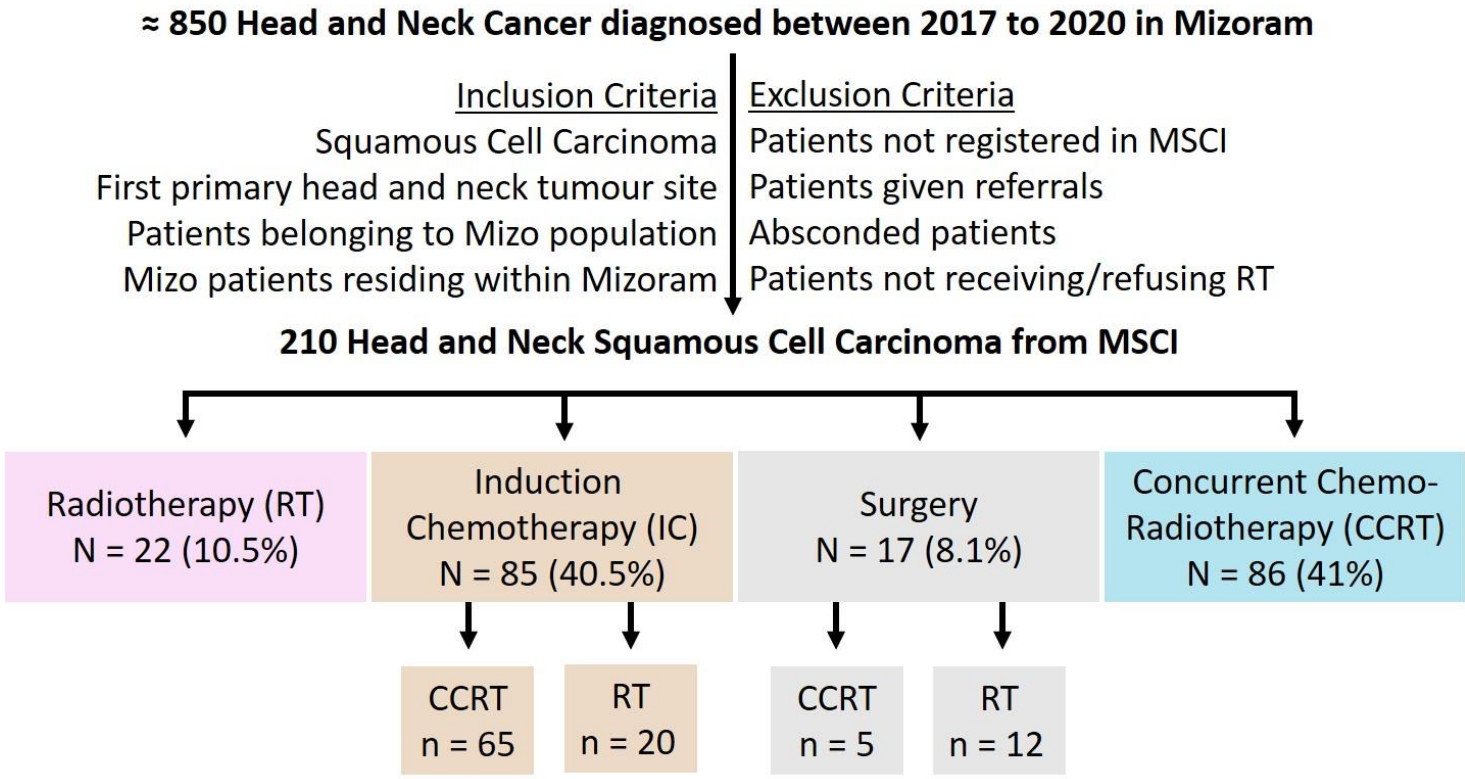



Figure 3. Flowchart for inclusion and exclusion criteria plus patient treatment distribution for survival analysis.

MSCI - Mizoram State Cancer Institute

Data extracted from medical records included clinical and demographic factors like age, sex, primary tumour site, TNM classification [tumour(T), nodes (N) and metastases (M)], total leukocyte count (TLC), absolute neutrophil count (ANC) and treatment regimen. Lifestyle factors included consumption of alcohol, betelnut chewing habits and tobacco habits (smoking/smokeless). Smokeless tobacco included snuffed tobacco (sahdah), liquified tobacco-infused water called tuibur or chewable gutkha products. Tumours were classified based on International Classification of Diseases, 10$^{th}$ Revision. TNM classification was done based on the American Joint Committee, 8$^{th}$ edition. The study cohort comprised heterogenous sites of head and neck cancer and each site had different classifications of T and N. To avoid misinterpretation of the stages for each site, the T and N classification was used independently instead of the stages defined by the TNM classification. Tumours were graded as well-differentiated, moderately differentiated, poorly differentiated or undifferentiated. There was no recorded information available for Human Papillomavirus and Epstein Barr virus. According to the treatment plan received, patients were grouped into four categories: i) Induction chemotherapy plus concurrent chemoradiotherapy/radiotherapy, (ii) concurrent chemoradiotherapy, (iii) radiotherapy only and (iv) surgery plus adjuvant concurrent chemoradiotherapy/radiotherapy. For routine evaluation, Computer Tomography scan was used before treatment and for follow up to determine tumour progression.

The patients who received RT were followed up 45 days after treatment, with subsequent follow-up every 6 months for 2 years. A CT scan was done at each follow up visits. The treatment response was evaluated using the RECIST v1.1 criteria (Response Evaluation Criteria in Solid Tumours) (Eisenhauer et al., 2009). Patients were categorized into four groups based on their response: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). Overall Survival (OS) was defined as the period (in months) from the start of treatment to death from any cause. PFS was defined as the time elapsed from initiation of treatment to either PD, SD or death. Initiation of treatment referred to CT, RT or Surgery, depending on which was administered first. Leukocytosis was defined as the TLC greater than

10,000 cells per cubic millimeter (thou/cumm) and Neutrophilia was defined as the ANC exceeding 7000 cells per cubic millimter (thou/cumm).

**Statistical Analysis**

Descriptive analysis was performed with gender, age, lifestyle habits including smoking (cigarette), dipping (sahdah), tobacco infused water (tuibur), chewing of areca nut (kuhva), consumption of smoked food and family history of cancer. Logistic Regression Analysis was done to calculate the adjusted Odds Ratio (OR) with 95% Confidence Interval (CI) to determine the risk associated with the factors on HNSCC cases against the controls. P-value greater than 0.05 was considered significant. The significant OR observed in univariate were treated as confounding factors in the multivariate analysis. The multivariate regression model was adjusted for smoking (cigarette), alcohol, areca nut (kuhva) and family history of cancer.

Frequencies for categorical variables and median values for numerical variables were generated. Univariate and multivariate analyses were done for OS and PFS using Statistical package for Social Sciences (SPSS). Univariate analysis was done for OS and PFS against each demographic, lifestyle, clinical factors and treatment administered. Unknown data for a variable were coded as missing data. The observed significant variables were considered as covariates in the subsequent multivariate analysis. Multicollinearity test was done on the predictors to ensure the accuracy and reliability of the multivariate models. To assess the multicollinearity among the predictors, a threshold of variance inflation factor (VIF) cut-off point 2 was used. Survival analysis plots using Kaplan-Meier Method and log-rank test were generated using R Studio. A p-value of $<0.05$ was considered statistically significant.

**Sampling for Proteome analysis**

Serum samples were collected from 20 patients, before treatment and two weeks (14 days) of treatment from Mizoram State Cancer Institute (MSC), Zemabawk. Samples were stored in cryovials with protease inhibitor cocktail in -20$^{o}$C.

Acetonitrile Precipitation - Depletion of highly abundant proteins (Das et al., 2020)

Chilled acetonitrile (ACN) was added to the serum in a 1:1 volume ratio. The samples were thoroughly mixed by tapping and kept on ice for 1 hour, with tapping every 15 minutes. They were centrifuged at 13,000 rpm for 1 hour at 4°C. The samples were dried using a SpeedVac for 1 hour and 30 minutes and resuspended with 80 µl of autoclaved Millipore water before being stored at -20°C. The ACN pellet was resuspended in 100 µl of water and pipetted vigorously.

**SDS - Polyacrylamide Gel Electrophoresis**

SDS-PAGE was run to check the depletion of highly abundant proteins (Figure 4). The resolving gel and stacking gel were prepared as per the Table 5. Protein markers, reagents such as gel loading buffers, running buffers, solutions for gel fixing, gel washing, solutions for staining, destaining and samples were also prepared according to the Table 6 - 9. The ACN pellet and enriched serum were heated at 100°C for 10 minutes before loading. SDS-PAGE gel was run at 130 V for 1 hour and 30 minutes. The gels were soaked in Gel Fixing Solution for 1 hour, washed with Gel Washing Solutions for 1 hour and stained in Staining Solution for 2 hours or until the edges of the gels turned blue. Destaining solutions were added and incubated for 1 hour, followed by overnight storage in Millipore water.

Table 5. SDS-PAGE gel preparation

| **15% Resolving Gel (8 ml)** | | **5% Stacking Gel (2 ml)** | |
|---|---|---|---|
| **Components** | **Quantity** | **Components** | **Quantity** |
| Water | 1.84 ml | Water | 1.4 ml |
| 30 % Acrylamide | 4 ml | 30% Acrylamide | 330 µl |
| 1.5 M Tris (pH 8.8) | 2 ml | 1M Tris (pH 6.8) | 250 µl |
| 10% SDS | 80 µl | 10% SDS | 20 µl |
| 10% APS | 80 µl | 10% APS | 20 µl |
| TEMED | 5 µl | TEMED | 4 µl |

Table 6. Buffers for SDS PAGE preparation

| **Gel loading buffer (2X) 10 ml** | | **Running buffer (pH 8.3) (1L)** | |
|---|---|---|---|
| **Components** | **Quantity** | **Components** | **Quantity** |
| Water | 3.55 ml | Tris base | 60.4 g |
| Tris HCl (0.5M, pH 6.8) | 1.25 ml | Glycine | 376 g |
| Glycerol | 2.50 ml | SDS | 20 g |
| 10% SDS | 2.00 ml | Dissolved in 1L millipore water and pH checked | |
| 0.5% Bromophenol blue | 0.05 ml | | |
| ß-Mercaptoethanol | 0.5 ml | | |

Table 7. Solutions for gel fixing and gel washing

| **Gel fixing solution** | | **Gel washing solution** | |
|---|---|---|---|
| **Components** | **Quantity** | **Components** | **Quantity** |
| Water | 300 ml | Water | 300 ml |
| 95% Ethanol | 500 ml | Methanol | 500 ml |
| Acetic Acid | 100 ml | Acetic Acid | 100 ml |
| Volume adjusted to 1000 ml | | Volume adjusted to 1000 ml | |

Table 8. Solutions for staining and destaining

| **Staining Solution** | **Destaining Solution** |
|---|---|
| 0.4 g of Coomassie blue R350 in 200 mL of 40% (v/v) methanol in water. | 500 mL of HPLC- grade methanol to 300 mL water |
| The solution was filtered to remove any insoluble material | 100 mL of reagent grade acetic acid |
| 200 mL of 20% (v/v) acetic acid in water. | Volume was adjusted to 1000 mL with water. |

Table 9. Preparation of samples for loading into SDS-PAGE

| **Components** | **Quantity loaded in gel** |
|---|---|
| Protein Marker (PM) | 10µl PM + 10 µl dye |
| Crude Serum (CS) | 2 µl CS + 8 µl water + 10 µl dye |
| ACN pellet | 20 µl pellet + 10 µl dye |
| Enriched serum (ES) (10 to 30 µg) | 5 -10 µl ES + 10 µl dye |

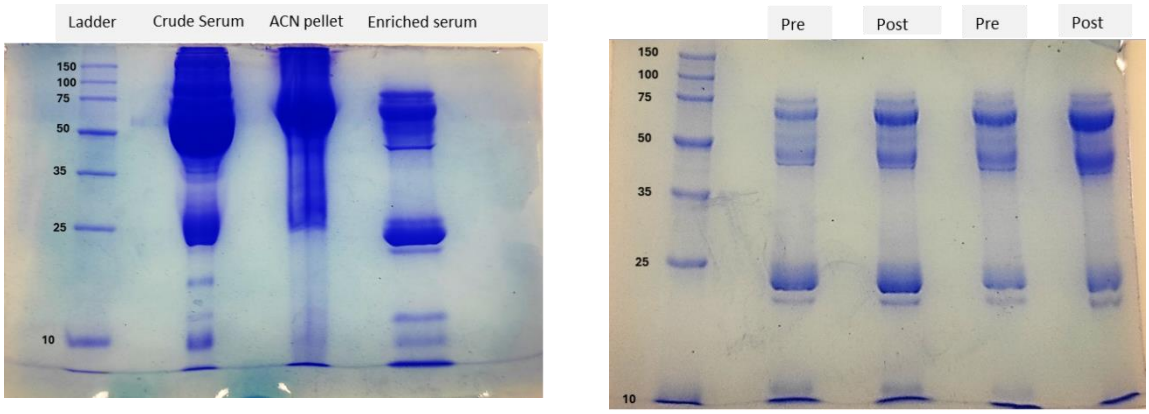



Figure 4. SDS PAGE gel showing depleted abundant proteins in pre-treated and treated samples.

**Trypsin Digestion**

Reduction and alkylation of the samples was done using 100 mM 1, 4-Dithiothreitol and 200 mM iodoacetamide respectively. Digestion was done overnight with MS grade trypsin in the ratio 1:25 (1 µg of trypsin to 25 µg of protein).

**Mass Spectrometry**

All samples were injected in duplicate (2 injections per sample) with an injection volume of 1.0 µl. The samples were loaded into the nanoACQUITY UPLC® chromatographic system (Waters, Manchester, UK) for analysis. The acquisition and analysis were performed using MassLynx 4.1 SCN781 software.

**Data processing**

Progenesis QI for Proteomics V4.2 (Non Linear Dynamics, Waters) was used for data processing. Peptides with a false positive rate of 1 were filtered out. Variable modifications, such as oxidation of methionine, were removed. Fixed modifications, such as cysteine carbamidomethylation, were also removed.

**Data analysis**

Peptides with a total number ≥2 and at least one unique peptide were selected, while abundant proteins were removed. A Volcano Plot (VolcaNoseR) was generated to identify significantly differentially expressed proteins with a fold change of 2. Proteins with a corrected p-value using the FDR approach (q-value ≤ 0.05) were

considered significant. Gene Ontology (GO) enrichment analysis and KEGG pathway analysis were performed using the DAVID Database to understand the involved pathways. A Protein-Protein Interaction (PPI) network was constructed using the STRING database. Protein clusters were identified using MCODE (Cytoscape) with a degree cutoff of 2, a K-score of 2 and a score cutoff of 0.2. Univariate Cox analysis was conducted for OS and PFS and Kaplan-Meier analysis using the log-rank test was performed.

**Sampling for Exome Analysis**

Biopsy tissue samples from treatment naïve patients along with their peripheral blood samples (2 ml in EDTA vial) were collected from 10 hypopharyngeal cancer patients at Civil Hospital Aizawl, Mizoram. All the patients were interviewed with a duly informed consent. Lifestyle factors like smoking habits, tobacco and alcohol consumption were obtained. Clinical factors including TNM staging and treatment administered were extracted from each patient's files. The patients were followed up 45 days after completion of treatment. CT scan was done to check the response to treatment. Treatment response was assessed based on RECIST v1.1 criteria into Complete response (CR), Partial response (PR), Stable disease (SD) and Progressive disease (PD).

**DNA isolation for Exome Analysis**

Genomic DNA from the tumour samples and blood samples were extracted using QIAamp® DNA Tissue Kit and QIAamp® Blood DNA mini kit, respectively. The genomic DNA was checked using Electrophoresis on a 0.8% agarose gel. Quantification of DNA was done using THERMO SCIENTIFIC μDROP PLATE on Multiskan SkyHigh Spectrophotometer. Samples were subjected to Whole Exome Sequencing at National Institute of Biomedical Genomics (NIBMG).

**Whole Exome Sequencing**

A 100 ng DNA sample was fragmented by mechanical shearing using Covaris, followed by end repair and A tailing. The DNA fragments were then ligated with unique dual indexing primer pairs with barcoded adapters, generating indexed DNA library amplicons. These were pooled and hybridized using Illumina TruSeq DNA

Exome Enrichment Reagents and coding exome probes. The probe and DNA library hybrids were captured using magnetic beads. PCR amplification was performed and the DNA was then sequenced using a Novaseq 6000 sequencer (Figure 5).

**Exome Data Analysis**

The quality of the raw FASTQ files was assessed using FastQC (Andrew, 2010). Adapter sequences and low-quality reads were subsequently trimmed with Trimmomatic software (Bolger et al., 2014). The trimmed files were rechecked using FastQC. These files were mapped to the Human Reference Genome (hg19) using BWA-MEM2 (Li, 2013). Post-alignment processing included marking duplicates with Picard (http://picard.sourceforge.net/) and sorting and indexing the reads with Samtools (Li et al., 2003). Variant calling was performed using Mutect2, and the identified variants were annotated with ANNOVAR (Figure 6).

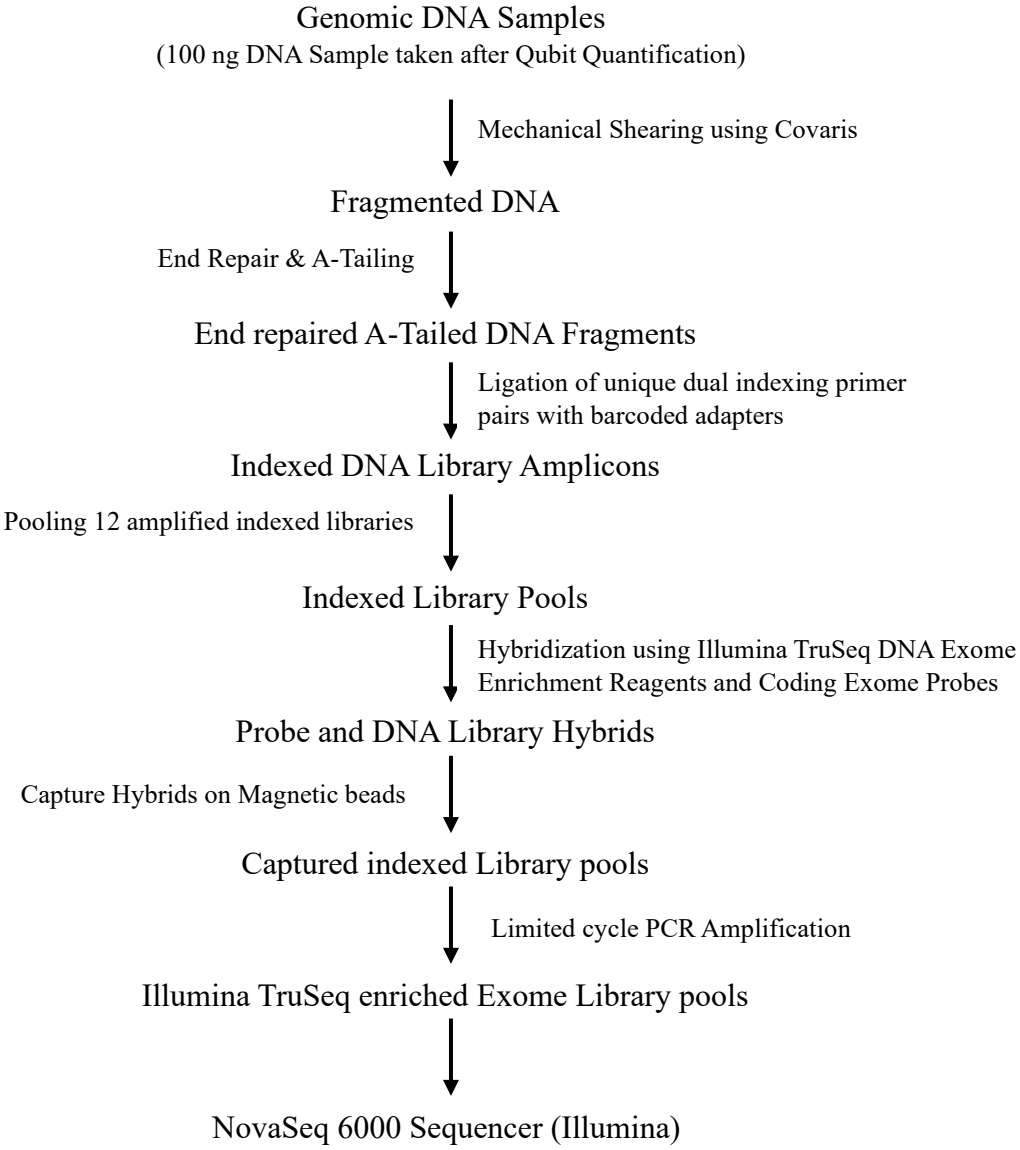



Figure 5. Flowchart representing steps in Whole Exome Sequencing

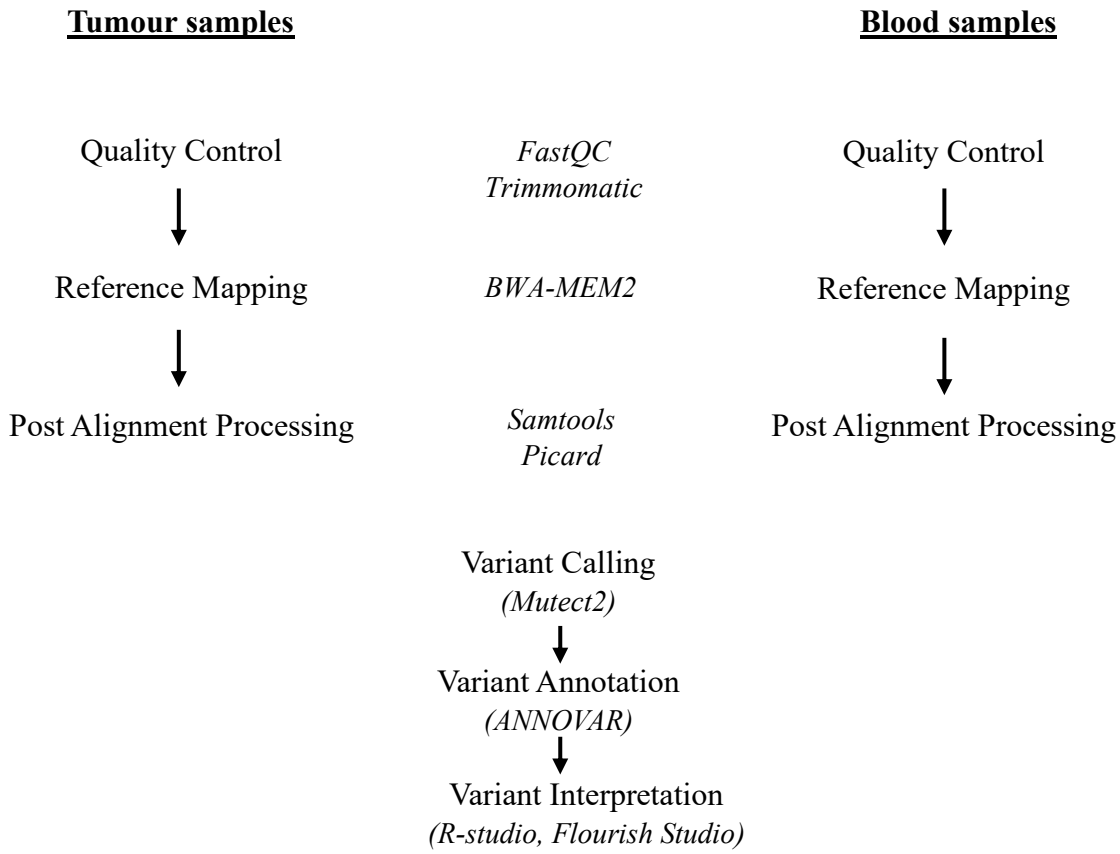



Figure 6. Flowchart representing data processing steps on Whole Exome Sequencing

## RESULTS

In our cohort, 40.7% were male, with a statistically significant odds ratio of 6.694 (Tables 10 & 11). The mean age of the participants was 54.66 years. The most commonly affected site was the oral cavity (41 patients), followed by the hypopharynx with 30 patients (Table 12). The sites with the lowest frequencies were the larynx and nasopharynx, each with 9 cases. In our study, 48.7% (146) of participants were smokers, with 85 in the case and 61 in the control groups. Among the cases, the most frequently smoked cigarette brand was Zozial (66 out of 85 smokers) (Table 12). Smoking levels were categorized based on an average of 70 pack-years. Smokers with pack-years above and below this average had significant p-values, with those above the average showing a higher risk compared to those below the average (Table 11).

The average level of alcohol consumption was 20 (Table 10). In this study, 71% of participants were non-drinkers (174 controls and 39 cases). Seventeen percent consumed alcohol below the average level, with 18 controls and 33 cases. The majority of patients (51 out of 63) consumed locally made alcohol (Table 12). Regression analysis indicated an increased risk with higher alcohol consumption, as the odds ratio for consumption above the average was 5.509, while for below-average consumption, it was 4.021.

Regression analysis indicated that the risk associated with kuhva (areca nut), dipping (sahdah), tuibur and smoked food were not statistically significant. Out of 300 participants, 105 were identified as having a First-Degree FHC, with 58 in the control group and 47 in the case group. A significant link to cancer risk was found for First-Degree FHC, with an OR of 1.92, whereas no significant association was observed for Second-Degree FHC. The duration of smoking was divided into categories: non-smokers, less than 20 years, 21 - 30 years, 31 - 40 years, 41 - 50 years, and more than 50 years. The frequency of FHC and First-Degree FHC cases was plotted against each smoking duration category. Except for those who smoked for less than 20 years, FHC accounted for over 50% of the total cases in each category, including non-smokers. Similarly, a graph was drawn for the duration of alcohol consumption, plotting FHC and First-Degree FHC across non-drinkers and those with up to 50 years of alcohol consumption, divided into ten-year intervals. Among alcohol consumers, FHC was

present in over 50% of the cases. Figures 7 & 8 shows the distribution of FHC in relation to the duration of smoking and alcohol consumption across different HNSCC sites.

The patients' ages ranged from 21 to 84 years and median age is 55 years (Table 13). Among the 210 cancer patients, the most prevalent site was the hypopharynx (67), followed by the nasopharynx (48), oral cavity (39), oropharynx (31) and larynx (25). The most frequent T classification observed in the study was T2 (86) while N1 was the most common N classification (89). In this study, 164 out of 210 patients (78.1%) smoked tobacco and 107 (51%) consumed alcohol. Most patients used alcohol, tobacco or cigarettes (Table 14). Additionally, 108 patients had leukocytosis, and 107 had neutrophilia. Some variables had missing data, such as grading, TLC, ANC, cigarette smoking, consumption of alcohol, smokeless tobacco use, betel nut chewing and FHC (Table 13).

Table 15 outlines the treatment modalities provided to the patients. The patients were divided into four categories: patients who received induction chemotherapy (IC) followed by concurrent chemoradiotherapy (CCRT) or radiotherapy (RT) (referred to as sequential chemoradiotherapy), those who received CCRT alone, those who received RT alone and those who underwent surgery followed by adjuvant CCRT or RT (Table 15). Among the 210 patients, 85 were treated with IC followed by CCRT or RT, while 86 received CCRT alone. RT alone was administered to 22 patients (10.5%), and surgery was performed on 17 oral cavity cancer patients. Of the 22 patients treated with RT only, 15 were in the early stages with no nodal involvement, while 7 had nodal involvement. Among the latter group, two received palliative RT without chemotherapy, three declined treatments and two were aged and weak for chemotherapy. The distribution of treatment modalities between tumour stages and nodal involvement is detailed in Table 16. Induction chemotherapy mostly involved the use of cisplatin or carboplatin along with paclitaxel or docetaxel. Of those receiving IC, 54 patients were treated with cisplatin plus paclitaxel. For CCRT, single agents such as cisplatin, carboplatin, or paclitaxel were used, with cisplatin being administered to 123 patients receiving CCRT. The dosages of the chemotherapeutic drugs used in the study are provided in Table 17. Palliative RT typically involved a

total dose of 30 Gy, delivered in 10 fractions. In contrast, those undergoing curative radical or adjuvant RT received doses ranging from 60 to 66 Gy in 30–33 fractions. Of the entire patient cohort, 184 individuals received radical RT, 17 received adjuvant RT, and 9 underwent palliative RT. CR was observed in 117 of the patients, PR in 8, SD in 1 and PD in 84 patients.

The 2-year OS rate for the 210 patients was 78.1% and the PFS rate was 57.4% as determined by Kaplan - Meier analysis (Table 18). Kaplan - Meier plot and Log-rank test of OS and PFS for treatment regimen, TLC, ANC and nodal involvement are shown in Figure 9 to 15. Among the different treatment approaches, the lowest OS rate was observed in patients receiving RT alone, at 70.4%. On the other hand, those receiving IC + CCRT or IC + RT had the lowest PFS rate, at 47.3% (Table 18). However, these differences were not statistically significant. Patients in the IC group who were treated with cisplatin plus 5-fluorouracil had the poorest OS and PFS [Figures 11 (a) & (b)]. In the IC group, the median PFS was 22.2 months. A statistically significant difference in PFS was found between the IC group and patients without IC [Figure 9 (b)]. However, there was no statistically significant difference in OS (Table 18). TLC of ≤ 10 thou/cumm was associated with a better survival rate compared to patients with TLC > 10 thou/cumm [Figure 12 (a)]. Similarly, patients with lower ANC had better OS and PFS rates [Figures 13 (a) & (b)]. Significant differences in survival probabilities were also observed across different N classifications ($p = 0.005$) [Figure 14 (a)]. Patients classified as N2 had the worst PFS, at 39.8%, with a median of 22.2 months. Among primary tumour locations, the oral cavity had the worst OS and PFS rates [Figures 15 (a) & (b)]. The highest OS rate was observed in patients with nasopharyngeal cancer, while hypopharyngeal cancer showed the best PFS rate.

Univariate Cox regression analysis identified T and N classifications, TLC and ANC as significant predictors of OS. Due to multicollinearity (VIF > 2) between TLC and ANC, ANC was adjusted for T and N classifications and it was excluded from the multivariate models for the other variables. Cancer site, TLC, N classification and type of treatment were identified as significant predictors of OS (Table 19). Hazard ratio (HR) indicated that laryngeal cancer was a strong predictor of poor survival (HR = 5.165). The HR increased with greater nodal involvement, with N2 classification

showing statistical significance (HR = 3.835). Leukocytosis was also a significant predictor of poor OS. N2 involvement and ANC were significant predictors for PFS in univariate analysis (Table 20). Since multicollinearity was found between ANC and TLC, TLC was adjusted for N classification only. Site (Larynx), N2 involvement, leukocytosis and neutrophilia emerged as significant predictors of PFS after adjusting for covariates. Laryngeal cancer (HR = 2.844) was a strong predictor of poor response. Similarly to OS, leukocytosis (HR = 2.035) and neutrophilia (HR = 1.946) were statistically significant predictors of PFS. Additionally, the N classification showed an increased HR with higher N involvement and was statistically significant for N2 (HR = 3.483).

Table 10. Characteristics of HNSCC cases and controls for epidemiological study

| Factors | Variables | Control n (%) | Case n (%) | Total N (%) |
|---|---|---|---|---|
| **Gender** | Female | 152 (85.4) | 26 (14.6) | 178 (59.3) |
| | Male | 48 (39.3) | 74 (60.7) | 122 (40.7) |
| **Age group (years)** | Below 45 | 58 (79.5) | 15 (20.5) | 73 (24.3) |
| | Above 45 | 142 (62.6) | 85 (37.4) | 227 (75.7) |
| **Smoking status** | No | 139 (90.3) | 15 (9.7) | 154 (51.3) |
| | Yes | 61 (41.8) | 85 (58.2) | 146 (48.7) |
| **Smoking Level** | Non-smoker | 139 (90.3) | 15 (15) | 154 (51.3) |
| | Below Average | 51 (49.5) | 52 (50.5) | 103 (34.3) |
| | Above Average | 10 (23.3) | 33 (76.7) | 43 (14.3) |
| **Alcohol status** | No | 174 (81.7) | 39 (18.3) | 213 (71.0) |
| | Yes | 26 (29.9) | 61 (70.1) | 87 (29.0) |
| **Alcohol Consumption Level** | Non-drinker | 174 (81.7) | 39 (18.3) | 213 (71.0) |
| | Below Average | 18 (35.3) | 33 (64.7) | 51 (17.0) |
| | Above Average | 8 (22.2) | 28 (77.8) | 36 (12.0) |
| **Snuff (Sahdah)** | No | 88 (61.1) | 56 (38.9) | 144 (48.0) |
| | Yes | 112 (71.8) | 44 (28.2) | 156 (52.0) |
| **Tuibur** | No | 159 (66.0) | 82 (34.0) | 241 (80.3) |
| | Yes | 41 (69.5) | 18 (30.5) | 59 (19.7) |
| **Kuhva (Areca Nut)** | No | 75 (80.6) | 18 (19.4) | 93 (31.0) |
| | Yes | 125 (60.4) | 82 (39.6) | 207 (69.0) |
| **Smoked food** | No | 43 (64.2) | 24 (35.8) | 67 (22.3) |
| | Yes | 157 (67.4) | 76 (32.6) | 233 (77.7) |
| **FHC** | Without | 118 (73.3) | 43 (26.7) | 161 (53.7) |
| | With | 82 (59.0) | 57 (41.0) | 139 (46.3) |
| **First Degree FHC** | Without | 142 (72.8) | 53 (27.2) | 195 (65.0) |
| | With | 58 (55.2) | 47 (44.8) | 105 (35.0) |
| **Second Degree FHC** | Without | 176 (67.7) | 84 (32.3) | 260 (86.7) |
| | With | 24 (60.0) | 16 (40.0) | 40 (13.3) |

FHC – Family history of Cancer

Average alcohol consumption level = 20

(no. of drinking days per week X duration of drinking years)

Average smoking level (pack years) = 70

Table 11. Regression analysis of risk factors with cases - controls

| Factors | Variables | Control (n) | Case (n) | p-value | Odds Ratio* (95% Confidence Interval) |
|---|---|---|---|---|---|
| **Gender** | Female | 152 | 26 | | Reference |
| | Male | 48 | 74 | <0.05 | 6.694 (3.278 - 13.669) |
| **Age group (years)** | Below 45 | 58 | 15 | | Reference |
| | Above 45 | 142 | 85 | <0.05 | 3.979 (1.768 - 8.955) |
| **Smoking status** | No | 139 | 15 | | Reference |
| | Yes | 61 | 85 | <0.05 | 6.703 (3.360 - 13.375) |
| **Smoking Level** | Non-smoker | 139 | 15 | | Reference |
| | Below Average | 51 | 52 | <0.05 | 4.896 (2.352 - 10.191) |
| | Above Average | 10 | 33 | <0.05 | 15.438 (5.989 - 39.793) |
| **Alcohol status** | No | 174 | 39 | | Reference |
| | Yes | 26 | 61 | <0.05 | 4.527 (2.354 - 8.706) |
| **Alcohol Consumption Level** | Non-drinker | 174 | 39 | | Reference |
| | Below Average | 18 | 33 | 0.001 | 4.021 (1.890 - 8.557) |
| | Above Average | 8 | 28 | <0.05 | 5.509 (2.189 - 13.918) |
| **Snuff (Sahdah)** | No | 88 | 56 | | Reference |
| | Yes | 112 | 44 | 0.364 | 0.753 (0.408 - 1.389) |
| **Tuibur** | No | 159 | 82 | | Reference |
| | Yes | 41 | 18 | 0.561 | 0.8 (0.377 - 1.697) |
| **Kuhva (Areca Nut)** | No | 75 | 18 | | Reference |
| | Yes | 125 | 82 | 0.590 | 1.218 (0.594 - 2.497) |
| **Smoked food** | No | 43 | 24 | | Reference |
| | Yes | 157 | 76 | 0.351 | 0.712 (0.349 - 1.454) |
| **FHC** | No | 118 | 43 | | Reference |
| | Yes | 82 | 57 | 0.948 | 1.021 (0.553 - 1.883) |
| **First Degree FHC** | No | 142 | 53 | | Reference |
| | Yes | 58 | 47 | 0.037 | 1.921 (1.040 - 3.547) |
| **Second Degree FHC** | No | 176 | 84 | | Reference |
| | Yes | 24 | 16 | 0.088 | 0.464 (0.192 - 1.122) |

FHC – Family History of Cancer

*OR adjusted with smoking, alcohol, kuhva and FHC

p-value is significant at 5% level (<0.05)

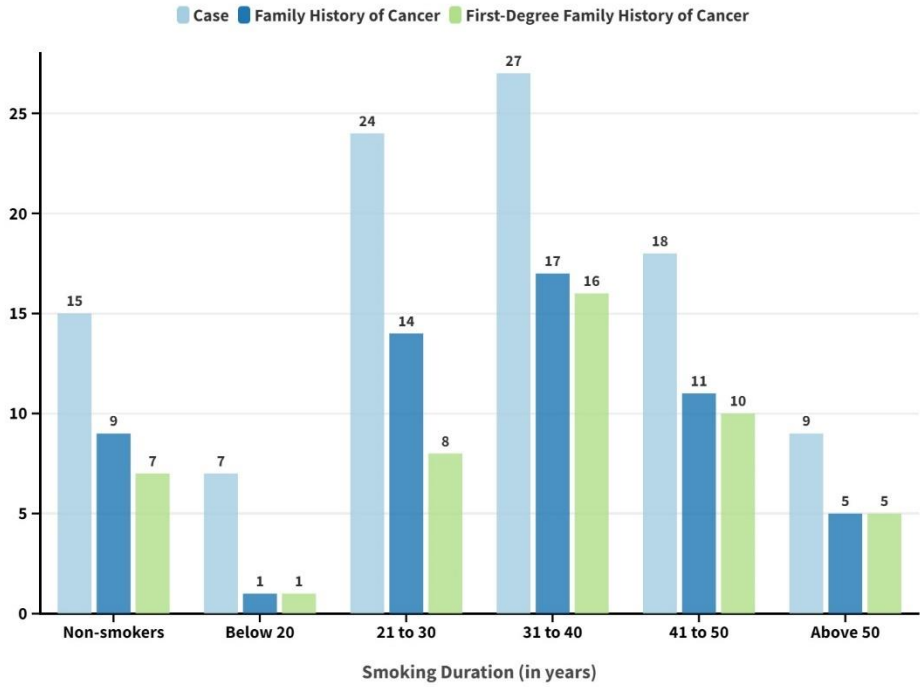



Figure 7 (a). Distribution of General FHC and First Degree FHC with smoking duration

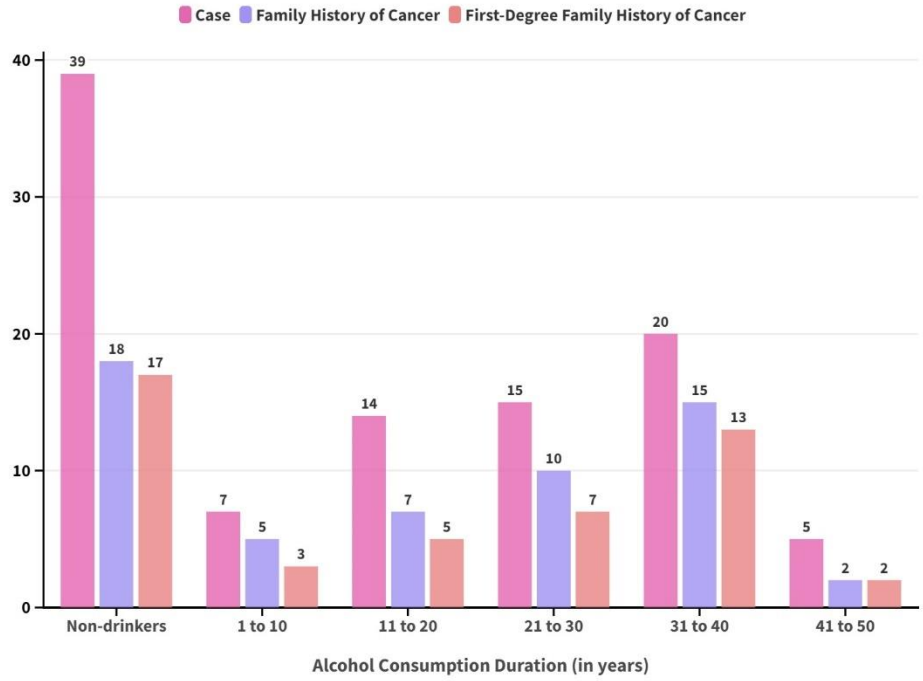



Figure 7 (b). Distribution of General FHC and First Degree FHC with alcohol consumption duration

Table 12. Distribution of cancer site and types of smoke and alcohol consumed in the cases alone for epidemiological study

| **Site** | **Total** |
|---|---|
| Oral Cavity | 41 |
| Hypopharynx | 30 |
| Oropharynx | 11 |
| Larynx | 9 |
| Nasopharynx | 9 |

| **Type of smoke** | **Total** |
|---|---|
| Non-smokers | 15 |
| Zozial (local cigarette) | 43 |
| Branded (Indian) | 19 |
| Both | 23 |

| **Type of Alcohol** | **Total** |
|---|---|
| Non-drinkers | 38 |
| Local liquor | 30 |
| Branded | 11 |
| Both | 21 |

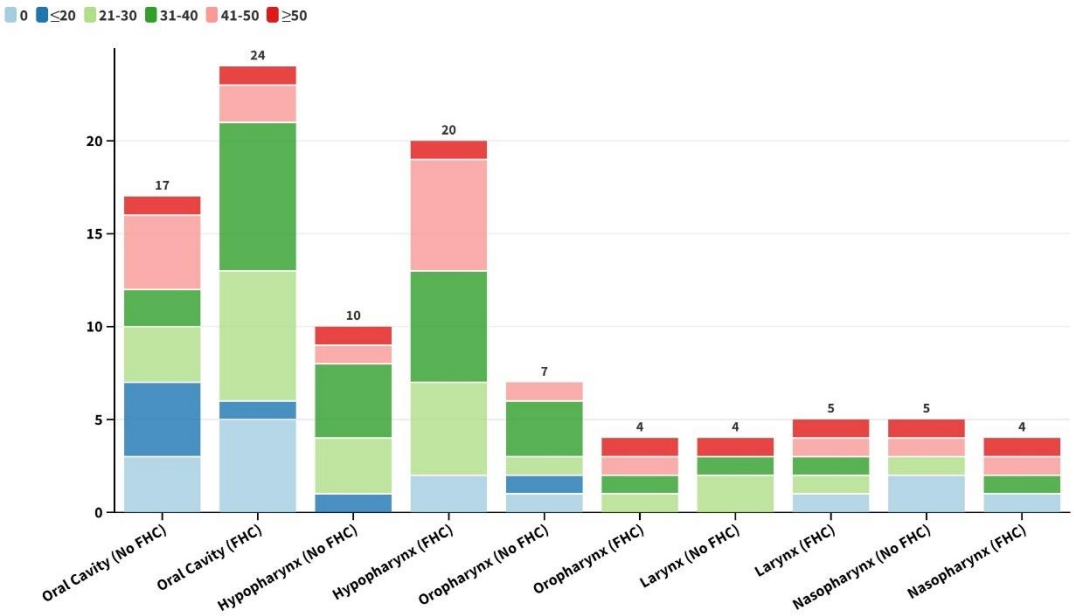



Figure 8 (a). Distribution of FHC and smoking duration with site-wise HNSCC ‡

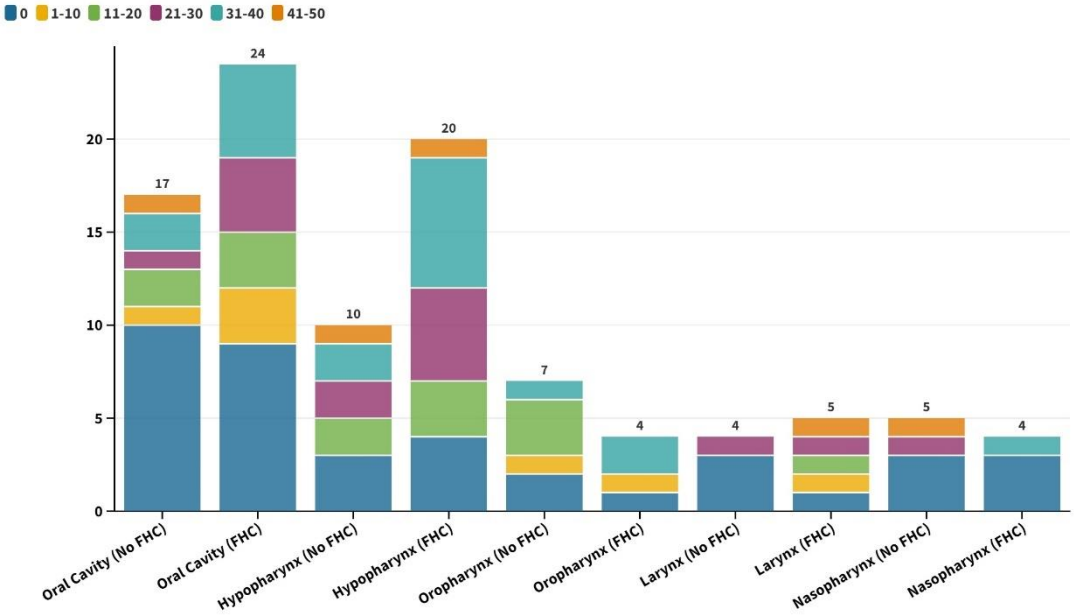



Figure 8 (b). Distribution of FHC and alcohol consumption duration with site-wise HNSCC ‡

‡ The bars indicate the total number of cases without family history (none) and with family history (FHC) for each site of cancer. Different colours represent the number of cases in different ranges of smoking or alcohol consumption.

Table 13. Characteristics of HNSCC patients included for survival analysis

| Characteristics | Variables | n | % |
|---|---|---|---|
| Age (years) | Minimum | 21 | |
| | Maximum | 84 | |
| | Median | 55 | |
| Gender | Male | 166 | 79 |
| | Female | 44 | 21 |
| Site | Hypopharynx | 67 | 31.9 |
| | Larynx | 25 | 11.9 |
| | Nasopharynx | 48 | 22.9 |
| | Oropharynx | 31 | 14.8 |
| | Oral Cavity | 39 | 18.6 |
| T Classification | I | 68 | 32.4 |
| | II | 86 | 41 |
| | III | 32 | 15.2 |
| | IV | 24 | 11.4 |
| N Classification | 0 | 62 | 29.5 |
| | I | 89 | 42.4 |
| | II | 55 | 26.2 |
| | III | 4 | 1.9 |
| Grading | Well differentiated | 20 | 9.5 |
| | Moderately differentiated | 63 | 30.5 |
| | Poorly differentiated | 23 | 11 |
| | Undifferentiated | 11 | 5.2 |
| | Not available | 93 | 43.8 |
| Total Leukocyte Count | ≤ 10 thou/cumm | 108 | 51.4 |
| | > 10 thou/cumm | 24 | 11.4 |
| | Not available | 78 | 37.1 |
| Absolute Neutrophil Count | ≤ 7 thou/cumm | 107 | 51 |
| | > 7 thou/cumm | 21 | 10 |
| | Not available | 82 | 39 |
| Cigarette Smoking | No | 29 | 14.3 |
| | Yes | 164 | 78.1 |
| | Not Available | 17 | 7.6 |
| Alcohol | No | 86 | 41 |
| | Yes | 107 | 51 |
| | Not Available | 17 | 8.1 |
| Smokeless tobacco | No | 82 | 39 |
| | Yes | 111 | 52.9 |
| | Not Available | 17 | 8.1 |
| Betelnut Chewing | No | 23 | 11 |
| | Yes | 170 | 81 |
| | Not Available | 17 | 8.1 |
| Family History of Cancer | No | 112 | 53.3 |
| | Yes | 71 | 33.8 |
| | Not Available | 27 | 12.9 |

Table 14. Distribution of patients consuming alcohol, smokeless tobacco and cigarette smoking

| Factors | Count |
|---|---|
| Smoking | 7 |
| Alcohol | 0 |
| Smokeless Tobacco | 3 |
| Betelnut | 3 |
| Smoking + Alcohol | 3 |
| Smoking + Smokeless Tobacco | 8 |
| Smoking + Betelnut | 14 |
| Alcohol + Betelnut | 1 |
| Smokeless Tobacco + Betelnut | 19 |
| Smoking + Alcohol + Smokeless Tobacco | 1 |
| Smoking + Alcohol + Betelnut | 53 |
| Alcohol + Smokeless Tobacco + Betelnut | 2 |
| Smoking + Smokeless Tobacco + Betelnut | 31 |
| Alcohol + Smoking + Smokeless Tobacco + Betelnut | 47 |
| None | 1 |
| Missing | 17 |
| Total | 210 |

Table 15. Treatment regime and response

| Characteristics | Variables | n | % |
|---|---|---|---|
| **Induction Chemotherapy Regimen** | Cisplatin + Paclitaxel | 54 | 25.7 |
| | Cisplatin + 5-Fluorouracil | 12 | 5.7 |
| | Cisplatin + Docetaxel | 3 | 7.1 |
| | Carboplatin + Paclitaxel | 15 | 1.4 |
| | Carboplatin + Docetaxel | 1 | 0.5 |
| | Not received | 125 | 59.5 |
| **Number of Induction Chemotherapy cycles** | Median | 3 | |
| | Range | 1 to 7 | |
| **Concurrent Chemo-Radiotherapy (CCRT)** | Cisplatin | 123 | 58.6 |
| | Carboplatin | 22 | 10.5 |
| | Paclitaxel | 4 | 1.9 |
| | Not Available | 4 | 1.9 |
| | Not received | 57 | 27.1 |
| **Number of CCRT weekly cycles** | Median | 6 | |
| | Range | 1 to 8 | |
| **Radiotherapy (RT) intention** | Radical | 184 | 87.6 |
| | Adjuvant | 17 | 8.1 |
| | Palliative | 9 | 4.3 |
| **RT dose (Gray)** | Median | 66 | |
| | Range | 24 to 70 | |
| **Overall Survival** | Alive | 168 | 80 |
| | Dead | 42 | 20 |
| **Progression Free Survival** | Complete Response | 117 | 55.7 |
| | Partial Response | 8 | 3.8 |
| | Stable Disease | 1 | 0.4 |
| | Progressive Disease | 84 | 40 |
| **Progression** | Distant Metastasis | 9 | |
| | Regional Metastasis | 3 | |
| | Recurrence | 20 | |

Table 16. Distribution of T and N classifications among different treatment modalities

| | Treatment Modalities | | | | |
|---|---|---|---|---|---|
| | IC+CCRT/IC+RT | CCRT | RT | S+CCRT/S+RT | Total |
| T1 | 27 | 28 | 10 | 3 | 68 |
| T2 | 31 | 42 | 7 | 6 | 86 |
| T3 | 13 | 10 | 4 | 5 | 32 |
| T4 | 14 | 6 | 1 | 3 | 24 |
| N0 | 12 | 30 | 15 | 5 | 62 |
| N1 | 42 | 37 | 3 | 7 | 89 |
| N2 | 27 | 19 | 4 | 5 | 55 |
| N3 | 4 | 0 | 0 | 0 | 4 |

T - Tumour classification, N - Nodal classification, IC - Induction Chemotherapy, CCRT - Concurrent Chemoradiotherapy, RT - Radiotherapy, S - Surgery.

Table 17. Doses of Chemotherapy drugs administered to the patients

| **Chemotherapy** | **Range (in mg)** |
|---|---|
| Induction Chemotherapy | |
| Cisplatin | 50 - 600 |
| Paclitaxel | 190 - 543 |
| Carboplatin | 110 - 570 |
| Docetaxel | 60 - 80 |
| 5 Fluorouracil | 600 - 1500 |
| Concurrent Chemoradiotherapy | |
| Carboplatin | 120 - 180 |
| Cisplatin | 40 - 150 |
| Paclitaxel | 30 - 90 |

Table 18. Kaplan–Meier estimates and log-rank test for two years overall survival (OS) and progression free survival (PFS) of treatment regimen

| Characteristics | Variables | N | Overall survival | | | Progression free survival | | |
|---|---|---|---|---|---|---|---|---|
| | | | Survival rates (%) | 95% CI | p-value | Survival rates (%) | 95% CI | p-value |
| | **Overall** | 210 | 78.1 | | | 57.4 | 50.8–64.8 | |
| **Treatment type** | **IC + CCRT/IC + RT** | 85 | 73.7 | 64.3–84.4 | 0.294 | 47.3 | 37.2–60.0 | 0.062 |
| | **CCRT** | 86 | 83.7 | 75.9–92.2 | | 66 | 56.5–77.1 | |
| | **RT** | 22 | 70.4 | 53.0–93.5 | | 61.8 | 44.1–86.7 | |
| | **S + CCRT/S + RT** | 17 | 80 | 62.1–100 | | 56.2 | 36.5–86.7 | |
| **IC regimen** | **CP + PAX** | 54 | 77.5 | 66.6–90.3 | 0.463 | 51.7 | 39.5–67.8 | 0.075 |
| | **CP + 5-FU** | 12 | 44.4 | 21.4–92.3 | | 18.8 | 05.4–65.0 | |
| | **CP + DOX** | 3 | 66.7 | 30.0–100 | | 66.6 | 30.0–100 | |
| | **CB + PAX** | 15 | 72.2 | 52.4–99.6 | | 30 | 12.3–73.4 | |
| **IC vs No IC** | **IC** | 85 | 73.7 | 64.3–84.4 | 0.216 | 47.3 | 37.2–60.0 | 0.010[a] |
| | **No IC** | 125 | 80.9 | 74.1–88.4 | | 63.9 | 55.8–73.1 | |
| **Total Leukocyte Count** | **≤10 thou/cumm** | 108 | 81.3 | 73.9–89.6 | 0.015[a] | 56.9 | 47.9–67.6 | 0.076 |
| | **>10 thou/cumm** | 24 | 58.4 | 40.7–83.9 | | 39.2 | 23.6–65.2 | |
| **Absolute Neutrophil Count** | **≤7 thou/cumm** | 107 | 81.1 | 73.6–89.4 | 0.014[a] | 57.5 | 48.4–68.2 | 0.043[a] |
| | **>7 thou/cumm** | 23 | 57 | 39.2–83.1 | | 36.4 | 21.0–63.3 | |
| **Site** | **Hypopharynx** | 67 | 83.7 | 74.9–93.5 | 0.101 | 65.7 | 55.0–78.4 | 0.525 |
| | **Nasopharynx** | 48 | 88 | 78.5–98.5 | | 51.9 | 38.7–69.7 | |
| | **Larynx** | 25 | 75.1 | 59.6–94.6 | | 63 | 46.4–85.6 | |
| | **Oropharynx** | 31 | 69.4 | 54.5–88.3 | | 53.9 | 38.8–75.0 | |
| | **Oral Cavity** | 39 | 66.7 | 52.8–84.1 | | 48.2 | 34.4–67.5 | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **N** | **N0** | 62 | 86.1 | 77.7–95.5 | 0.062 | 68.3 | 57.5–81.2 | 0.005[a] |
| | **N1** | 89 | 79.4 | 71.1–88.6 | | 60.2 | 50.4–71.7 | |
| | **N2** | 55 | 65.7 | 53.1–81.4 | | 39.8 | 28.1–56.5 | |
| | **N3** | 4 | 75 | 42.6–100 | | 50 | 18.8–100 | |

IC, Induction Chemotherapy; CCRT, Concurrent Chemoradiotherapy; RT, Radiotherapy; S, Surgery. CP, Cisplatin; PAX, Paclitaxel; 5-FU, 5-Flourouracil; DOX, Docetaxel; CB, Carboplatin. thou/cumm, thousand cells per mm3. [a]Statistically significant (p-value <0.05).

Table 19. Univariate and Multivariate analysis for characteristics of patients, tumour and treatment regimen with overall survival

| **Characteristics** | **Variables** | **Univariate** | | **Multivariate** | |
|---|---|---|---|---|---|
| | | **HR (95% CI)** | **p-value** | **HR[a] (95% CI)** | **p-value** |
| | **Age** | 1.007 (0.977–1.038) | 0.658 | 1.003 (0.963–1.046) | 0.88 |
| **Sex** | **Male** | Reference | | | |
| | **Female** | 0.914 (0.423–1.976) | 0.82 | 0.404 (0.092–1.773) | 0.23 |
| **Site** | **Hypopharynx** | Reference | | | |
| | **Larynx** | 1.760 (0.639–4.843) | 0.274 | 5.165 (1.518–17.570) | 0.009[b] |
| | **Nasopharynx** | 0.732 (0.250–2.141) | 0.569 | 0.452 (0.117–1.744) | 0.249 |
| | **Oropharynx** | 2.109 (0.857–5.192) | 0.104 | 1.655 (0.518–5.285) | 0.395 |
| | **Oral Cavity** | 2.226 (0.962–5.154) | 0.062 | 2.273 (0.768–6.728) | 0.138 |
| **T Classification** | **1** | Reference | | | |
| | **2** | 1.305 (0.611–2.786) | 0.492 | 0.883 (0.347–2.248) | 0.794 |
| | **3** | 1.052 (0.366–3.028) | 0.925 | 0.740 (0.156–3.510) | 0.705 |
| | **4** | 2.822 (1.169–6.815) | 0.021[b] | 2.073 (0.698–6.156) | 0.189 |

| | | | | | |
|---|---|---|---|---|---|
| **N Classification** | **0** | Reference | | | |
| | **1** | 1.621 (0.699–3.756) | 0.26 | 2.329 (0.814–6.664) | 0.115 |
| | **2** | 2.954 (1.263–6.908) | 0.012[b] | 3.835 (1.231–11.946) | 0.020[b] |
| | **3** | 2.506 (0.313–20.051) | 0.387 | | – |
| **Total Leukocyte Count (TLC)** | **≤10 thou/cumm** | Reference | | | |
| | **>10 thou/cumm** | 2.603 (1.167–5.803) | 0.019[b] | 2.951 (1.290–6.748) | 0.010[b] |
| **Absolute Neutrophil Count (ANC)** | **≤7 thou/cumm** | Reference | | | |
| | **>7 thou/cumm** | 2.625 (1.177–5.852) | 0.018[b] | 2.500 (1.100–5.684) | 0.029[b] |
| **Alcohol intake** | **No** | Reference | | | |
| | **Yes** | 1.888 (0.956–3.727) | 0.067 | 2.487 (0.957–6.460) | 0.061 |
| **Smoking** | **No** | Reference | | | |
| | **Yes** | 1.666 (0.592–4.689) | 0.333 | 1.423 (0.326–6.211) | 0.639 |
| **Betelnut use** | **No** | Reference | | | |
| | **Yes** | 1.148 (0.408–3.230) | 0.794 | 2.446 (0.523–11.430) | 0.256 |
| **Smokeless tobacco** | **No** | Reference | | | |
| | **Yes** | 0.798 (0.426–1.496) | 0.418 | 0.788 (0.359–1.731) | 0.553 |
| **Family history of cancer** | **No** | Reference | | | |
| | **Yes** | 0.545 (0.255–1.169) | 0.119 | 0.935 (0.340–2.572) | 0.896 |
| **Grading** | **Well differentiated** | Reference | | | |
| | **Moderately differentiated** | 0.790 (0.278–2.243) | 0.658 | 2.826 (0.302–26.475) | 0.363 |
| | **Poorly differentiated** | 0.623 (0.149–2.609) | 0.517 | 1.942 (0.136–27.677) | 0.624 |
| | **Undifferentiated** | 0.332 (0.040–2.960) | 0.332 | 0.622 (0.034–11.261) | 0.748 |

[a]Hazard Ratio adjusted for T classification, N Classification and Total Leukocyte Count (TLC) except for Absolute Neutrophil Count (ANC). ANC was adjusted for T and N Classification. [b]Statistically significant (p-value <0.05).

Table 20. Univariate and Multivariate analysis for characteristics of patients, tumour and treatment regimen with progression free survival

| | Characteristics | Univariate | | Multivariate | |
|---|---|---|---|---|---|
| | | HR (95% CI) | p-value | HRa (95% CI) | p-value |
| | **Age** | 1.003 (0.982–1.024) | 0.789 | 1.624 (0.807–3.268) | 0.174 |
| **Sex** | **Male** | Reference | | | |
| | **Female** | 0.890 (0.517–1.534) | 0.675 | 0.439 (0.174–1.107) | 0.081 |
| **Site** | **Hypopharynx** | Reference | | | |
| | **Larynx** | 1.180 (0.544–2.564) | 0.675 | 2.844 (1.117–7.244) | 0.028[b] |
| | **Nasopharynx** | 1.439 (0.791–2.617) | 0.233 | 1.236 (0.603–2.532) | 0.563 |
| | **Oropharynx** | 1.537 (0.786–3.004) | 0.209 | 1.853 (0.793–4.329) | 0.154 |
| | **Oral Cavity** | 1.636 (0.885–3.023) | 0.116 | 1.757 (0.752–4.103) | 0.193 |
| **T Classification** | **1** | Reference | | | |
| | **2** | 1.023 (0.615–1.701) | 0.93 | 0.806 (0.424–1.531) | 0.51 |
| | **3** | 1.109 (0.572–2.149) | 0.76 | 1.917 (0.799–4.600) | 1.145 |
| | **4** | 1.555 (0.788–3.071) | 0.203 | 1.038 (0.447–2.409) | 0.931 |
| **N Classification** | | Reference | | | |
| | **1** | 1.366 (0.777–2.403) | 0.279 | 1.582 (0.782–3.198) | 0.202 |
| | **2** | 2.574 (1.452–4.562) | 0.001[b] | 3.483 (1.706–7.110) | 0.001[b] |
| | **3** | 2.104 (0.490–9.034) | 0.317 | 6.527 (0.830–51.347) | 0.075 |
| **Total Leukocyte Count (TLC)** | **≤10 thou/cumm** | Reference | | | |
| | **>10 thou/cumm** | 1.718 (0.939–3.144) | 0.079 | 2.035 (1.095–3.782) | 0.025[b] |

| | | | | | |
|---|---|---|---|---|---|
| **Absolute Neutrophil Count (ANC)** | **≤7 thou/cumm** | Reference | | | |
| | **>7 thou/cumm** | 1.849 (1.009–3.389) | 0.047[b] | 1.946 (1.056–3.586) | 0.033[b] |
| **Alcohol intake** | **No** | Reference | | | |
| | **Yes** | 1.167 (0.746–1.825) | 0.499 | 1.501 (0.849–2.651) | 0.162 |
| **Smoking** | **No** | Reference | | | |
| | **Yes** | 1.719 (0.827–3.570) | 0.147 | 2.182 (0.670–7.105) | 0.195 |
| **Betelnut** | **No** | Reference | | | |
| | **Yes** | 0.801 (0.424–1.515) | 0.495 | 1.443 (0.656–3.176) | 0.362 |
| **Smokeless tobacco** | **No** | Reference | | | |
| | **Yes** | 1.082 (0.692–1.691) | 0.731 | 1.124 (0.645–1.958) | 0.68 |
| **Family history of cancer** | **No** | Reference | | | |
| | **Yes** | 0.755 (0.462–1.235) | 0.263 | 0.991 (0.502–1.955) | 0.979 |
| **Grading** | **Well differentiated** | Reference | | | |
| | **Moderately differentiated** | 0.861 (0.398–1.862) | 0.704 | 2.682 (0.557–12.924) | 0.219 |
| | **Poorly differentiated** | 1.427 (0.601–3.391) | 0.42 | 2.506 (0.466–13.478) | 0.284 |
| | **Undifferentiated** | 0.819 (0.252–2.661) | 0.74 | 2.052 (0.362–11.648) | 0.417 |

[a]Hazard Ratio adjusted for N classification and Absolute Neutrophil Count (ANC) except Total Leukocyte Count (TLC) which was adjusted for N classification only. thou/cumm, thousand cells per mm3. [b]Statistically significant (p-value <0.05)

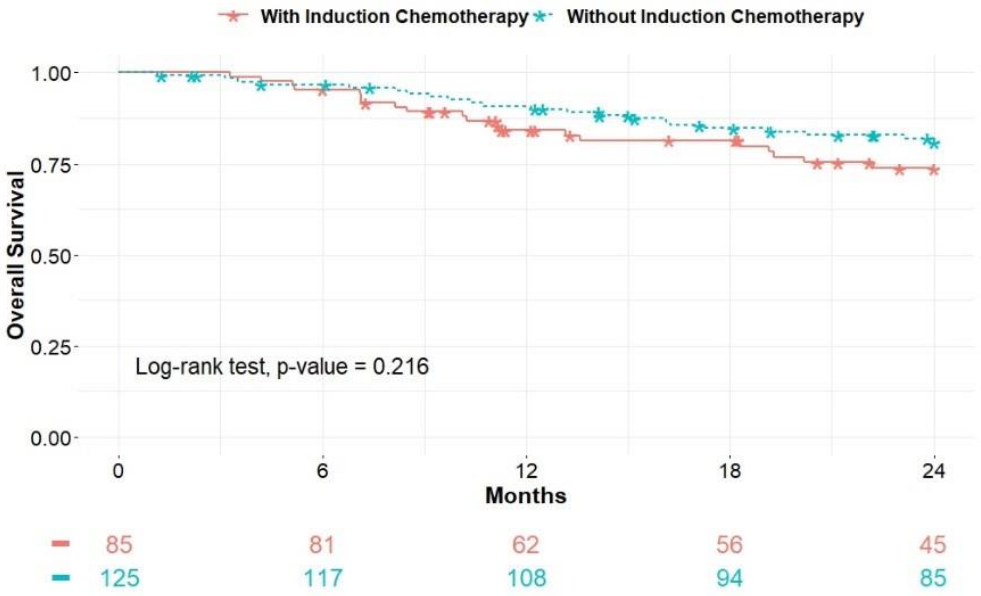



Figure 9 (a). Overall Survival in patients who received Induction Chemotherapy against those who did not

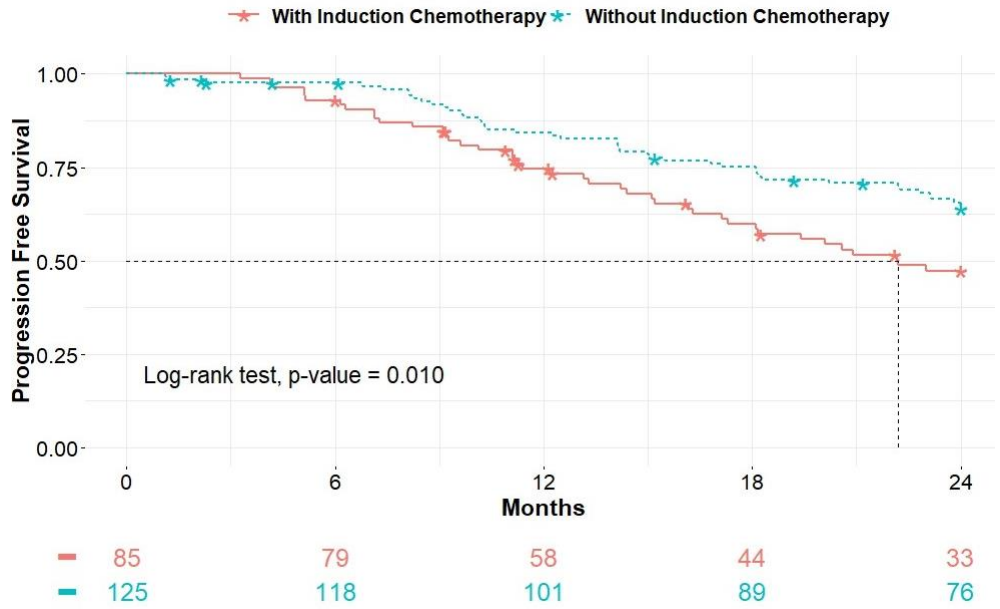



Figure 9 (b). Progression Free Survival in patients who received Induction Chemotherapy against those who did not

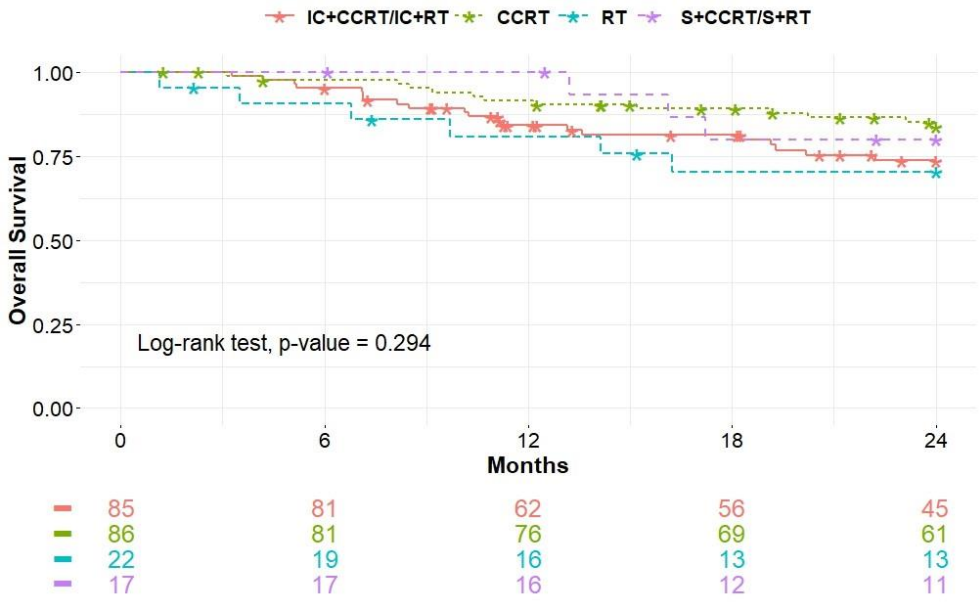



Figure 10 (a). Overall Survival between treatment modalities

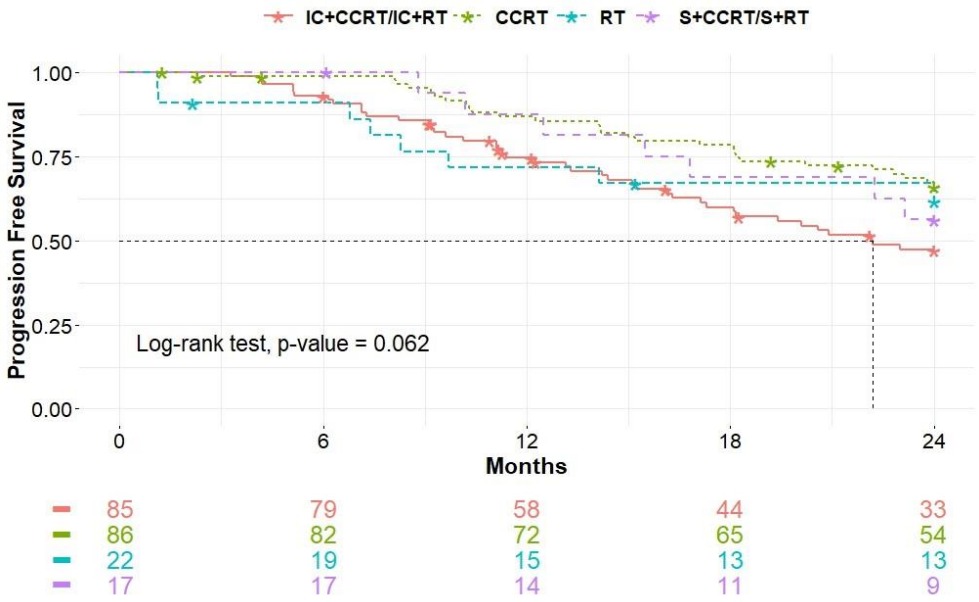



Figure 10 (b). Progression Free Survival between treatment modalities

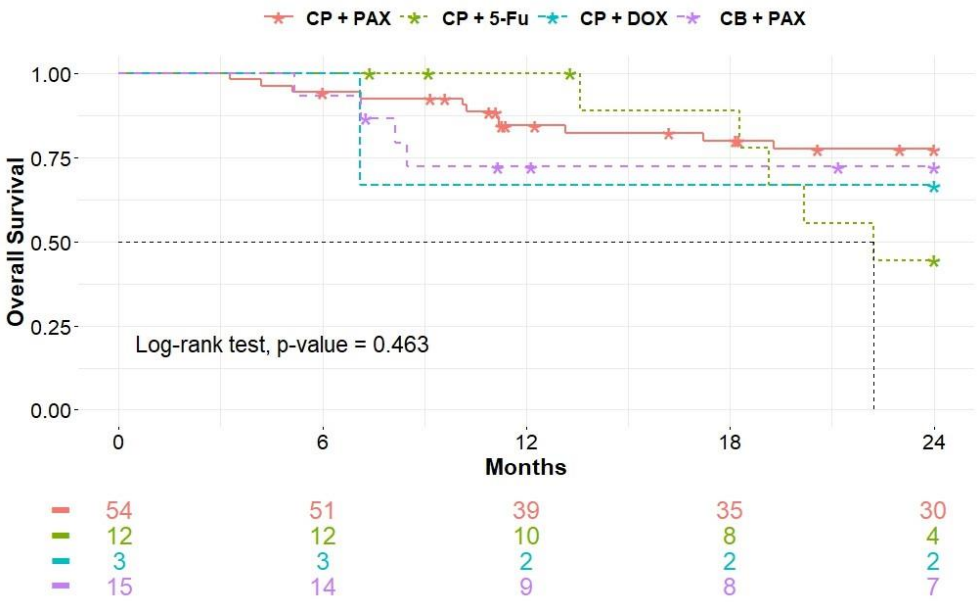



Figure 11 (a). Overall Survival between Induction Chemotherapy regimen

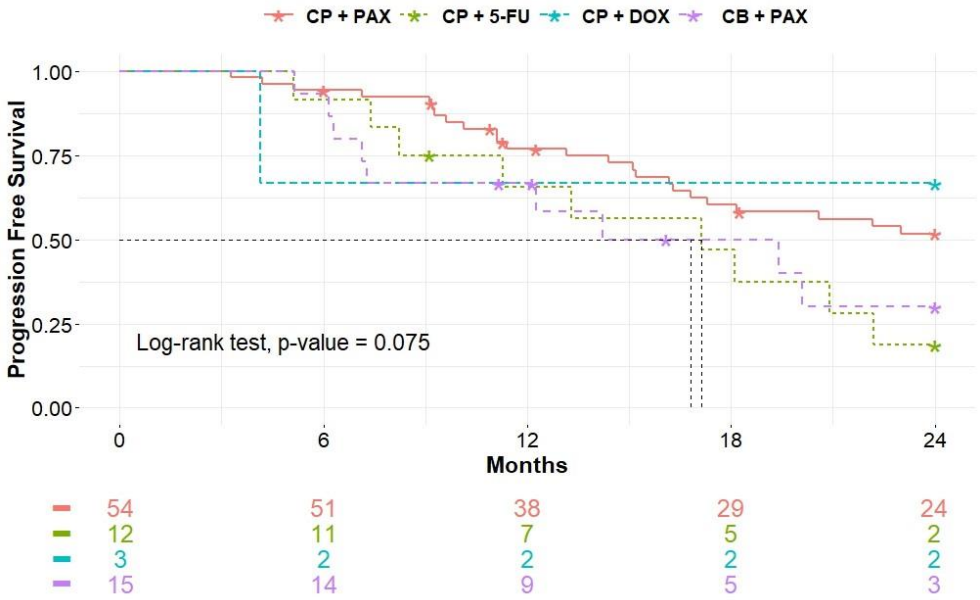



Figure 11 (b). Progression Free Survival between Induction Chemotherapy regimen

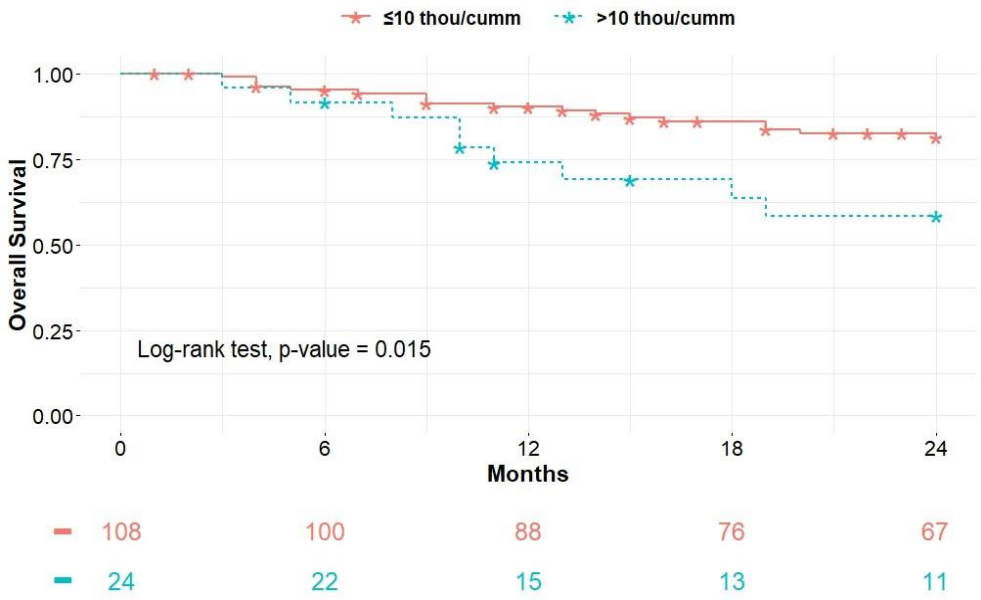



Figure 12 (a). Overall Survival between Total Leukocyte Count (TLC) ≤10 and >10 thou/cumm

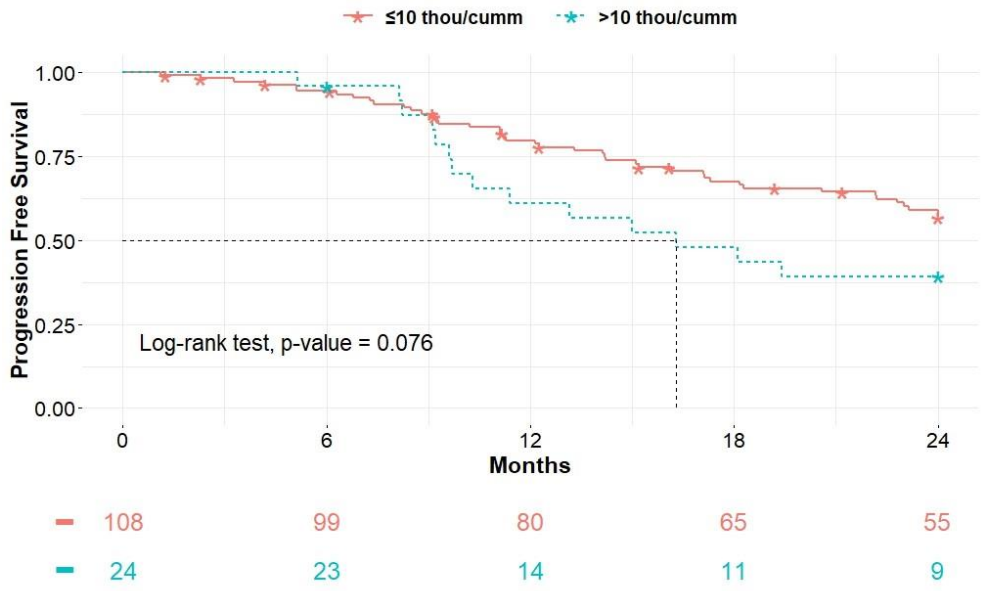



Figure 12 (b). Progression Free Survival between Total Leukocyte Count (TLC) ≤10 and >10 thou/cumm

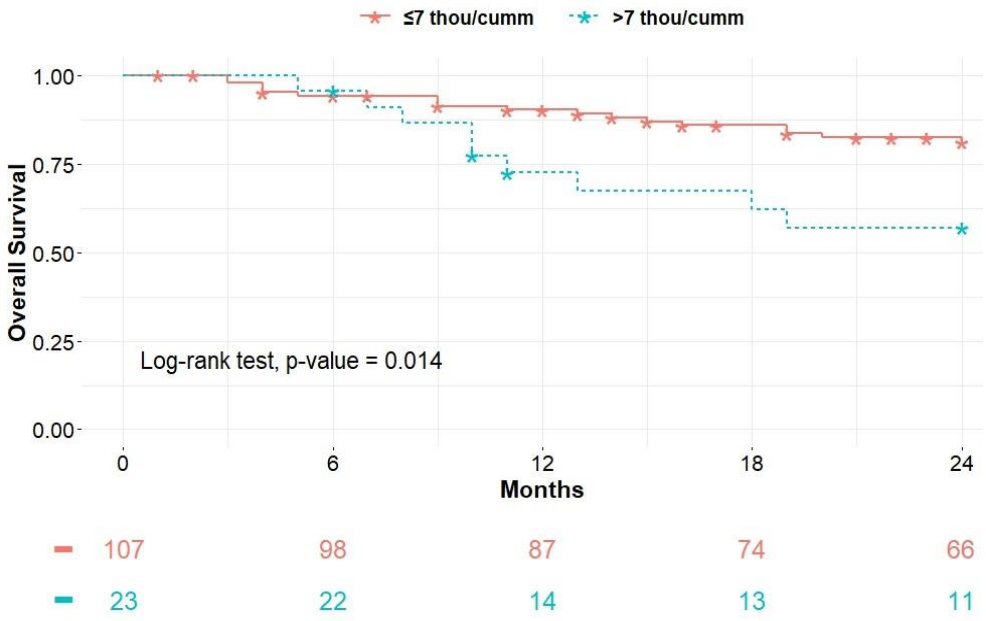



Figure 13 (a). Overall Survival between Absolute Neutrophil Count (ANC) ≤7 and >7 thou/cumm

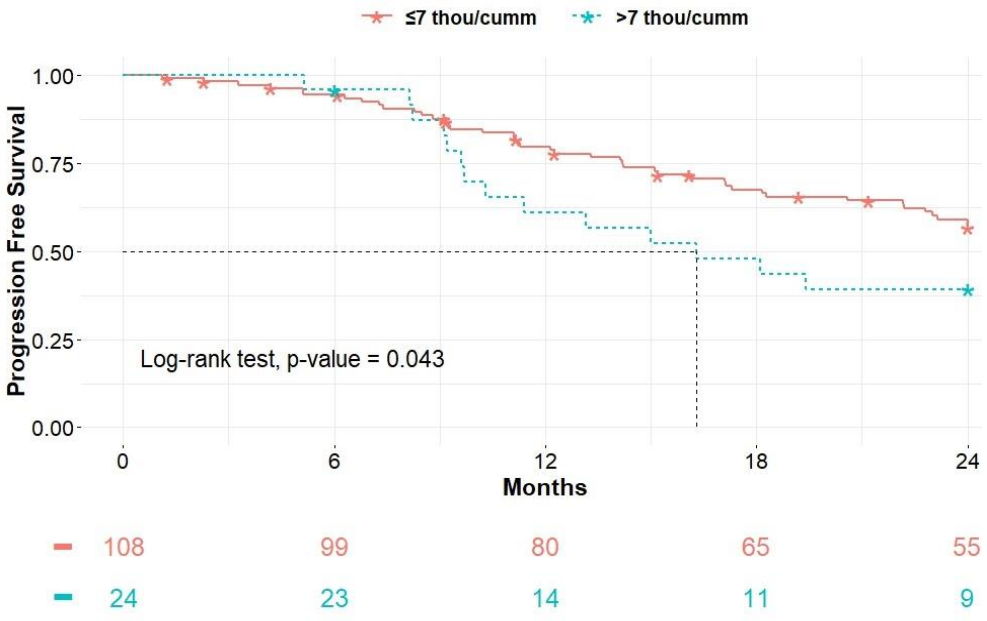



Figure 13 (b). Progression Free Survival between Absolute Neutrophil Count (ANC) ≤7 and >7 thou/cumm

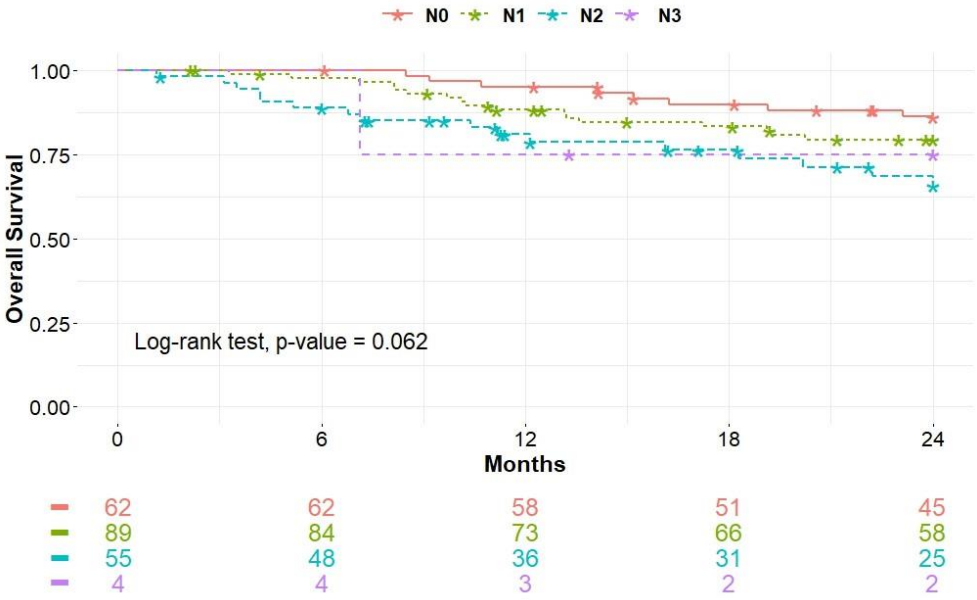



Figure 14 (a). Overall Survival between levels of Nodal (N) involvement

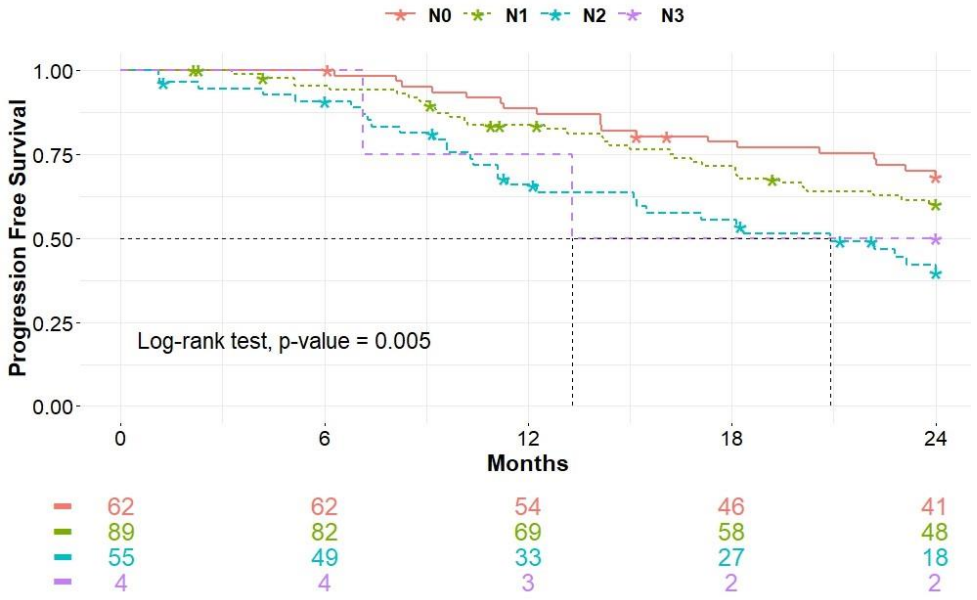



Figure 14 (b). Progression Free Survival between levels of Nodal (N) involvement

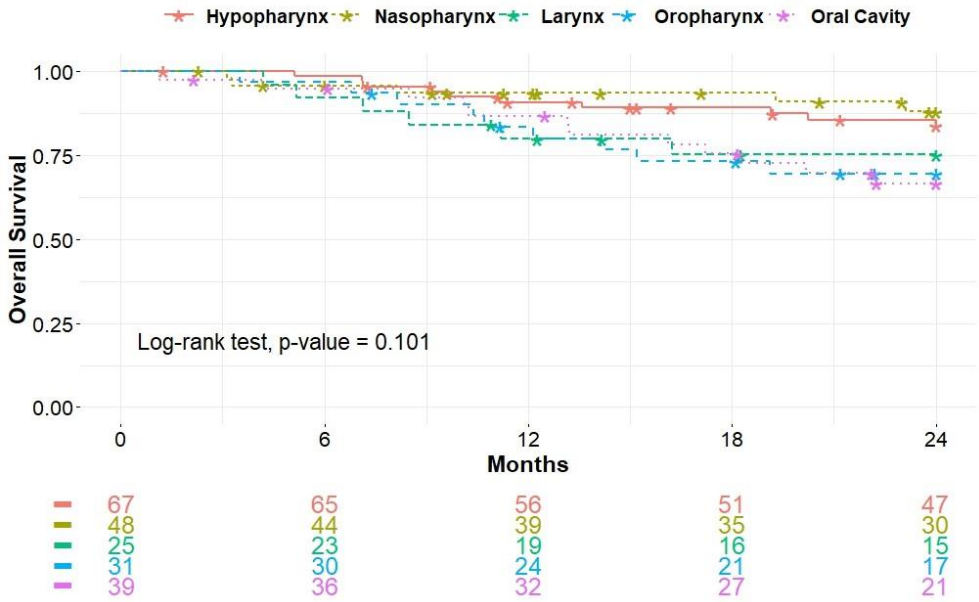



Figure 15 (a). Overall Survival between the cancer sites

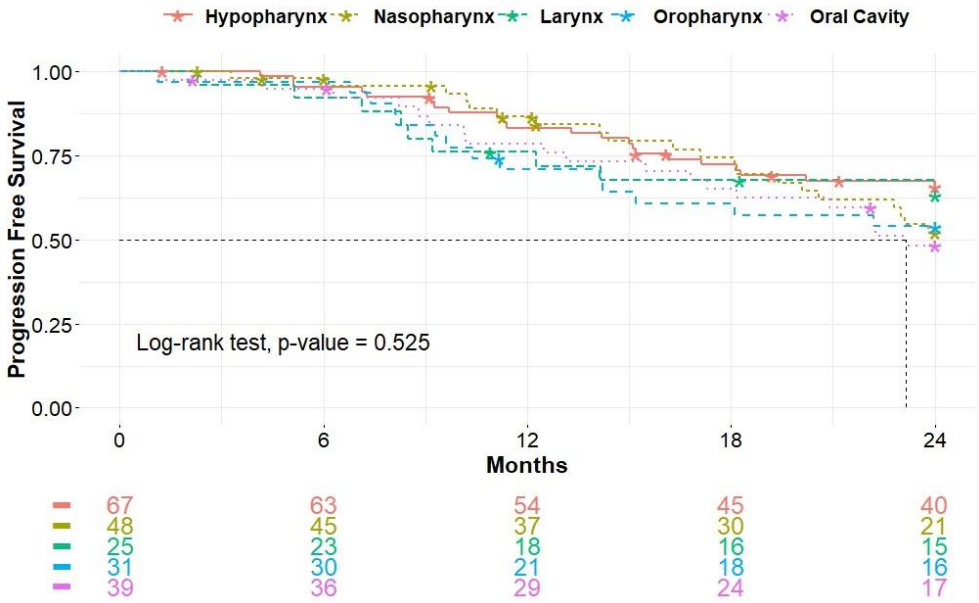



Figure 15 (b). Progression Free Survival between the cancer sites

Table 21. Characteristics of patients' clinical and lifestyle factors for proteome analysis

| Patient Number | Age (years) | Gender | Cancer Site | Stage | T | N | TLC | ANC | Platelets | Smoking | Alcohol | Tobacco | Betelnut | Family History |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 53 | Male | Hypopharynx | III | 3 | 1 | 7200 | 3900 | 2.1 | Yes | Yes | Yes | Yes | NA |
| 2 | 65 | Male | Larynx | I | 1 | 0 | 6300 | 4000 | 2.5 | Yes | No | Yes | Yes | NA |
| 3 | 50 | Male | Oral Cavity | IV | 1 | 2 | NA | NA | NA | Yes | Yes | Yes | Yes | Yes |
| 4 | 49 | Female | Nasopharynx | II | 2 | 1 | 8800 | 5100 | 3.6 | Yes | Yes | Yes | Yes | No |
| 5 | 39 | Male | Oropharynx | I | 1 | 0 | 8600 | 5000 | 2.04 | Yes | Yes | Yes | Yes | NA |
| 6 | 54 | Male | Hypopharynx | II | 2 | 0 | 9000 | 5400 | 2.9 | Yes | Yes | Yes | Yes | No |
| 7 | 44 | Male | Hypopharynx | III | 3 | 1 | 11100 | 6900 | 3.2 | Yes | Yes | No | Yes | No |
| 8 | 54 | Male | Oropharynx | II | 2 | 1 | 10700 | 8300 | 4.8 | Yes | Yes | Yes | Yes | NA |
| 9 | 57 | Male | Oropharynx | IV | 3 | 2 | 7400 | 4366 | 2.4 | Yes | Yes | No | Yes | NA |
| 10 | 55 | Male | Nasopharynx | III | 1 | 2 | 12500 | 6900 | 1.7 | Yes | Yes | No | Yes | NA |
| 11 | 38 | Male | Salivary Gland | IV | 2 | 2 | NA | NA | NA | NA | NA | NA | NA | NA |
| 12 | 56 | Male | Larynx | II | 2 | 0 | 6100 | 4400 | 3.1 | Yes | Yes | No | Yes | NA |
| 13 | 60 | Male | Hypopharynx | II | 2 | 1 | NA | NA | NA | Yes | Yes | Yes | Yes | Yes |
| 14 | 55 | Male | Oral Cavity | III | 3 | 1 | 8300 | 6400 | 2 | Yes | Yes | Yes | Yes | Yes |
| 15 | 70 | Male | Hypopharynx | III | 2 | 2 | 5300 | NA | 1.74 | Yes | Yes | Yes | Yes | No |
| 16 | 63 | Male | Larynx | I | 1 | 0 | 8900 | 7000 | 2.5 | Yes | Yes | No | Yes | No |
| 17 | 49 | Male | Hypopharynx | III | 2 | 1 | 11500 | 8800 | 3.6 | Yes | Yes | Yes | Yes | No |
| 18 | 64 | Male | Hypopharynx | III | 1 | 1 | 9400 | 5500 | 3.7 | Yes | Yes | Yes | Yes | No |
| 19 | 71 | Male | Hypopharynx | III | 2 | 2 | 15500 | 11900 | 3.2 | Yes | Yes | Yes | Yes | No |
| 20 | 56 | Male | Larynx | II | 2 | 0 | 8400 | 4788 | 2.1 | Yes | Yes | Yes | Yes | No |

T - Tumour Classification, N - Nodal Classification, TLC - Total Leukocyte Count, ANC - Absolute Neutrophil Count

Table 22. Characteristics of patients' treatment regimen and response for proteome analysis

| Patient Number | Radiation intention | Radiation (Gy) | Radiation Fractions | Induction Chemotherapy | Concurrent Chemotherapy | Overall Survival | Progression Free Survival |
|---|---|---|---|---|---|---|---|
| 1 | Radical | 66 | 33 | Paclitaxel (230 mg) & Cisplatin (100 mg) | Cisplatin (40 mg) | Death | Progressive Disease |
| 2 | Radical | 66 | 33 | No Induction Chemotherapy | NA | Death | Progressive Disease |
| 3 | Radical | 66 | 33 | No Induction Chemotherapy | Carboplatin (150 mg) | Lost | Lost |
| 4 | Radical | 70 | 35 | No Induction Chemotherapy | Cisplatin (40 mg) | Alive | Complete Response |
| 5 | Radical | 60 | 30 | No Induction Chemotherapy | Cisplatin (60 mg) | Alive | Progressive Disease |
| 6 | Radical | 66 | 33 | No Induction Chemotherapy | Cisplatin (48 mg) | Alive | Complete Response |
| 7 | Radical | 66 | 33 | No Induction Chemotherapy | Cisplatin (40 mg) | Lost | Lost |
| 8 | Radical | 66 | 33 | No Induction Chemotherapy | Cisplatin (40mg) | Death | Partial Response |
| 9 | Palliative | 30 | 10 | No Induction Chemotherapy | No Chemotherapy | Death | Progressive Disease |
| 10 | Radical | 70 | 35 | No Induction Chemotherapy | Cisplatin (40 mg) | Death | Progressive Disease |
| 11 | Palliative | 30 | 15 | No Induction Chemotherapy | No Chemotherapy | Alive | Progressive Disease |
| 12 | Radical | 66 | 33 | No Induction Chemotherapy | NA | Lost | Lost |
| 13 | Radical | 60 | 30 | No Induction Chemotherapy | Cisplatin (55 mg) | Alive | Complete Response |
| 14 | Radical | 66 | 33 | No Induction Chemotherapy | Cisplatin (50 mg) | Alive | Progressive Disease |
| 15 | Radical | 66 | 33 | No Induction Chemotherapy | Cisplatin (50 mg) | Lost | Lost |
| 16 | Radical | 60 | 30 | No Induction Chemotherapy | Cisplatin (50 mg) | Death | Complete Response |
| 17 | Radical | 70 | 35 | No Induction Chemotherapy | refused | Death | Partial Response |
| 18 | Radical | 60 | 30 | No Induction Chemotherapy | Cisplatin (50 mg) | Death | Progressive Disease |
| 19 | Radical | 70 | 35 | Paclitaxel (260 mg) & Cisplatin (110 mg) | Cisplatin (45 mg) | Death | Progressive Disease |
| 20 | Radical | 66 | 33 | No Induction Chemotherapy | Cisplatin (50 mg) | Alive | Complete Response |

134 (924) differentially expressed proteins were identified between pre-treated and during treatment

Total number of peptides ≥ 2
Number of unique peptides ≥ 1
Abundant proteins removed

110 (823) proteins

Volcano Plot
Fold change = 2
(q-value) ≤ 0.05

78 (304)

Out of 304 protein hits
155 proteins were downregulated
149 proteins were upregulated.

Figure 16. Flowchart representing the filtering steps of 78 differentially expressed proteins in the 20 samples

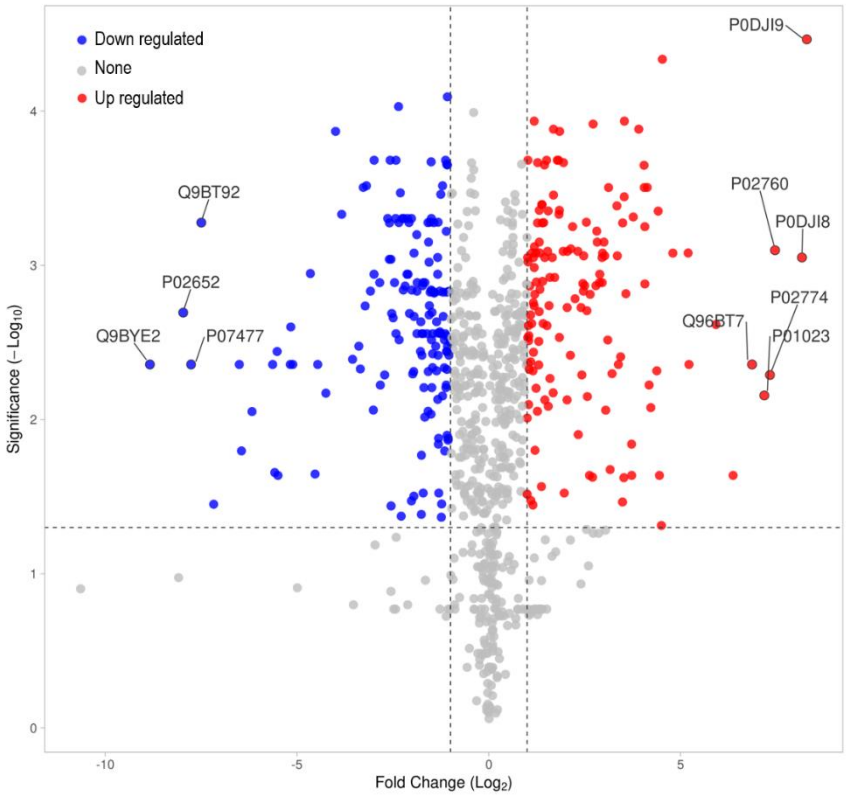



Figure 17. Volcano Plot showing the upregulated and downregulated proteins

A total of 40 serum samples were analysed to compare the differential protein expression between pretreated samples and samples collected 2 weeks during treatment from 20 patients with head and neck cancer. The characteristics of the patients' clinical and lifestyle factors and their treatment regimen and response are given in Table 21 & 22. The pretreated protein expression was treated as baseline for the changes observed in proteins expressions in 2 weeks treated samples. As shown in Figure 16, 134 differentially proteins were found in all the samples. After removing abundant proteins such as Albumin and Immunoglobulin G and filtering for proteins with at least 2 peptide counts and 1 or more unique peptides, we identified a total of 110 proteins. A Volcano Plot was employed to further refine the selection, narrowing down the significant differentially expressed proteins from 110 to 78 (Figure 17). This filtering was based on a fold change cutoff of 2 and a q-value of $\leq 0.05$. Among the 78 proteins, differential expression was observed 304 times across the samples. Of these, 155 proteins were downregulated, while 149 proteins were upregulated (Table 23, Figure 18).

The PPI analysis of 78 differentially expressed proteins identified 77 nodes and 691 interactions, selected based on a confidence score of $\geq 0.4$ using STRING analysis. Using Cytoscape, nodes with a degree $\geq 2$ were filtered, resulting in a network of 60 nodes and 690 interactions [Figure 19 (A)]. Two significant modules with scores greater than 5 were identified using MCODE. Module 1 contains 30 nodes and 389 interactions, with an MCODE score of 26.83 [Figure 19 (B)]. Module 2 comprises 7 nodes and 20 interactions, with an MCODE score of 6.68 [Figure 19 (C)].

To further explore the functional roles of the 78 differentially expressed proteins, Gene Ontology (GO) analysis was performed using the DAVID database. In the Cellular Component category, the proteins were predominantly enriched in blood microparticles, extracellular exosomes, the extracellular region, and extracellular space [Figure 20 (a)]. For Molecular Function, the proteins were primarily associated with organic acid binding, haptoglobin binding, oxygen transporter activity, and peroxidase activity [Figure 20 (b)]. In the Biological Process category, the proteins were mainly linked to the acute phase response, cellular oxidant detoxification, and cellular oxidant detoxification [Figure 20 (c)]. Additionally, KEGG pathway analysis

revealed that these proteins were mainly involved in pathways such as complement and coagulation cascades, cholesterol metabolism, and African trypanosomiasis [Figure 20 (d)].

Kaplan-Meier estimates and the log-rank test indicate that patients with downregulated SAA1 protein expression exhibit better OS (log-rank, p = 0.010) and PFS (log-rank, p = 0.005) compared to patients with upregulated or unchanged SAA1 expression during treatment relative to baseline pre-treatment levels (Figure 21). Similarly, patients with downregulated B2M expression show a better prognosis (log-rank, p = 0.047) compared to those with upregulated or unchanged expression during treatment (Figure 22). Additionally, patients with consistent HBB expression had better PFS (log-rank, p = 0.035) than those whose expression levels changed during treatment (Figure 22). Cox proportional hazards analysis showed that patients with higher levels of ASGH expression have a worse prognosis; however, the log-rank test comparing the categorical expression levels was not significant (Figure 23).

Table 23. Distribution of top 78 differentially expressed proteins across the 20 samples

| Protein name | Protein ID | Up | Down | No Change |
|---|---|---|---|---|
| Acylpyruvase FAHD1_ mitochondrial | FAHD1 | | 2 | |
| ADAMTS-like protein 1 | ADAMTSL1 | 1 | | |
| Afamin | AFM | | 1 | 1 |
| Alkylated D repair protein alkB homolog 8 | ALKBH8 | | 1 | |
| Alpha-1-acid glycoprotein 1 | ORM1 | 3 | 6 | 6 |
| Alpha-1-acid glycoprotein 2 | ORM2 | 3 | 6 | 9 |
| Alpha-1-antitrypsin | SERPINA1 | 4 | 4 | 8 |
| Alpha-1B-glycoprotein | A1BG | 2 | 2 | 9 |
| Alpha-1-microglobulin/bikunin precursor | AMBP | 2 | 7 | 4 |
| Alpha-2-HS-glycoprotein | AHSG | 3 | 2 | 14 |
| Alpha-2-macroglobulin | A2M | | 1 | |
| Angiotensinogen | AGT | 1 | 2 | 2 |
| Apolipoprotein A | APOA4 | 6 | 2 | 10 |
| Apolipoprotein A-I | APOA1 | 3 | 1 | 10 |
| Apolipoprotein A-II | APOA2 | 5 | 2 | 9 |
| Apolipoprotein C-I | APOC1 | 6 | 3 | 6 |
| Apolipoprotein C-II | APOC2 | 5 | 4 | 4 |

| | | | | |
|---|---|---|---|---|
| Apolipoprotein C-III | APOC3 | 6 | 3 | 6 |
| Apolipoprotein D | APOD | 5 | 3 | 9 |
| Apolipoprotein F | APOF | 2 | 1 | 11 |
| Apolipoprotein M | APOM | 4 | 4 | 4 |
| Beta-2-glycoprotein 1 | APOH | | 3 | 2 |
| Beta-2-microglobulin | B2M | | 3 | 1 |
| Ceruloplasmin | CP | 1 | 1 | |
| Clusterin | CLU | 1 | | |
| Coiled-coil domain-containing protein 185 | CCDC185 | | 1 | |
| Dual oxidase 1 | DUOX1 | | 3 | 6 |
| E3 ubiquitin-protein ligase TRIM35 | TRIM35 | | 1 | |
| Fibrinogen alpha chain | FGA | 2 | 4 | 6 |
| Galectin-10 | CLC | | 1 | 1 |
| GRIP and coiled-coil domain-containing protein 2 | GCC2 | 5 | 2 | 7 |
| Haptoglobin | HP | | 1 | |
| Haptoglobin-related protein | HPR | | 1 | |
| Hemoglobin subunit alpha | HBA1 | 2 | 2 | 2 |
| Hemoglobin subunit beta | HBB | 4 | 2 | 10 |
| Hemoglobin subunit delta | HBD | 1 | | 2 |
| Hemoglobin subunit gamma-1 | HBG1 | 3 | 1 | 4 |
| Hemoglobin subunit gamma-2 | HBG2 | 1 | | 1 |
| Hemopexin | HPX | 1 | 3 | 11 |
| Histidine-rich glycoprotein | HRG | 3 | 3 | 2 |
| Inter-alpha-trypsin inhibitor heavy chain H4 | ITIH4 | | 1 | |
| Kallikrein-11 | KLK11 | 2 | 1 | 1 |
| Keratin_ type I cuticular Ha2 | KRT32 | 1 | | |
| Keratin_ type I cytoskeletal 16 | KRT16 | | 1 | |
| Keratin_ type I cytoskeletal 17 | KRT17 | | 1 | |
| Keratin_ type I cytoskeletal 19 | KRT19 | | 1 | |
| Keratin_ type II cytoskeletal 1 | KRT1 | 3 | 1 | 2 |
| Keratin_ type II cytoskeletal 1b | KRT77 | 1 | | |
| Keratin_ type II cytoskeletal 2 epidermal | KRT2 | 1 | | |
| Keratin_ type II cytoskeletal 6C | KRT6C | 1 | | |
| Kinesin-like protein KIF21B | KIF21B | | 1 | 2 |
| Kininogen-1 | KNG1 | 3 | 2 | 3 |
| Leucine-rich alpha-2-glycoprotein | LRG1 | 3 | 7 | 6 |
| Lumican | LUM | 4 | 1 | 6 |
| Nuclear mitotic apparatus protein 1 | NUMA1 | 6 | | 9 |
| Plasma protease C1 inhibitor | SERPING1 | | 1 | |
| Platelet basic protein | PPBP | | 1 | 8 |
| Prothrombin | F2 | 3 | 4 | 7 |
| Putative alpha-1-antitrypsin-related protein | SERPINA2 | 1 | 1 | 3 |
| Rab-interacting lysosomal protein | RILP | 1 | | |

| | | | | |
|---|---|---|---|---|
| Reticulocalbin-1 | RCN1 | 1 | | 2 |
| Retinol-binding protein 4 | RBP4 | 5 | 3 | 10 |
| Serine/threonine-protein kinase 3 | STK3 | | 1 | |
| Serotransferrin | TF | | 4 | 12 |
| Serum amyloid A-1 | SAA1 | 4 | 4 | 3 |
| Serum amyloid A-2 protein | SAA2 | 1 | 1 | |
| Serum amyloid A-4 protein | SAA4 | 1 | | |
| Squamous cell carcinoma antigen recognized by T-cells 3 | SART3 | | 1 | 3 |
| Tetranectin | CLEC3B | 1 | 5 | 5 |
| Transmembrane protease serine 13 | TMPRSS13 | 2 | 2 | 9 |
| Transthyretin | TTR | 5 | 3 | 7 |
| Trichoplein keratin filament-binding protein | TCHP | 5 | 1 | 8 |
| Trypsin-1 | PRSS1 | 1 | | 1 |
| Trypsin-3 | PRSS3 | 3 | | 2 |
| Vitamin D-binding protein | GC | 3 | 3 | 7 |
| Vitronectin | VTN | | 1 | |
| Zinc finger HIT domain-containing protein 2 | ZNHIT2 | 2 | 3 | 5 |
| Zinc-alpha-2-glycoprotein | AZGP1 | 3 | 3 | 9 |

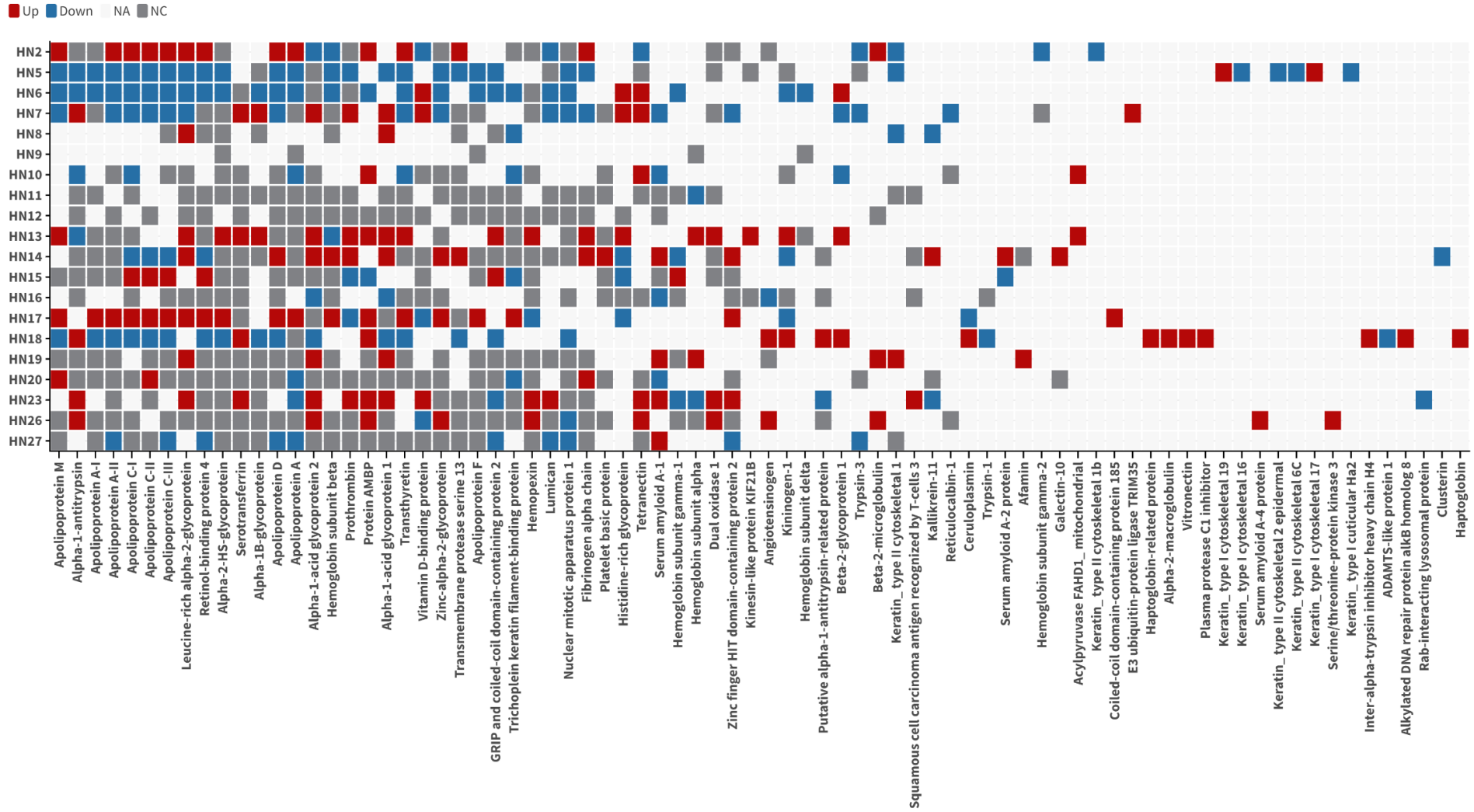



Figure 18. Top 78 differentially expressed proteins across the 20 samples

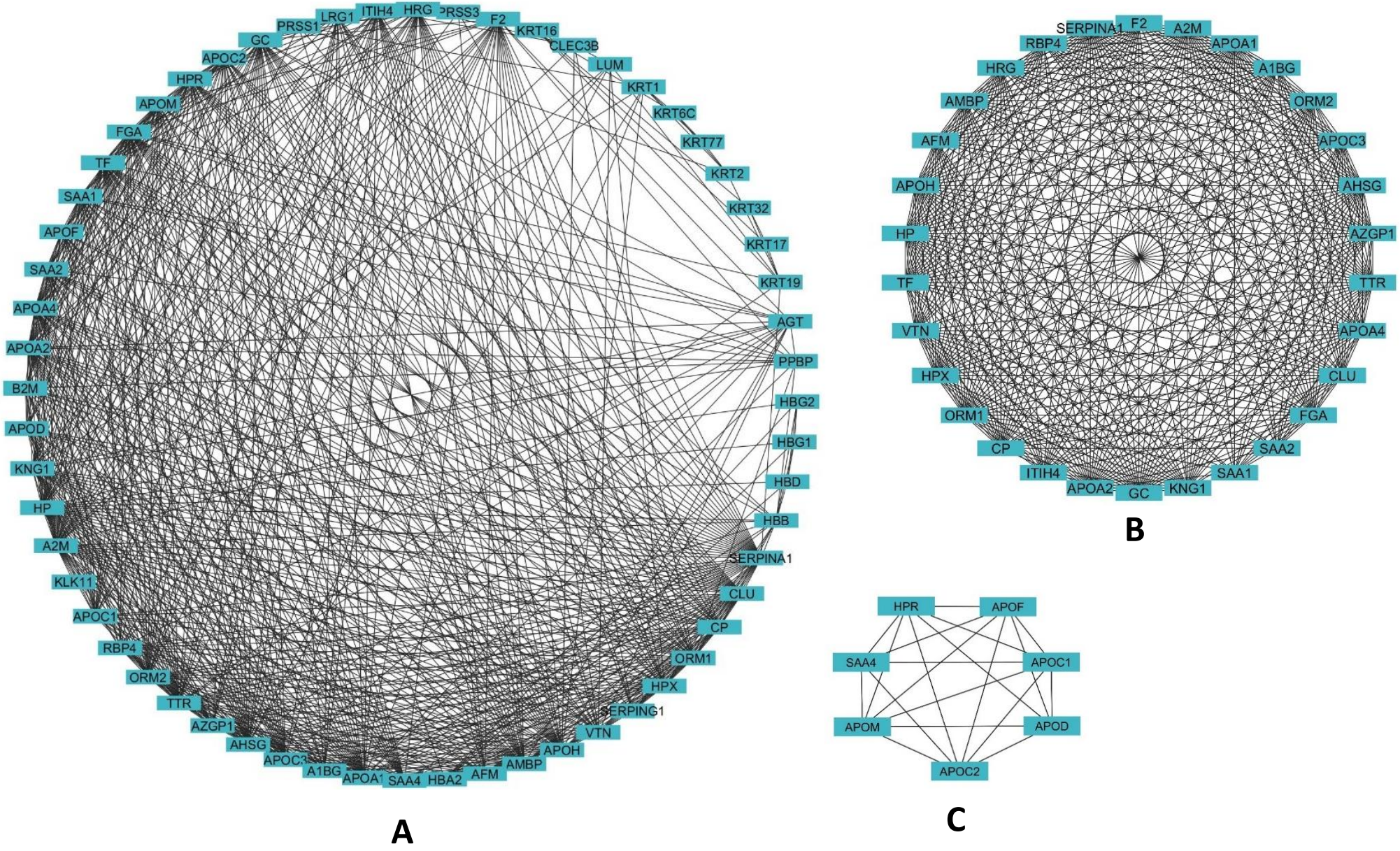



Figure 19. Protein-protein interaction (PPI) network construction (A) PPI network constructed with degree cutoff = 2, score cutoff = 0.4 (60 nodes and 690 edges. (B) Module 1 with MCODE score of 26.83 (30 nodes and 389 edges). (C) Module 2 with MCODE score of 6.67 (7 nodes and 20 edges).

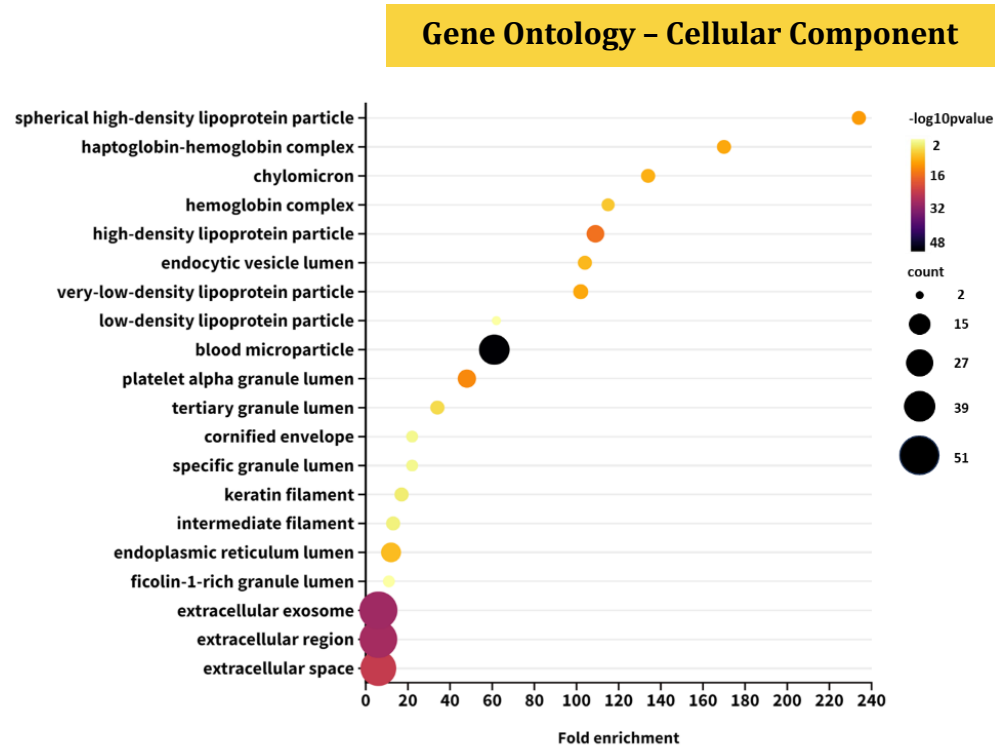



Figure 20 (a). Gene ontology (GO) analysis - Cellular Component

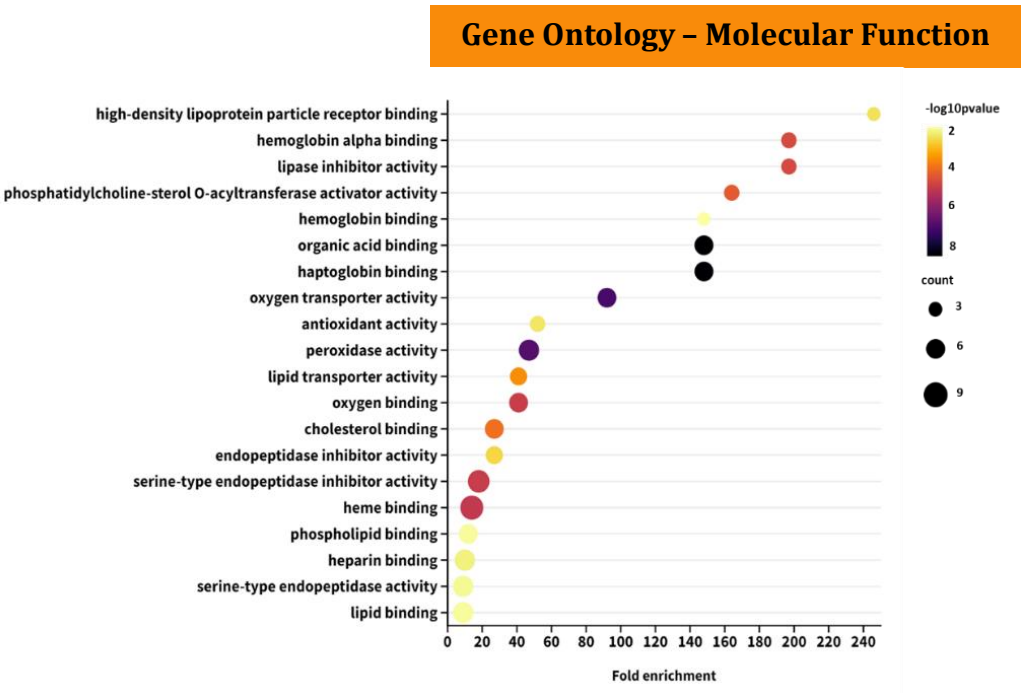



Figure 20 (b). Gene ontology (GO) analysis – Molecular Function

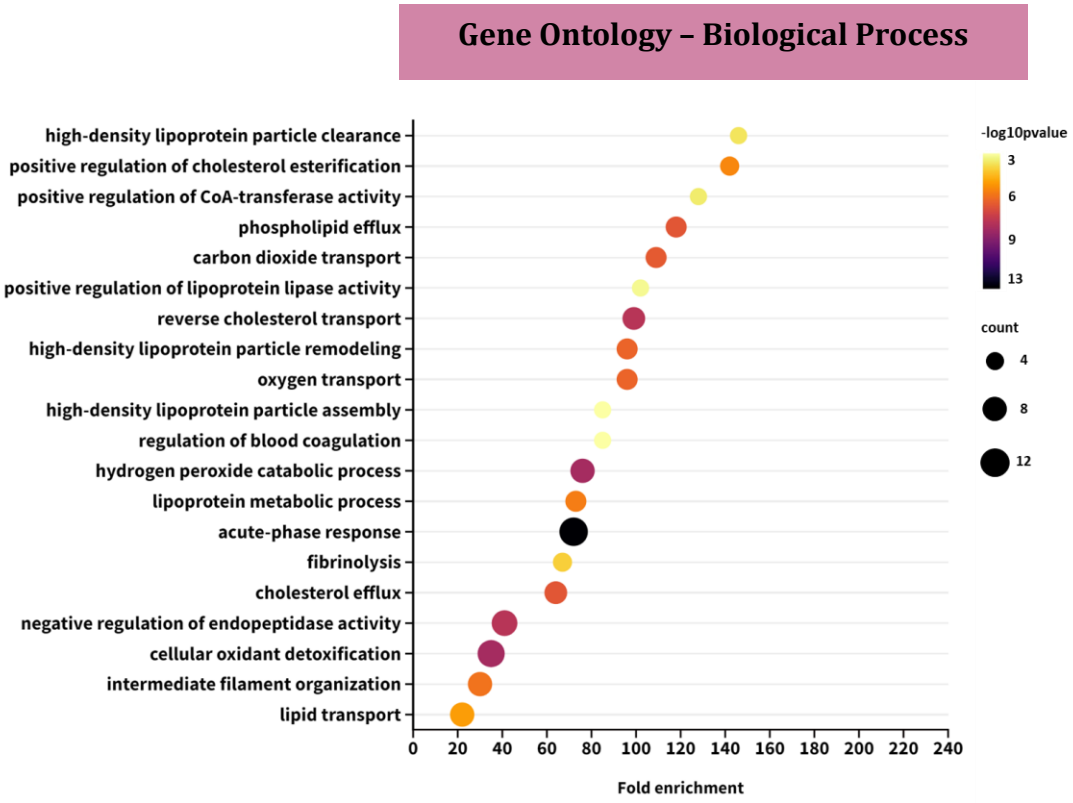



Figure 20 (c). Gene ontology (GO) analysis – Biological Process

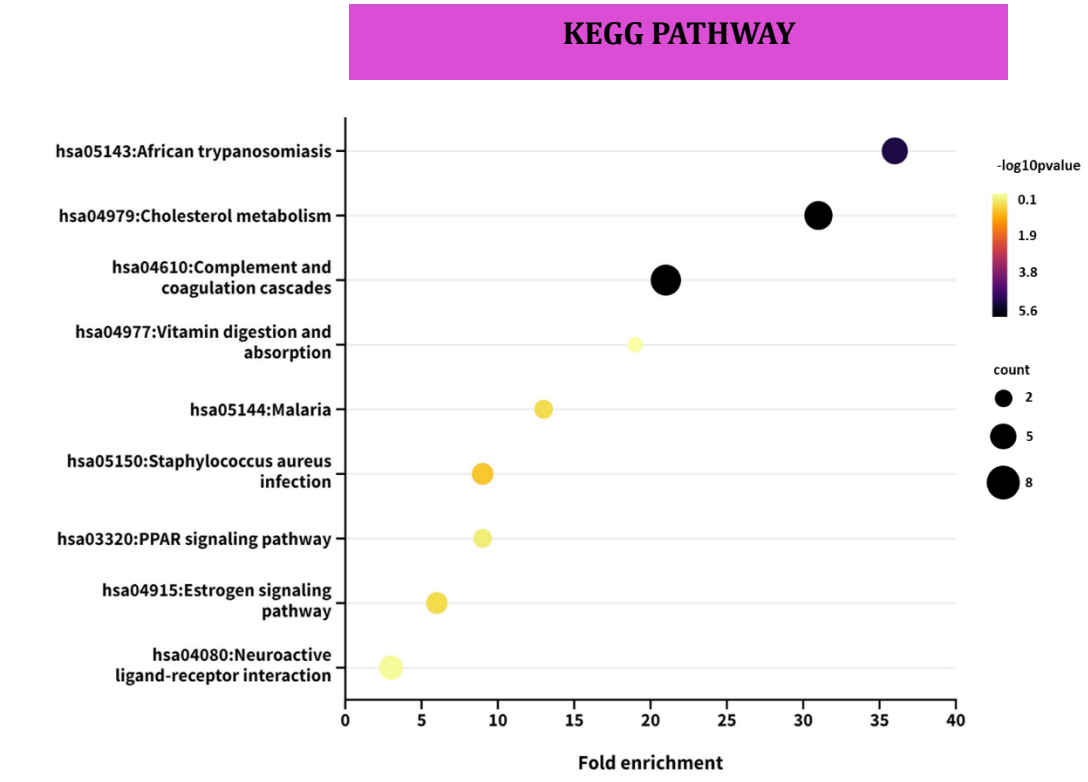
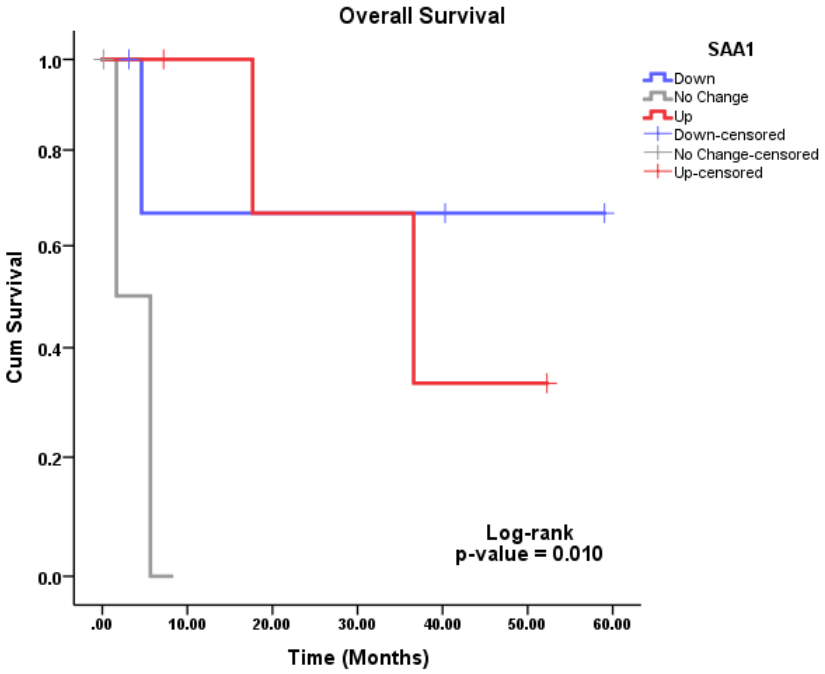



Figure 20 (d). KEGG Pathway analysis

**Kaplan-Meier estimates and log-rank test**

| SAA1 | OS rate (%) | PFS rate (%) |
|---|---|---|
| Down | 66 | 66 |
| No Change | 0 | 0 |
| Up | 33 | 50 |



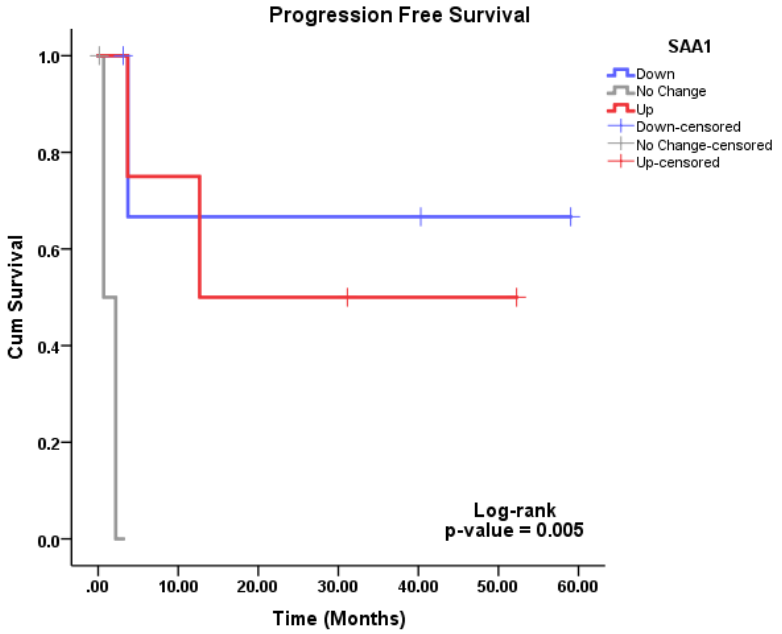
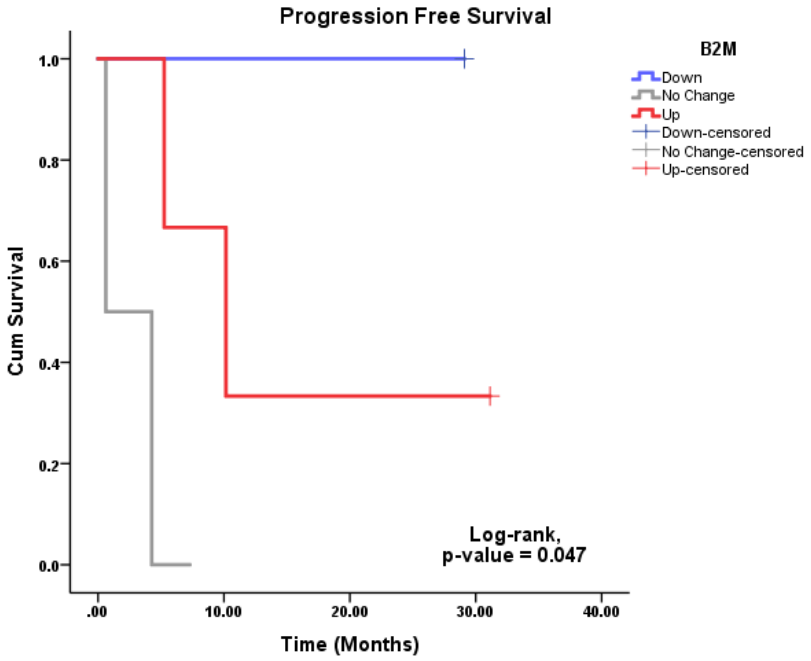




Figure 21. Kaplan-Meier plot and Log-rank test for OS and PFS on SAA1 protein

| β2M | PFS rate (%) |
|---|---|
| Down | 100 |
| No Change | 0 |
| Up | 33 |

| HBB | PFS rate (%) |
|---|---|
| Down | 0 |
| No Change | 45 |
| Up | 0 |



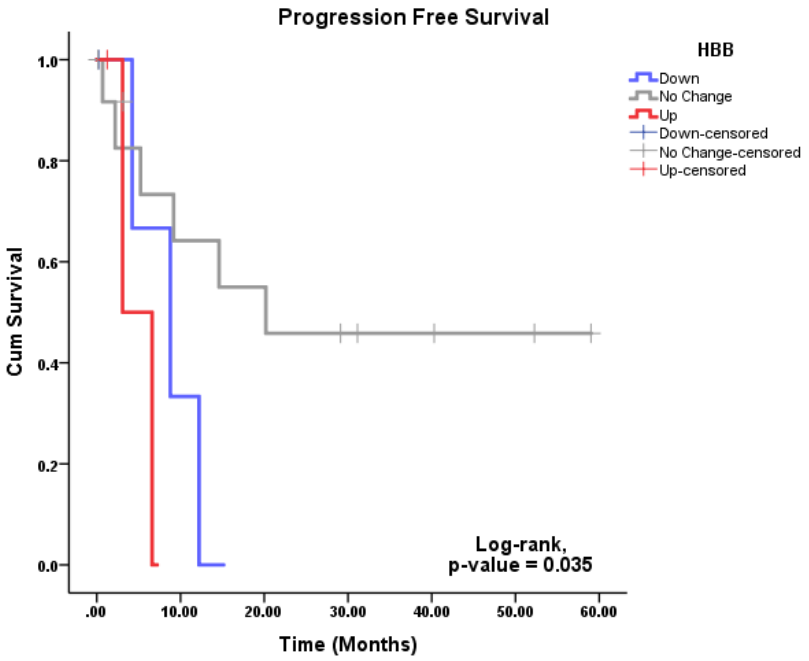




Figure 22. Kaplan-Meier plot and Log-rank test for PFS on β2M protein and HBB protein

| ASHG | PFS rate (%) |
|---|---|
| Down | 0 |
| No Change | 38 |
| Up | 0 |

| Protein | p-value | Hazard Ratio | Lower CI | Upper CI |
|---|---|---|---|---|
| AHSG | 0.048 | 1.938 | 1.006 | 3.733 |

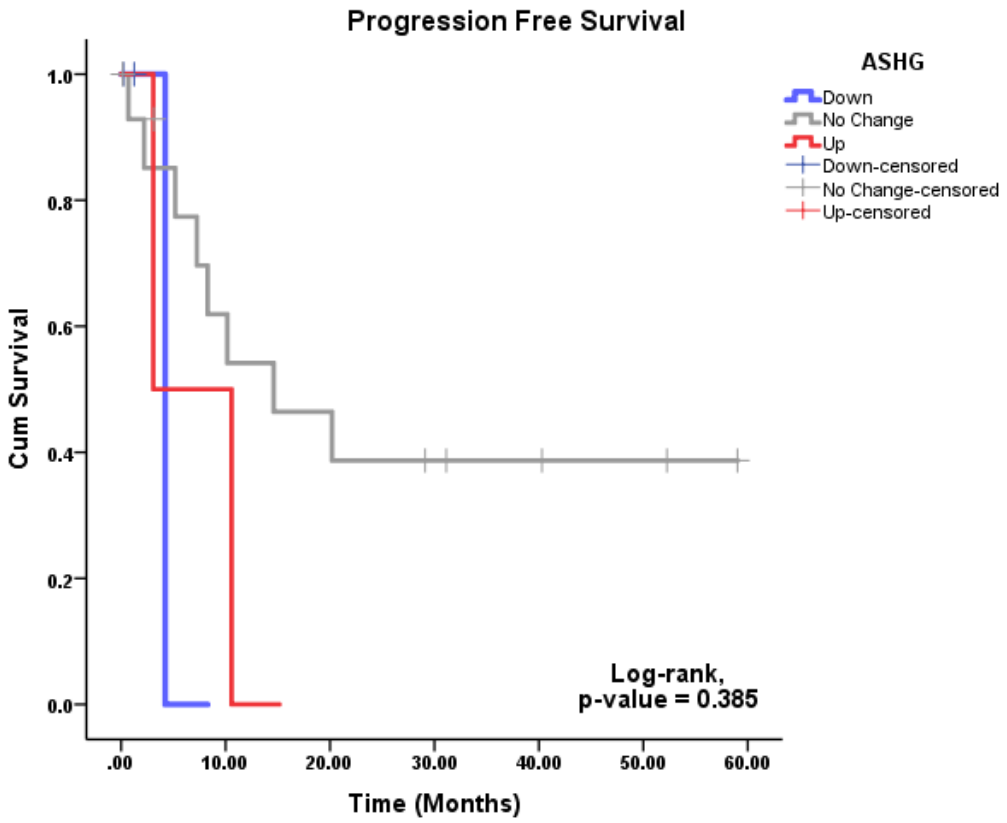



Figure 23. Hazard Ratio, Kaplan-Meier plot and Log-rank test for PFS on ASHG protein

The clinicopathological and lifestyle factors for the ten hypopharyngeal cancer patients is shown in Table 24. The factors included age, stage, smoking habits, alcohol, betelnut and tobacco consumption, treatment regimen and response. Cisplatin was mainly administered for chemotherapy.

Table 24. Characteristics of ten hypopharyngeal cancer patients in the study

| Variables | | Number of s (N=10) |
|---|---|---|
| Age (in years) | Range | 44 - 71 |
| | Mean ± SD | 53 ± 2.55 |
| Stage | I | 0 |
| | II | 0 |
| | III | 4 |
| | IV | 5 |
| | NA | 1 |
| Smoking | Yes | 8 |
| | No | 2 |
| Alcohol | Yes | 5 |
| | No | 5 |
| Tobacco | Yes | 6 |
| | No | 4 |
| Betelnut | Yes | 8 |
| | No | 0 |
| | NA | 2 |
| Induction Chemotherapy | Cisplatin & Paclitaxel | 4 |
| | Cisplatin & Fluorouracil | 1 |
| | Cisplatin, Docetaxel and Fluorouracil | 1 |
| | None | 2 |
| | NA | 2 |
| Concurrent Chemotherapy | Cisplatin | 4 |
| | Paclitaxel | 1 |
| | None | 2 |
| | NA | 3 |
| Radiotherapy | Range | 33 - 66 |
| Treatment Response | Complete Response | 2 |
| | Partial Response | 3 |
| | Progressive Disease | 1 |
| | Incomplete treatment | 2 |
| | Lost to follow up | 2 |

Whole Exome Sequencing (WES) was performed on tumour tissues and matched blood DNA from ten hypopharyngeal cancer samples. Our data showed that C/A transversion and C/T transitions are larger than C/G, T/A transversions and T/G transition (Figure 24). About 56,88,669 variants were identified collectively in all the ten samples (Figure 25). Exonic variants accounted for 1.10 % of all the variants with 62,403 variantes spanning across 12,122 genes (Figures 25 & 26). The different types of alterations observed among the exonic variants are provided in Figures 26 & 27, where 29, 279 non-synonymous Single Nucleotide Variations (SNV) in 8333 genes were found. Figure 28 shows the total variants identified in each sample where T7 has the highest mutational burden.

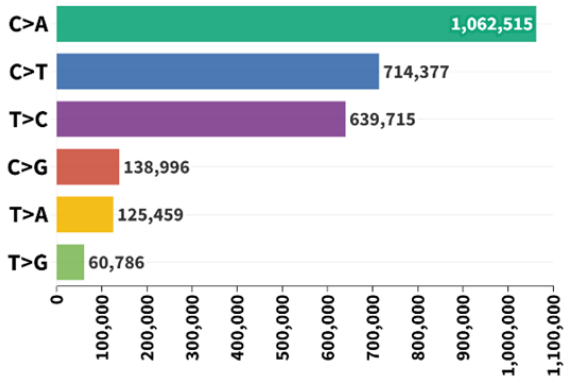



Figure 24. Number of transitions and transversion found in ten hypopharyngeal cancer

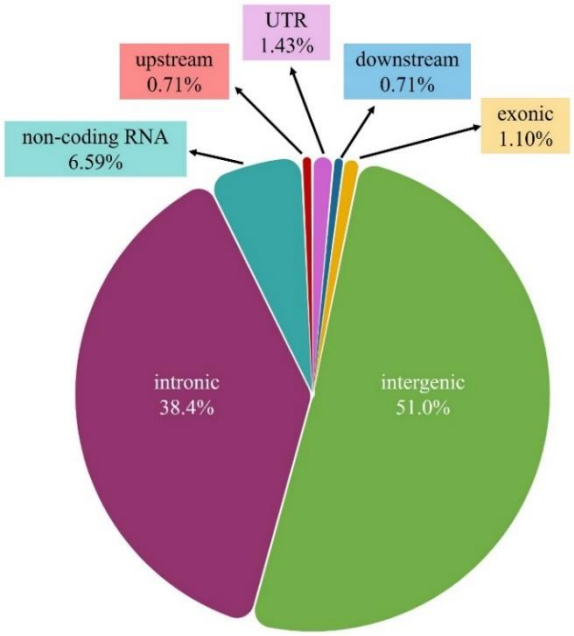



Figure 25. The proportion of variants in ten hypopharyngeal cancer

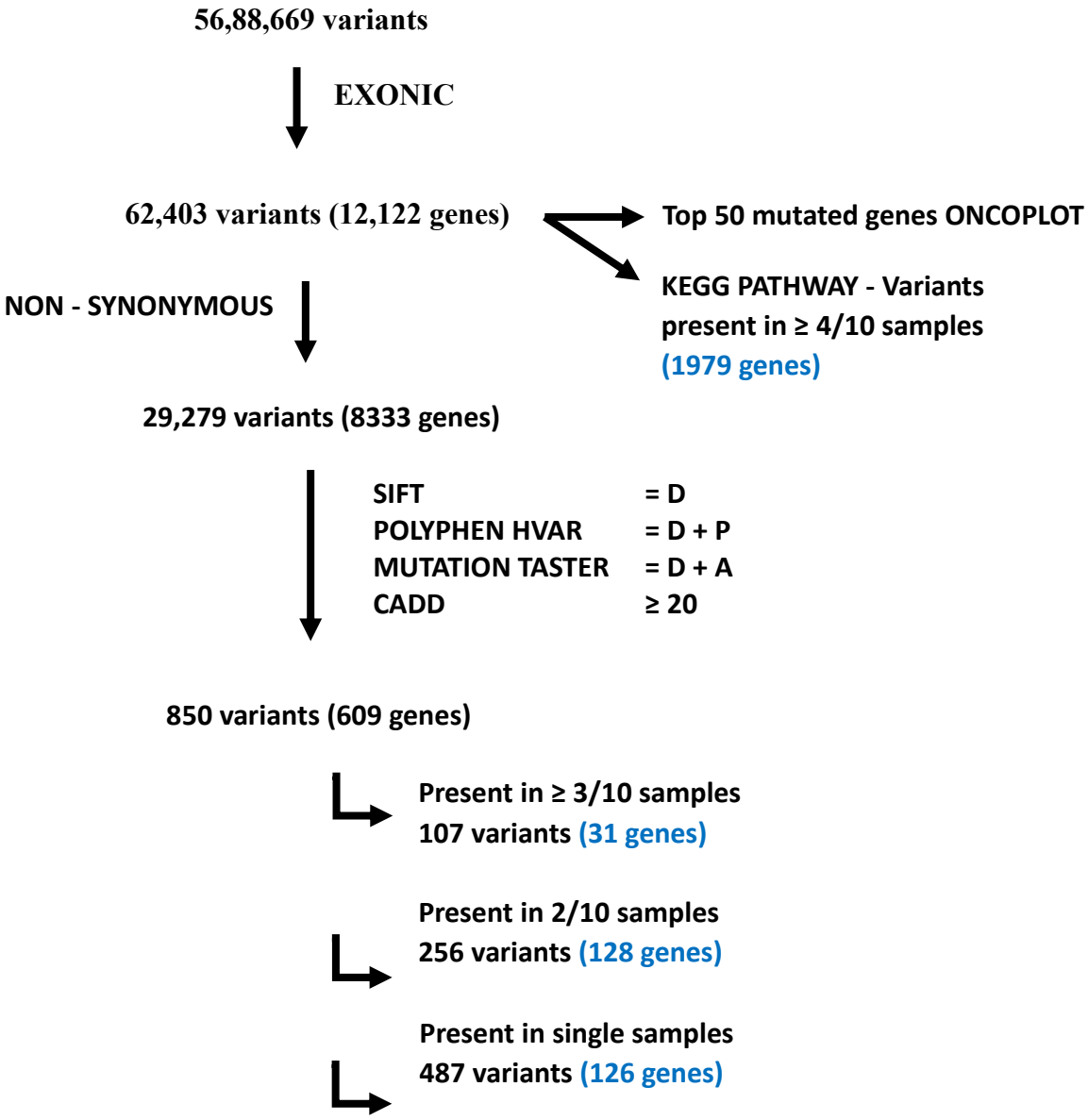



Figure 26. Flowchart illustrating the steps for filtering variants data

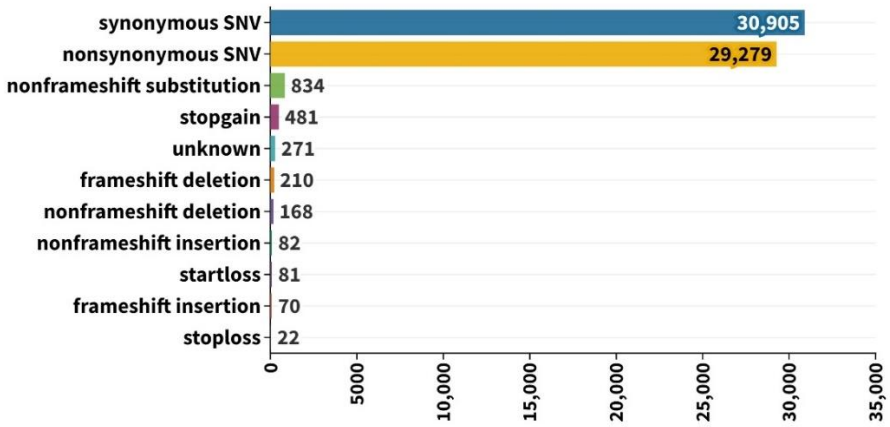



Figure 27. Number of different variants in exonic regions in ten hypopharyngeal cancer

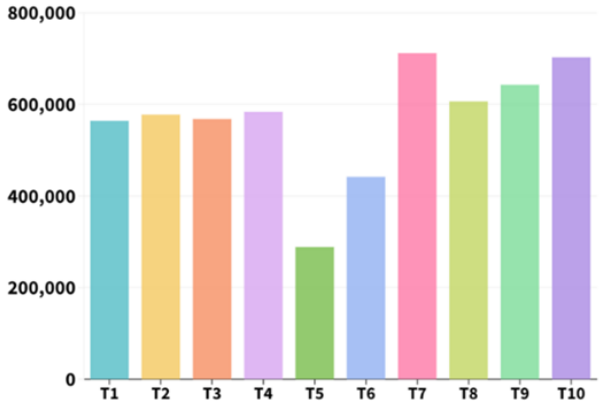



Figure 28. Number of total variants per sample in ten hypopharyngeal cancer

Figure 29 shows the top 50 mutated genes across the ten hypopharyngeal cancers. The top four genes namely *FSIP2, HMCN2, MUC3A* and *ZNF705E* were found to be altered in all the ten samples. The majority of the alterations were multi hit events across the samples. The *ZNF705E* gene was mostly altered by nonsense mutation alone. Also, patient T6 has a nonsense mutation alone in TP53 gene. Patient T10 harbours variants in all the top 50 genes while T9, T8 and T7 have 49 out of 50 genes alterations. Genomic changes in *ATP5F1A, CCDC187* and *FOXD4L4* were found in 9 out of 10 samples, with the exception of patient T3. The genes that were observed to be altered in 9 out of the 10 samples were *LRP2, MUC12, MUC16, NEB* and *TTN*.

From the 62,403 variants in 12,122 genes, variants that were present in at least 4 out of 10 samples were filtered out for KEGG pathway analysis. Table 25 shows the top 12 pathways that are significantly enriched in our samples. The genes that are involved in pathways related to cancer such as the ECM receptor interaction, Human Papillomavirus, ABC transporters, Lung Cancer, Complement and Coagulation Cascades and P13K-AKT pathways are depicted in Figures 30 to 32. Furthermore, 850 variants from 609 genes were predicted to be deleterious using SIFT, Polyphen, Mutation Taster and CADD. From the 850 variants, those that are present in three or more samples were filtered out and plotted in Figure 33 along with lifestyle habits and treatment and their response (Table 26). The gene *NT5C3A* was found to be the most frequently mutated in all the ten samples (7/10), followed by *MTMR4* and *AP1G2* observed in 5/10 samples. *AZIN2, IRX6, KMT2C, NUDT12, POP5* and *SHANK2* variants were present in 4/10 samples. Samples T7, T8, T9 and T10 were observed to be more mutated among the ten samples. There are 128 genes altered in 2 out of 10 samples (Table 27) and 126 genes altered in single sample (Table 28).

Table 25. KEGG Pathways significantly enriched in our samples

| **Pathway** | **No. of genes** | **p-value** |
|---|---|---|
| hsa04512:ECM-receptor interaction | 36 | 0.000 |
| hsa04814:Motor proteins | 45 | 0.000 |
| hsa04974:Protein digestion and absorption | 26 | 0.000 |
| hsa04510:Focal adhesion | 39 | 0.000 |
| hsa05165:Human papillomavirus infection | 49 | 0.001 |
| hsa04740:Olfactory transduction | 61 | 0.001 |
| hsa02010:ABC transporters | 12 | 0.002 |
| hsa05146:Amoebiasis | 20 | 0.002 |
| hsa05168:Herpes simplex virus 1 infection | 65 | 0.006 |
| hsa04610:Complement and coagulation cascades | 16 | 0.011 |
| hsa05222:Small cell lung cancer | 16 | 0.019681 |
| hsa04151:PI3K-Akt signalling pathway | 44 | 0.041091 |

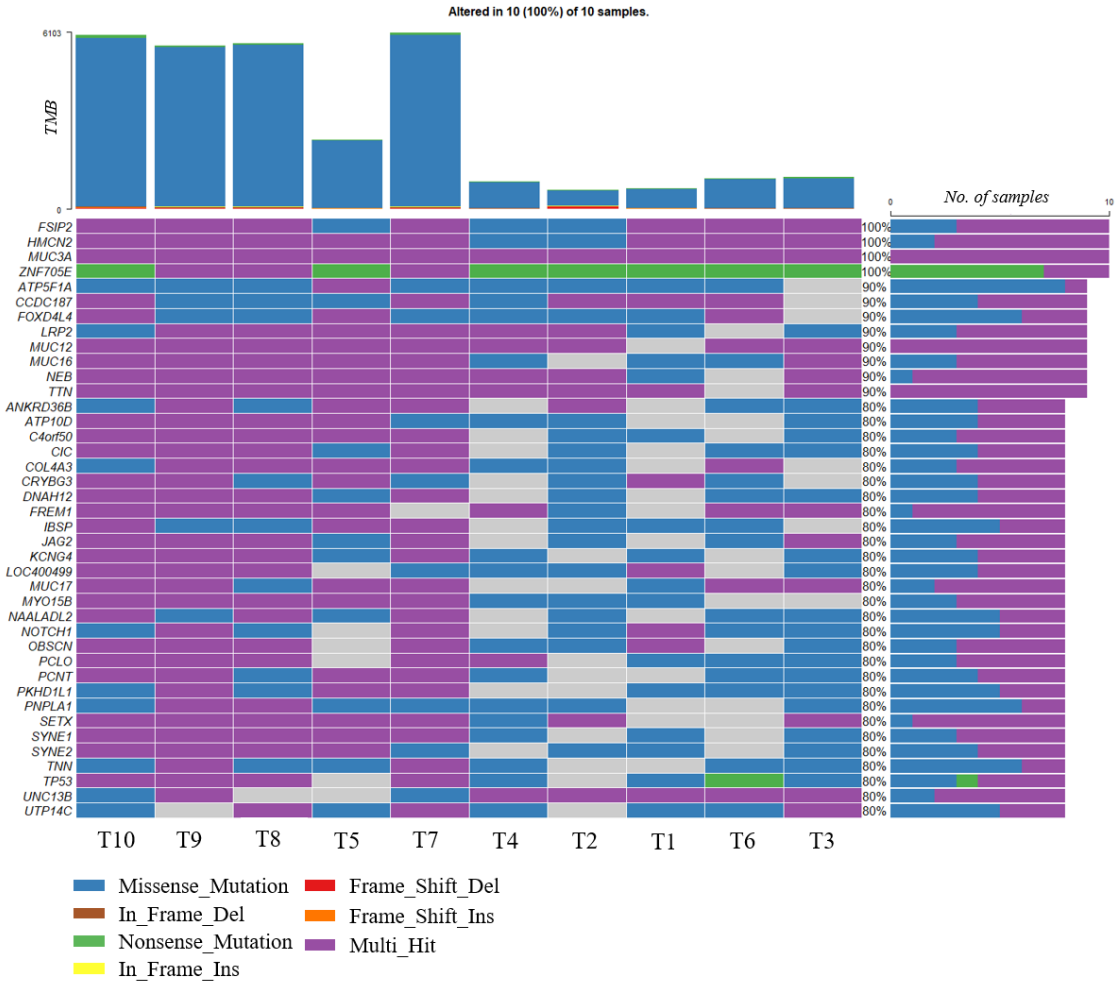



Figure 29. Top 50 altered genes across the ten hypopharyngeal samples

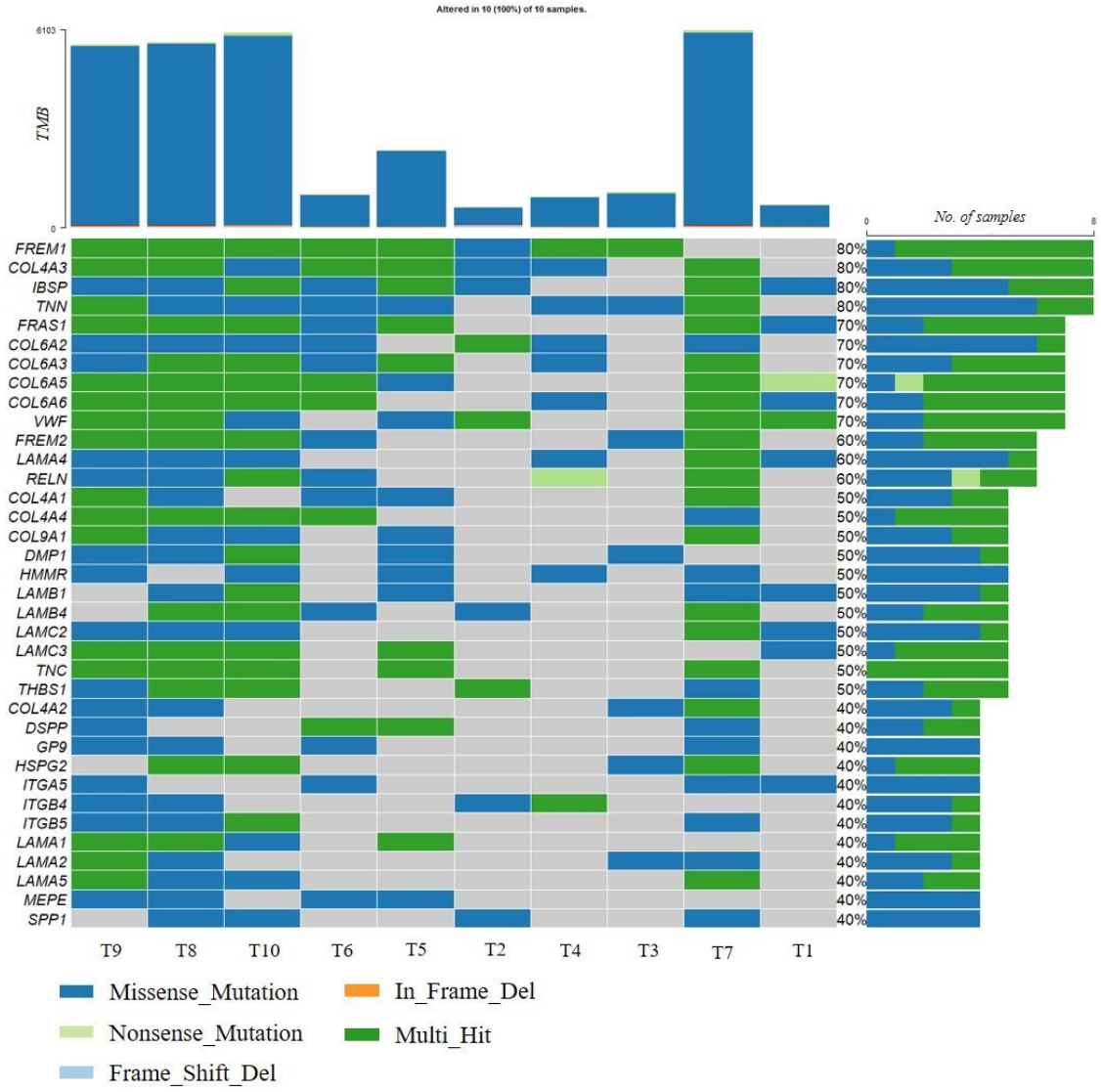



Figure 30 (a). Genes enriched in ECM Receptor Interaction pathway present in ≥ 4/10 samples

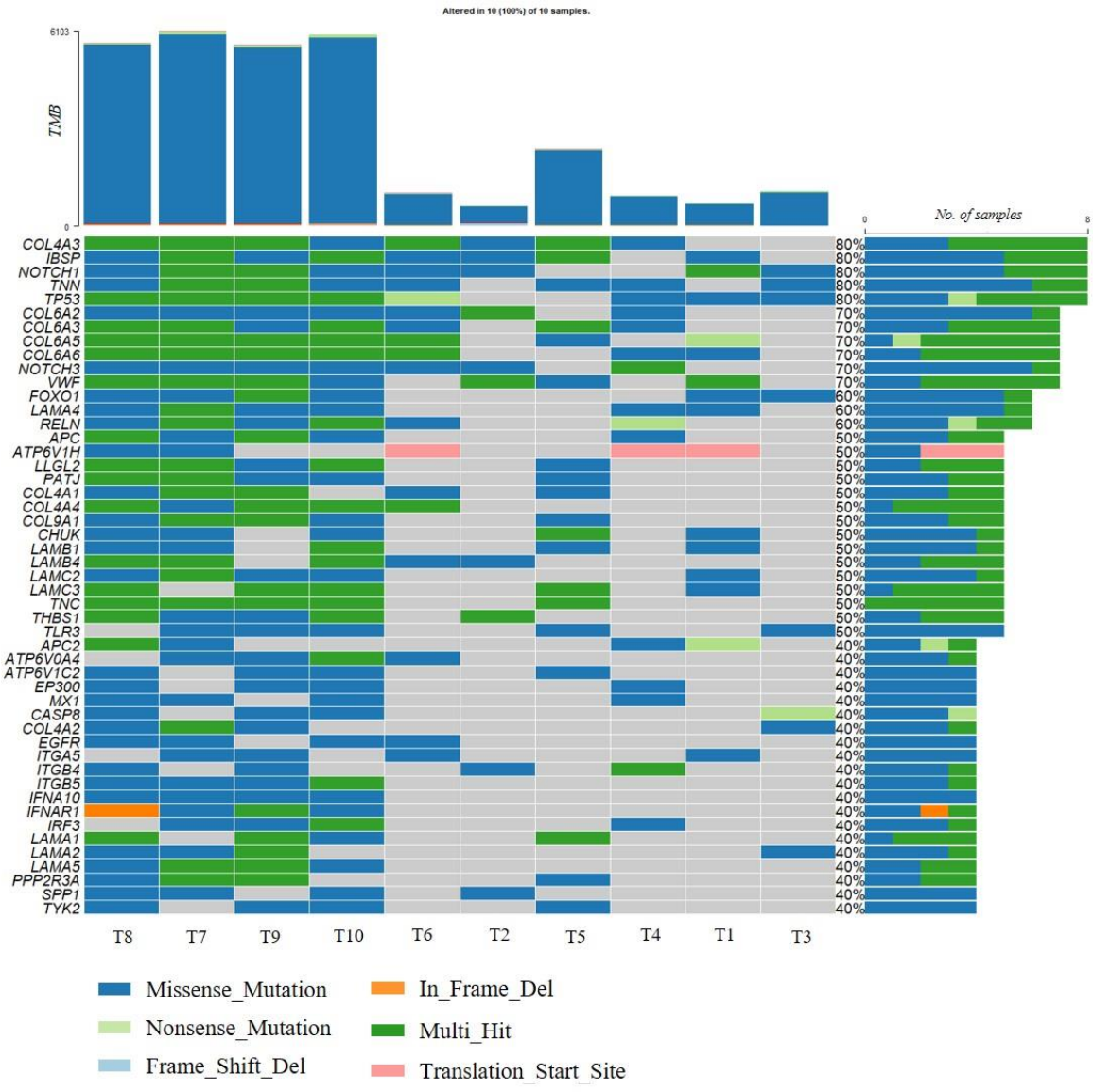



Figure 30 (b). Genes enriched in Human Papillomavirus infection pathway present in ≥ 4/10 samples

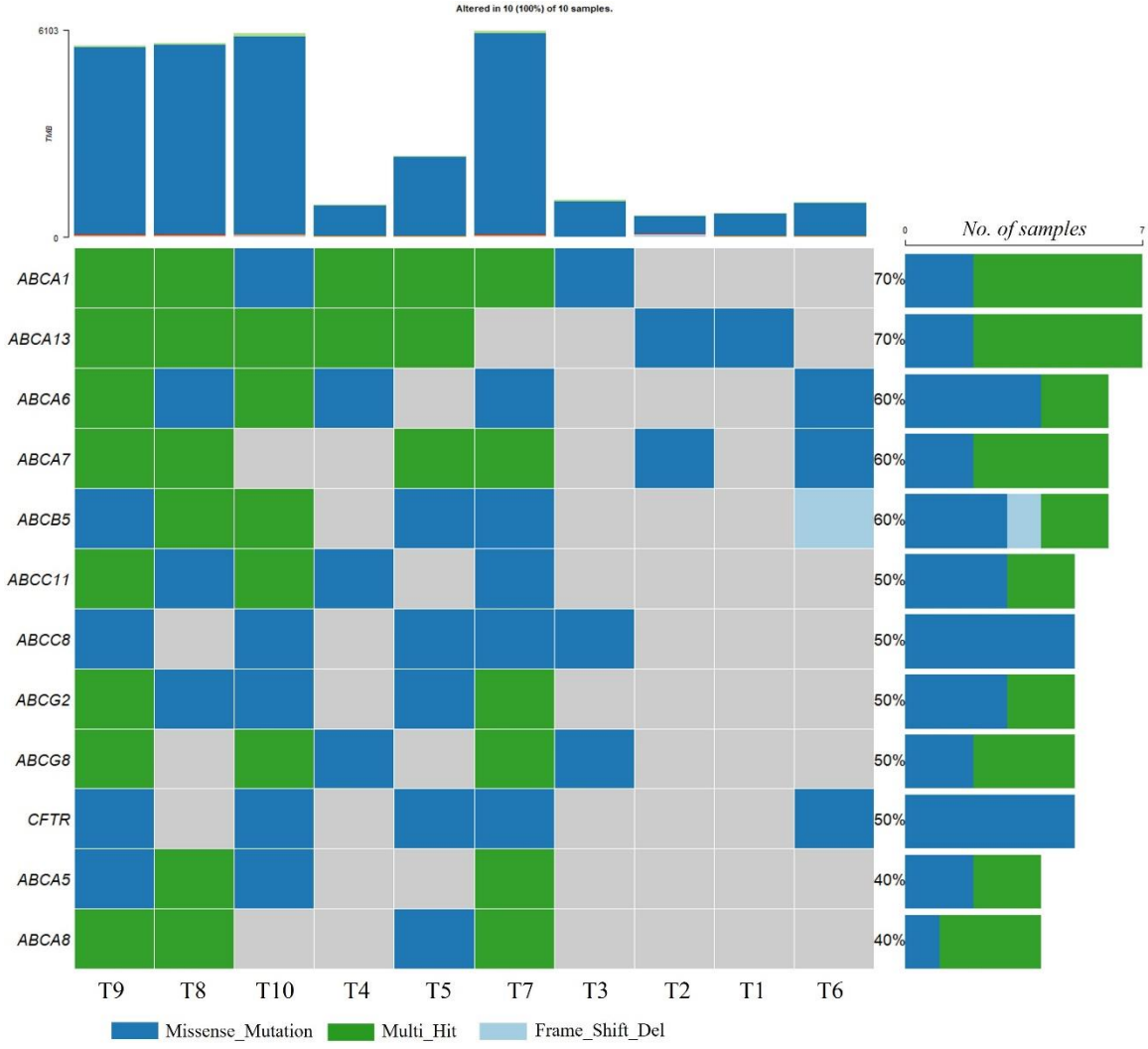



Figure 30 (c). Genes enriched in ABC transporters pathway present in ≥ 4/10 samples

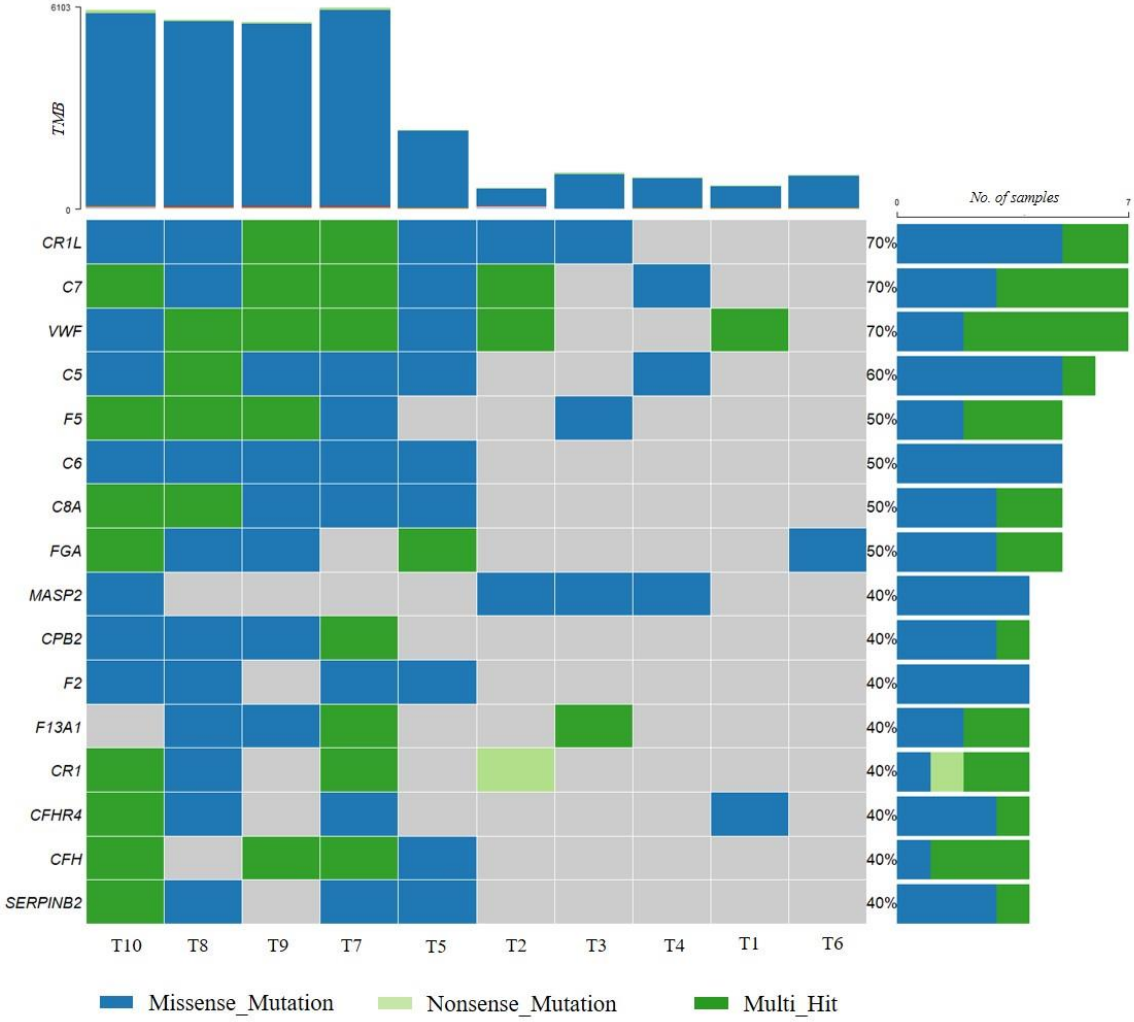



Figure 30 (d). Genes enriched in Complement and coagulation cascades pathway present in ≥ 4/10 samples

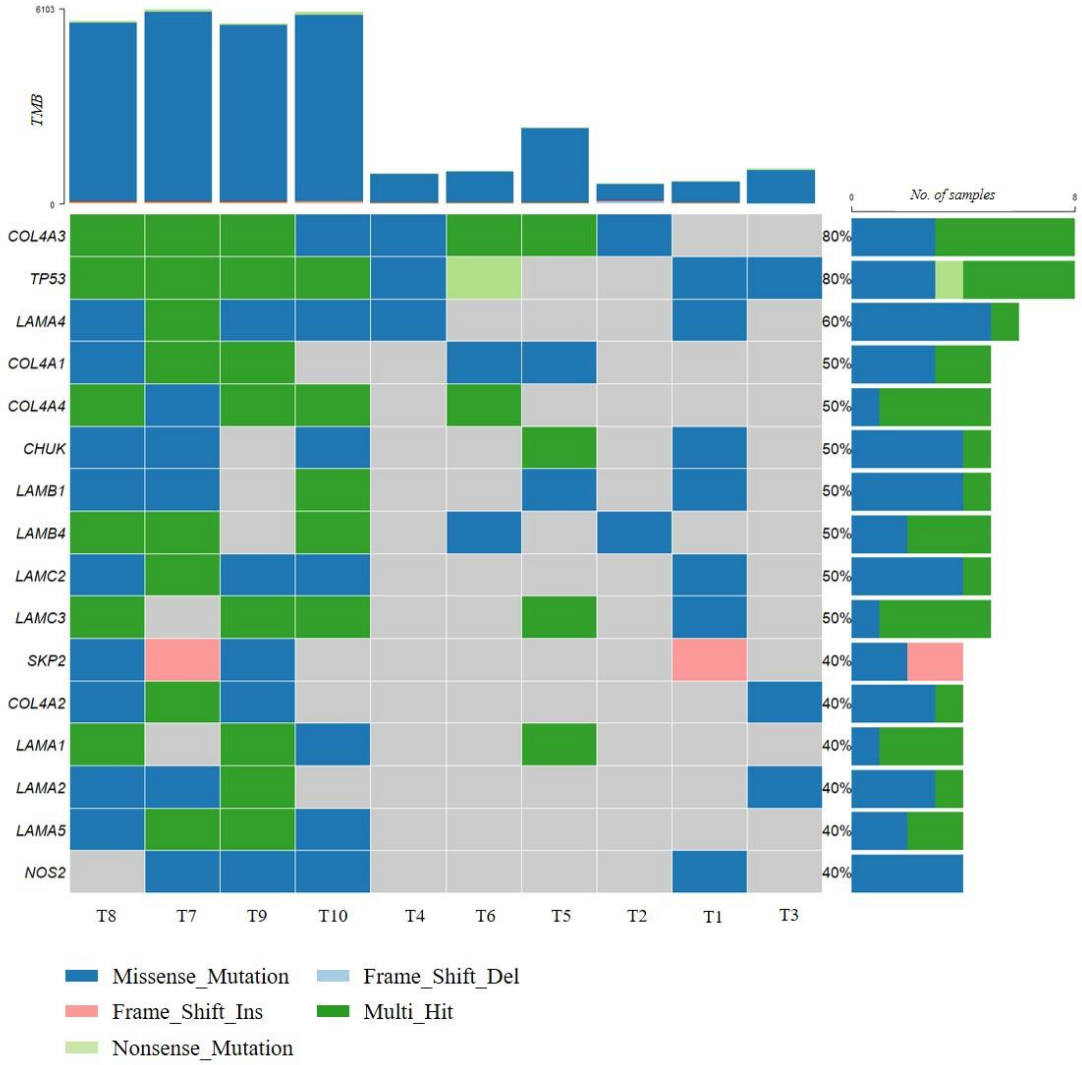



Figure 30 (e). Genes enriched in Lung cancer pathway present in ≥ 4/10 samples

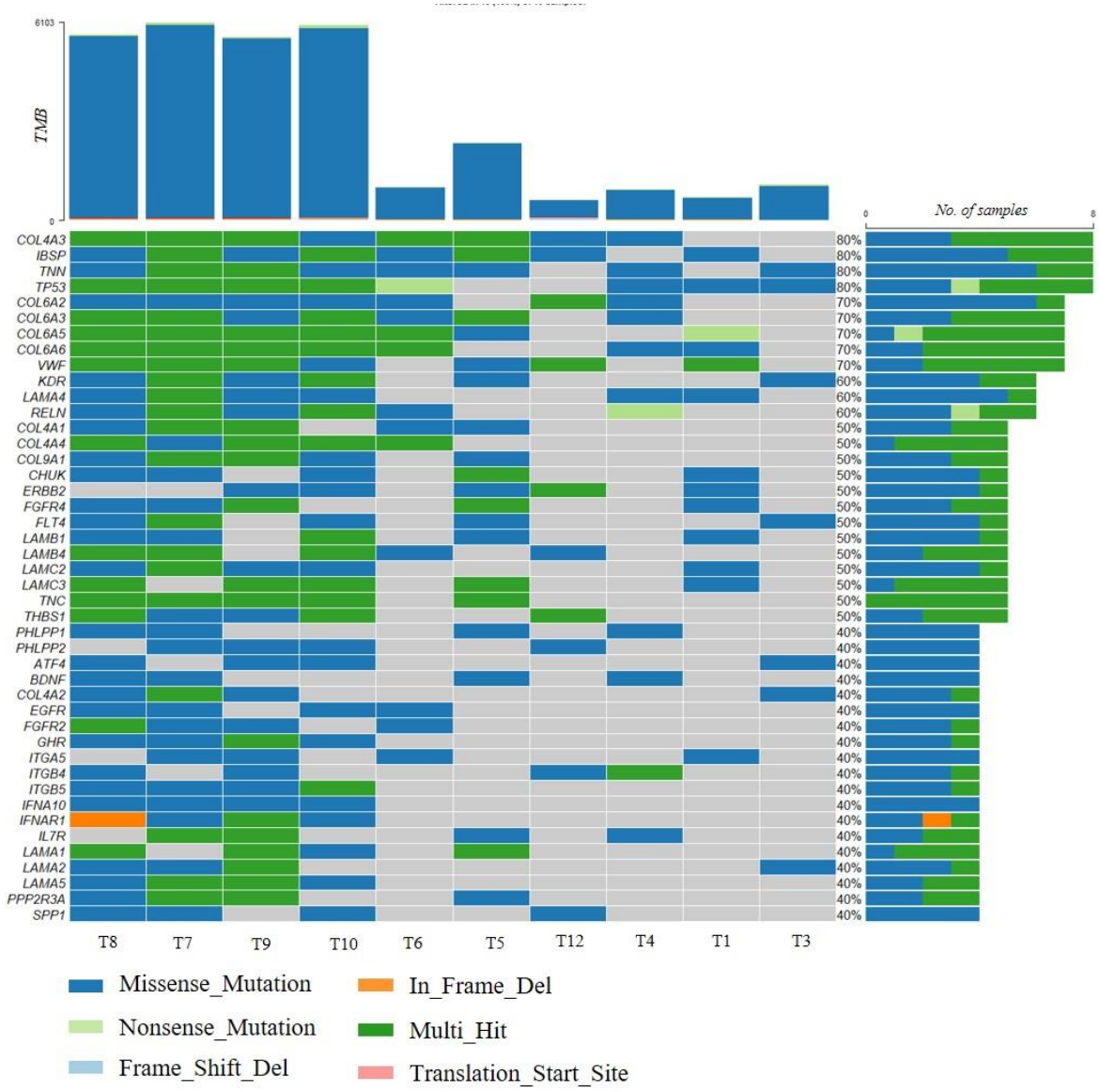



Figure 30 (f). Genes enriched in PI3K-Akt signalling pathway pathway present in ≥ 4/10 samples

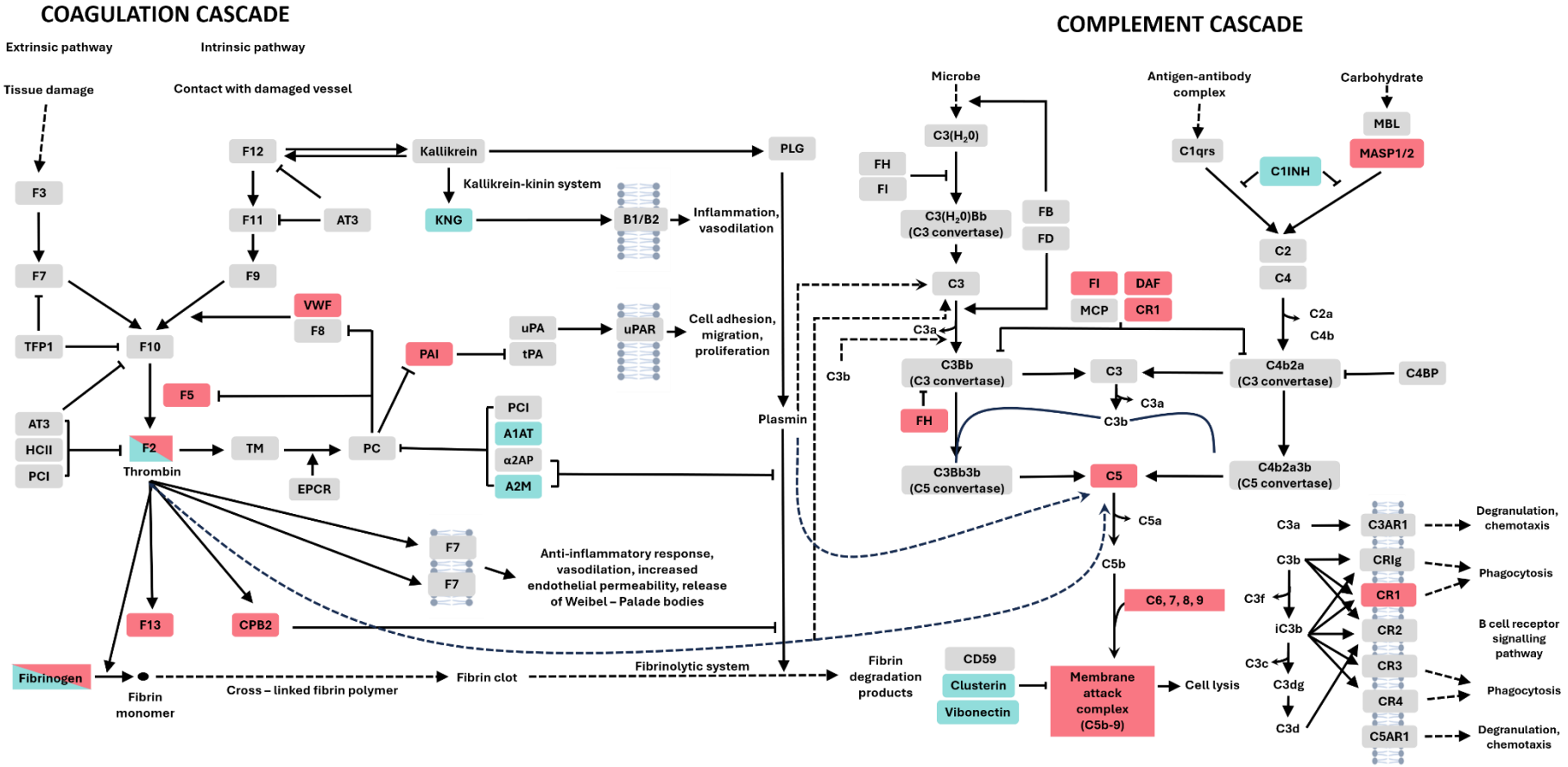
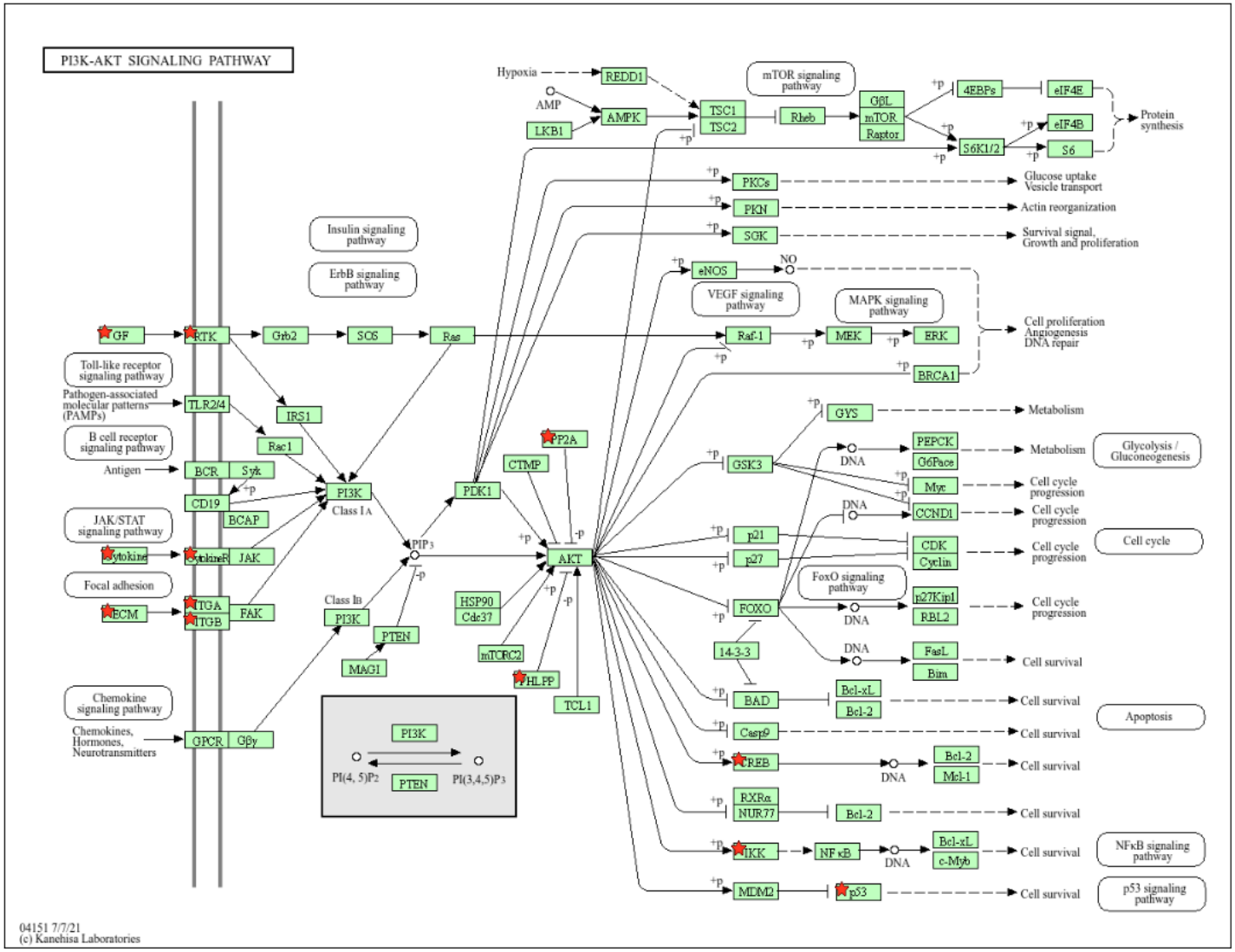



Figure 31. Genes mutated in ten hypopharyngeal samples (Highlighted pink) and genes differentially expressed in serum HNSCC samples (Highlighted blue)


PI3K-AKT SIGNALING PATHWAY
Hypoxia
REDD1
AMP
LKB1
AMPK
mTOR signaling pathway
TSC1
TSC2
Rheb
GβL
mTOR
Raptor
4EBPs
eIF4E
eIF4B
S6K1/2
S6
Protein synthesis
PKCs
Glucose uptake
Vesicle transport
PKN
Actin reorganization
SGK
Survival signal,
Growth and proliferation
Insulin signaling pathway
ErbB signaling pathway
eNOS
NO
VEGF signaling pathway
MAPK signaling pathway
GF
RTK
Grb2
SOS
Ras
Raf-1
MEK
ERK
Cell proliferation
Angiogenesis
DNA repair
BRCA1
Toll-like receptor signaling pathway
Pathogen-associated molecular patterns (PAMPs)
TLR2/4
IRS1
Rac1
B cell receptor signaling pathway
Antigen
BCR
Syk
CD19
BCAP
PI3K
Class IA
PDK1
PIP3
AKT
PP2A
CTMP
GYS
Metabolism
GSK3
PEPCK
G6Pase
Metabolism
Glycolysis / Gluconeogenesis
DNA
Myc
Cell cycle progression
CCND1
Cell cycle progression
JAK/STAT signaling pathway
Cytokine
CytokineR
JAK
p21
p27
CDK
Cyclin
Cell cycle progression
Cell cycle
Focal adhesion
ECM
ITGA
ITGB
FAK
FoxO signaling pathway
FOXO
p27Kip1
RBL2
Cell cycle progression
Class Ib
PI3K
PTEN
MAG1
HSP90
Cdc37
mTORC2
HLPP
TCL1
14-3-3
FasL
Bim
Cell survival
Chemokine signaling pathway
Chemokines, Hormones, Neurotransmitters
GPCR
Gβγ
BAD
Bcl-xL
Bcl-2
Cell survival
Apoptosis
Casp9
Cell survival
CREB
Bcl-2
Mcl-1
Cell survival
RXRα
NUR77
Bcl-2
Cell survival
IKK
NFκB
Bcl-xL
c-Myb
Cell survival
NFκB signaling pathway
MDM2
p53
Cell survival
p53 signaling pathway
PI3K
PI(4, 5)P2
PTEN
PI(3,4,5)P3
04151 7/7/21
(c) Kanehisa Laboratories

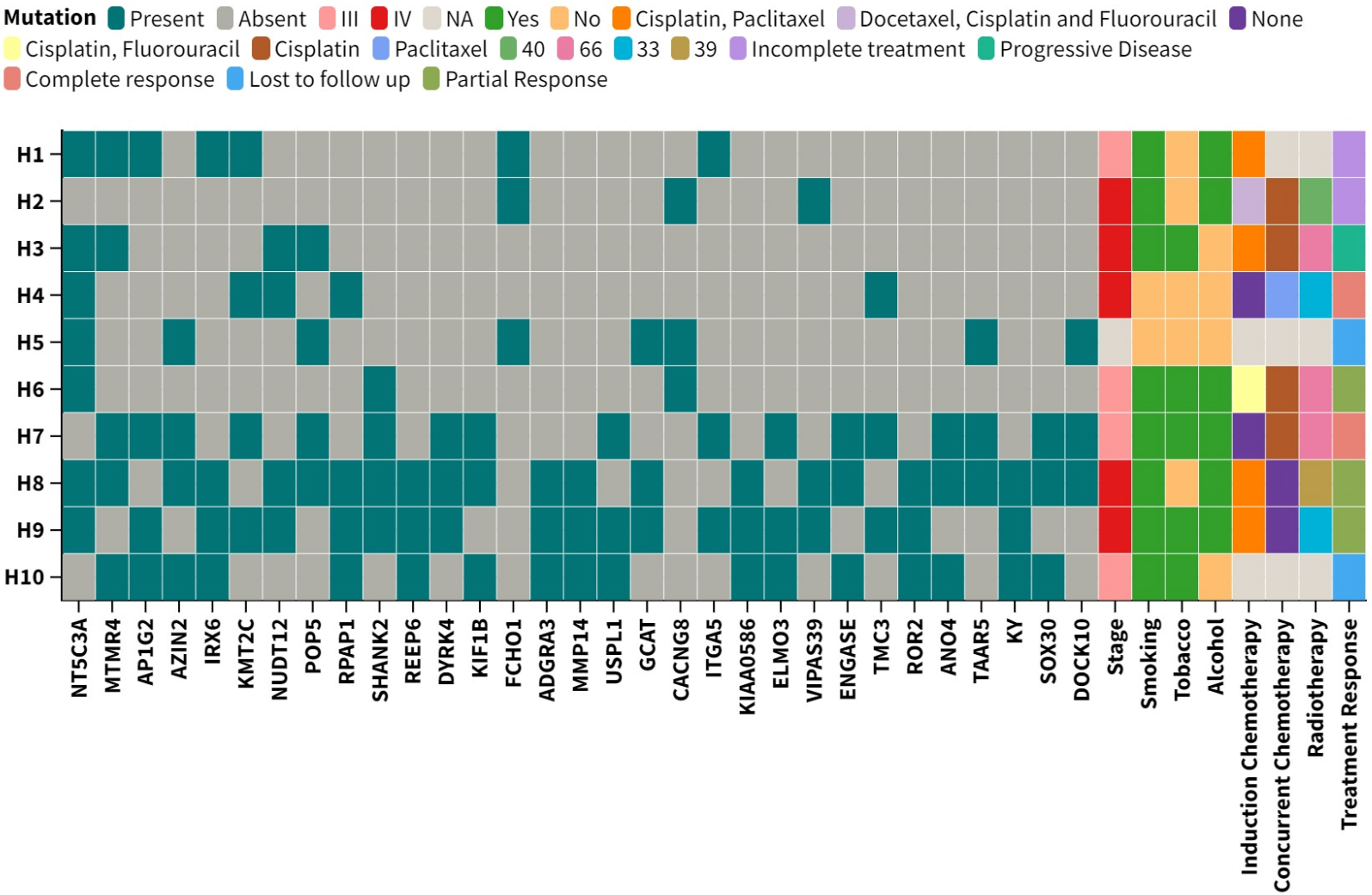



Figure 33. Illustrates 31 genes mutated across the samples (≥3/10 samples)

Table 26. Variants present in ≥ 3 out of 10 samples (31 genes)

| Gene | Chr | Start | End | Ref | Alt | rsID |
|---|---|---|---|---|---|---|
| **NT5C3A** | chr7 | 33014776 | 33014776 | T | C | rs79747830 |
| **MTMR4** | chr17 | 58504145 | 58504145 | G | A | rs183891875 |
| **AP1G2** | chr14 | 23559828 | 23559828 | C | T | rs74849041 |
| **AZIN2** | chr1 | 33096815 | 33096815 | G | T | rs16835244 |
| **IRX6** | chr16 | 55327738 | 55327738 | C | A | rs141194029 |
| **KMT2C** | chr7 | 152247975 | 152247975 | G | A | rs200598064 |
| **NUDT12** | chr5 | 103552319 | 103552319 | C | T | rs372191550 |
| **POP5** | chr12 | 120579809 | 120579809 | C | T | rs142190925 |
| **RPAP1** | chr15 | 41522937 | 41522937 | G | A | rs147507787 |
| **SHANK2** | chr11 | 70485855 | 70485855 | G | T | rs557121637 |
| **REEP6** | chr19 | 1495527 | 1495527 | G | C | rs76079175 |
| **DYRK4** | chr12 | 4599033 | 4599033 | A | G | rs3741927 |
| **KIF1B** | chr1 | 10337509 | 10337509 | A | G | rs2297881 |
| **FCHO1** | chr19 | 17787765 | 17787765 | G | A | rs773887321 |
| **ADGRA3** | chr4 | 22392616 | 22392616 | T | G | rs117922332 |
| **MMP14** | chr14 | 22844771 | 22844771 | G | A | rs3751489 |
| **USPL1** | chr13 | 30646969 | 30646969 | C | T | rs3742303 |
| **GCAT** | chr22 | 37810117 | 37810117 | G | T | rs182422563 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **CACNG8** | chr19 | 53979945 | 53979945 | C | T | rs370446891 |
| **ITGA5** | chr12 | 54403913 | 54403913 | A | G | rs772305445 |
| **KIAA0586** | chr14 | 58492205 | 58492205 | A | G | rs3783697 |
| **ELMO3** | chr16 | 67201748 | 67201748 | C | T | rs79736950 |
| **VIPAS39** | chr14 | 77428381 | 77428381 | G | A | rs75211061 |
| **ENGASE** | chr17 | 79081987 | 79081987 | G | A | rs12937557 |
| **TMC3** | chr15 | 81362283 | 81362283 | G | A | rs76735592 |
| **ROR2** | chr9 | 91724411 | 91724411 | C | T | rs34431454 |
| **ANO4** | chr12 | 100942423 | 100942423 | G | C | rs34162417 |
| **TAAR5** | chr6 | 132589346 | 132589346 | T | A | rs80078646 |
| **KY** | chr3 | 134610379 | 134610379 | C | T | rs753733495 |
| **SOX30** | chr5 | 157646739 | 157646739 | G | T | rs12188040 |
| **DOCK10** | chr2 | 224841896 | 224841896 | C | T | rs144638149 |

Table 27. Variants present in 2 out of 10 hypopharyngeal cancer samples (256 variants in 128 genes)

| Gene | Chr | Start | End | Ref | Alt | rsID |
|---|---|---|---|---|---|---|
| **ABCA3** | 16 | 2279097 | 2279097 | C | T | rs201955122 |
| **ABCC4** | 13 | 95247049 | 95247049 | G | C | rs11568689 |
| **ABCG2** | 4 | 88131113 | 88131113 | C | T | rs528655917 |
| **AGPS** | 2 | 177513838 | 177513838 | A | C | . |
| **ANKRD23** | 2 | 96839581 | 96839581 | C | T | . |
| **ANKRD44** | 2 | 197147072 | 197147072 | C | T | rs200088798 |
| **ANXA9** | 1 | 150987947 | 150987947 | C | T | rs138102093 |
| **AOAH** | 7 | 36618320 | 36618320 | G | A | rs80187893 |
| **APMAP** | 20 | 24971511 | 24971511 | G | A | rs78661674 |
| **ASXL2** | 2 | 25799530 | 25799530 | A | T | rs150279913 |
| **ATP8A1** | 4 | 42507084 | 42507084 | G | A | rs3792687 |
| **AVPI1** | 10 | 97679697 | 97679697 | G | T | rs117859077 |
| **C1orf158** | 1 | 12759337 | 12759337 | C | T | rs151319866 |
| **CAMSAP1** | 9 | 135811589 | 135811589 | T | C | rs201218402 |
| **CD34** | 1 | 207888710 | 207888710 | C | T | rs146829371 |
| **CDC123** | 10 | 12215798 | 12215798 | G | A | . |
| **CEMIP2** | 9 | 71709323 | 71709323 | C | T | rs147272925 |
| **CENPE** | 4 | 103198273 | 103198273 | A | G | rs779727126 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **CHRNA2** | 8 | 27463458 | 27463458 | C | A | . |
| **D2HGDH** | 2 | 241767928 | 241767928 | G | T | . |
| **DCAF1** | 3 | 51443810 | 51443810 | T | C | . |
| **DES** | 2 | 219418833 | 219418833 | A | G | rs564121737 |
| **DGKH** | 13 | 42168733 | 42168733 | G | A | . |
| **DIP2B** | 12 | 50678709 | 50678709 | C | T | rs752300911 |
| **DIP2B** | 12 | 50698401 | 50698401 | C | T | . |
| **DIXDC1** | 11 | 112016769 | 112016769 | C | T | rs184718561 |
| **DMPK** | 19 | 45778500 | 45778500 | C | T | rs145330026 |
| **DOCK6** | 19 | 11238284 | 11238284 | G | A | rs12609039 |
| **DPH6** | 15 | 35373564 | 35373564 | G | C | rs10519996 |
| **DPPA2** | 3 | 109304611 | 109304611 | G | A | rs748847987 |
| **DYNC2H1** | 11 | 103155401 | 103155401 | G | A | . |
| **ENPP6** | 4 | 184091272 | 184091272 | C | T | rs187347918 |
| **EPHA5** | 4 | 65602256 | 65602256 | A | G | . |
| **ETV3L** | 1 | 157099250 | 157099250 | C | T | rs568226360 |
| **FAAP100** | 17 | 81547646 | 81547646 | C | T | rs550951998 |
| **FAT2** | 5 | 151529320 | 151529320 | C | T | rs771350129 |
| **FCSK** | 16 | 70466181 | 70466181 | C | T | rs140084649 |
| **GDPD5** | 11 | 75449053 | 75449053 | A | T | rs757943424 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **GLT6D1** | 9 | 135624282 | 135624282 | C | T | rs61739510 |
| **GNE** | 9 | 36236961 | 36236961 | A | C | . |
| **GOLGA2** | 9 | 128260659 | 128260659 | G | A | rs372407712 |
| **GPATCH2** | 1 | 217431312 | 217431312 | T | C | rs35737297 |
| **GPR55** | 2 | 230910607 | 230910607 | C | T | . |
| **GRN** | 17 | 44352187 | 44352187 | C | T | rs752428000 |
| **GSR** | 8 | 30708107 | 30708107 | G | A | rs8190955 |
| **HKDC1** | 10 | 69250401 | 69250401 | T | C | . |
| **HMCN1** | 1 | 186074822 | 186074822 | G | A | rs149435109 |
| **HMGCLL1** | 6 | 55541767 | 55541767 | C | G | rs376191221 |
| **HPX** | 11 | 6431685 | 6431685 | G | A | rs779098616 |
| **HYLS1** | 11 | 125899652 | 125899652 | A | C | . |
| **IFIH1** | 2 | 162288137 | 162288137 | T | C | rs117608083 |
| **INSM1** | 20 | 20368641 | 20368641 | C | T | . |
| **JMJD7-PLA2G4B;** | 15 | 41840888 | 41840888 | T | G | rs201696797 |
| **KALRN** | 3 | 124657464 | 124657464 | T | C | rs772703622 |
| **L3MBTL3** | 6 | 130057478 | 130057478 | C | T | rs199734515 |
| **LGI4** | 19 | 35126275 | 35126275 | A | C | . |
| **LGR6** | 1 | 202276481 | 202276481 | C | T | rs117583891 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **LTBP4** | 19 | 40613979 | 40613979 | C | A | rs775003342 |
| **LYST** | 1 | 235733680 | 235733680 | C | T | rs200353560 |
| **MBTPS1** | 16 | 84095677 | 84095677 | G | A | . |
| **MCM10** | 10 | 13209117 | 13209117 | G | A | rs187685058 |
| **MCM2** | 3 | 127618067 | 127618067 | G | A | rs2307311 |
| **ME1** | 6 | 83253682 | 83253682 | C | T | rs118081320 |
| **MELTF** | 3 | 197017123 | 197017123 | G | A | rs2276790 |
| **METTL6** | 3 | 15414101 | 15414101 | A | G | . |
| **MINDY4** | 7 | 30840772 | 30840772 | G | A | rs755386534 |
| **MINDY4** | 7 | 30882350 | 30882350 | G | A | rs200111157 |
| **MMP27** | 11 | 102695000 | 102695000 | C | T | rs139446845 |
| **MOCS2** | 5 | 53101507 | 53101507 | T | C | rs2233215 |
| **MRPL22** | 5 | 154966736 | 154966736 | G | A | rs3749671 |
| **MYH15** | 3 | 108492565 | 108492565 | T | C | rs374250280 |
| **MYH4** | 17 | 10443493 | 10443493 | C | T | rs201770076 |
| **MYO18A** | 17 | 29111520 | 29111520 | G | A | rs116906886 |
| **MYO1A** | 12 | 57043086 | 57043086 | T | C | rs778384327 |
| **NCBP3** | 17 | 3846117 | 3846117 | T | G | rs12449334 |
| **OR51S1** | 11 | 4848824 | 4848824 | C | T | . |
| **OR6Q1** | 11 | 58031136 | 58031136 | T | C | rs372602969 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **OVCH2** | 11 | 7700427 | 7700427 | C | T | rs61759818 |
| **PDGFRB** | 5 | 150117817 | 150117817 | T | G | rs879255377 |
| **PELI2** | 14 | 56118723 | 56118723 | G | C | rs117942237 |
| **PHYKPL** | 5 | 178213038 | 178213038 | G | C | rs781761949 |
| **PIGX** | 3 | 196728068 | 196728068 | C | T | rs2291397 |
| **PKHD1** | 6 | 51746881 | 51746881 | A | G | . |
| **PLA2G7** | 6 | 46708081 | 46708081 | A | T | rs201842579 |
| **PLEKHA6** | 1 | 204228095 | 204228095 | G | A | rs775252585 |
| **PWP1** | 12 | 107688476 | 107688476 | G | A | rs117231086 |
| **PXN** | 12 | 120213958 | 120213958 | G | A | rs200828118 |
| **RAMAC** | 15 | 82989160 | 82989160 | G | C | rs763072846 |
| **RASGEF1C** | 5 | 180128449 | 180128449 | G | T | . |
| **REPIN1** | 7 | 150371765 | 150371765 | G | A | . |
| **RNF113B** | 13 | 98177131 | 98177131 | G | A | . |
| **RRH** | 4 | 109833215 | 109833215 | T | A | . |
| **SAMD9L** | 7 | 93135106 | 93135106 | A | G | rs2073793 |
| **SCN10A** | 3 | 38756807 | 38756807 | A | C | rs78555408 |
| **SERPINB9** | 6 | 2890296 | 2890296 | G | A | rs146773314 |
| **SH3TC1** | 4 | 8227684 | 8227684 | G | A | rs755241625 |
| **SLC12A6** | 15 | 34245819 | 34245819 | C | G | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **SLC6A18** | 5 | 1232300 | 1232300 | G | A | rs148359610 |
| **SMARCA1** | X | 129481079 | 129481079 | G | T | rs186125032 |
| **SNX19** | 11 | 130914429 | 130914429 | G | A | . |
| **SORBS1** | 10 | 95384090 | 95384090 | C | T | rs370290679 |
| **SORT1** | 1 | 109342050 | 109342050 | C | A | rs2228605 |
| **SPECC1L** | 22 | 24321443 | 24321443 | C | T | rs148203655 |
| **SPTB** | 14 | 64767821 | 64767821 | C | A | . |
| **SREBF2** | 22 | 41898747 | 41898747 | G | C | rs17848351 |
| **SSC5D** | 19 | 55490945 | 55490945 | G | A | rs576454141 |
| **STAB1** | 3 | 52523733 | 52523733 | C | T | . |
| **STK25** | 2 | 241498777 | 241498777 | G | A | rs199688267 |
| **SYNE1** | 6 | 152293733 | 152293733 | C | T | rs80265744 |
| **TCIRG1** | 11 | 68041789 | 68041789 | G | C | rs118141250 |
| **TEP1** | 14 | 20383549 | 20383549 | T | C | rs117775572 |
| **THUMPD3** | 3 | 9371528 | 9371528 | G | C | rs369103372 |
| **TMED6** | 16 | 69347921 | 69347921 | C | G | rs187873309 |
| **TMEM117** | 12 | 44299657 | 44299657 | G | A | rs756921980 |
| **TNS3** | 7 | 47303111 | 47303111 | G | A | rs999907801 |
| **TRPM2** | 21 | 44390991 | 44390991 | G | A | rs745524892 |
| **TSNAXIP1** | 16 | 67825718 | 67825718 | G | A | rs754752368 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **TSR3** | 16 | 1350821 | 1350821 | G | A | rs142807232 |
| **TXNDC15** | 5 | 134888114 | 134888114 | G | C | rs79262456 |
| **UGT1A1** | 2 | 233768226 | 233768226 | C | T | rs34946978 |
| **UPK1B** | 3 | 119194251 | 119194251 | C | G | rs557080777 |
| **UTRN** | 6 | 144447729 | 144447729 | T | C | . |
| **VWA8** | 13 | 41670984 | 41670984 | A | G | rs375522079 |
| **ZNF106** | 15 | 42422605 | 42422605 | G | A | rs751935490 |
| **ZNF133** | 20 | 18315881 | 18315881 | C | T | . |
| **ZNF362** | 1 | 33295202 | 33295202 | G | A | rs374616798 |
| **ZNF574** | 19 | 42080053 | 42080053 | C | T | . |
| **ZPR1** | 11 | 116778930 | 116778930 | G | A | rs757653816 |

Table 28. Variants found in single samples across ten hypopharyngeal cancer samples (487 variants in 126 genes)

| Gene | Chr | Start | End | Ref | Alt | rsID |
|---|---|---|---|---|---|---|
| **AAK1** | chr2 | 69642958 | 69642958 | G | A | rs372153554 |
| **ABCA2** | chr9 | 137028136 | 137028136 | C | T | . |
| **ABCC5** | chr3 | 183951453 | 183951453 | C | G | . |
| **ABCF2;ABCF2-H2BE1** | chr7 | 151223816 | 151223816 | C | A | . |
| **ABCG8** | chr2 | 43851712 | 43851712 | G | T | . |
| **ACAA2** | chr18 | 49797570 | 49797570 | C | T | . |
| **ACBD3** | chr1 | 226152493 | 226152493 | C | G | . |
| **ACTL7B** | chr9 | 108855741 | 108855741 | C | T | rs748279247 |
| **ADAM10** | chr15 | 58611971 | 58611971 | C | A | . |
| **ADAM22** | chr7 | 88153245 | 88153245 | G | T | . |
| **ADAM33** | chr20 | 3673605 | 3673605 | C | T | . |
| **ADCYAP1** | chr18 | 905495 | 905495 | A | T | . |
| **ADGRB2** | chr1 | 31730840 | 31730840 | C | T | rs746715738 |
| **ADGRB3** | chr6 | 68956066 | 68956066 | C | T | . |
| **ADGRL3** | chr4 | 61909615 | 61909615 | G | T | rs192210727 |
| **ADRM1** | chr20 | 62308030 | 62308030 | C | A | . |
| **AEBP2** | chr12 | 19514774 | 19514774 | A | C | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **AGXT2** | chr5 | 35010017 | 35010017 | C | G | . |
| **AK2** | chr1 | 33036756 | 33036756 | C | T | rs199641857 |
| **AKAP13** | chr15 | 85543861 | 85543861 | C | T | rs114777682 |
| **AKR1D1** | chr7 | 138113718 | 138113718 | A | C | . |
| **ALDH1L1** | chr3 | 126114580 | 126114580 | C | A | rs765800160 |
| **ANGPTL3** | chr1 | 62597690 | 62597690 | G | C | . |
| **ANK2** | chr4 | 113274627 | 113274627 | C | A | . |
| **ANK2** | chr4 | 113318556 | 113318556 | C | T | rs200524407 |
| **ANK3** | chr10 | 60173110 | 60173110 | G | A | rs190581397 |
| **ANKRD50** | chr4 | 124678735 | 124678735 | C | G | . |
| **ANKRD55** | chr5 | 56111211 | 56111211 | C | G | . |
| **ANKRD9** | chr14 | 102507730 | 102507730 | G | A | rs867729152 |
| **ANO2** | chr12 | 5615199 | 5615199 | A | G | . |
| **APOB** | chr2 | 21012072 | 21012072 | C | T | rs746414462 |
| **APOBEC2** | chr6 | 41061590 | 41061590 | G | A | rs142866037 |
| **AQR** | chr15 | 34874851 | 34874851 | C | A | . |
| **ARFIP2** | chr11 | 6478767 | 6478767 | C | A | . |
| **ARHGAP28** | chr18 | 6824836 | 6824836 | G | A | rs188975691 |
| **ARID3C** | chr9 | 34622035 | 34622035 | C | T | rs142710890 |
| **ARL9** | chr4 | 56518680 | 56518680 | G | C | rs77571713 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **ARMC2** | chr6 | 108854364 | 108854364 | G | A | rs796107747 |
| **ARPP21** | chr3 | 35792528 | 35792528 | C | G | . |
| **ASCC2** | chr22 | 29789107 | 29789107 | C | T | rs373433240 |
| **ATL2** | chr2 | 38314660 | 38314660 | A | T | . |
| **ATP11C** | chrX | 139804502 | 139804502 | G | T | . |
| **ATP5MF-PTCD1;PTCD1** | chr7 | 99433333 | 99433333 | A | G | rs368133338 |
| **ATP6V1A** | chr3 | 113805418 | 113805418 | C | T | . |
| **ATP7A** | chrX | 78003219 | 78003219 | G | T | . |
| **AVIL** | chr12 | 57810901 | 57810901 | T | G | . |
| **B4GALNT1** | chr12 | 57627768 | 57627768 | A | T | . |
| **BAX** | chr19 | 48955794 | 48955794 | G | A | . |
| **BCS1L** | chr2 | 218662232 | 218662232 | C | G | . |
| **BIRC3** | chr11 | 102325495 | 102325495 | T | A | . |
| **BIVM-ERCC5;ERCC5** | chr13 | 102866790 | 102866790 | T | A | rs761058720 |
| **BLMH** | chr17 | 30249088 | 30249088 | C | T | rs116194289 |
| **BLOC1S4** | chr4 | 6716294 | 6716294 | G | C | . |
| **BMPER** | chr7 | 34058137 | 34058137 | T | C | . |
| **BMS1** | chr10 | 42791700 | 42791700 | G | A | rs2272881 |
| **BRI3** | chr7 | 98291212 | 98291212 | G | A | rs772300526 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **BRINP2** | chr1 | 177273495 | 177273495 | A | C | . |
| **BRINP3** | chr1 | 190226168 | 190226168 | G | T | . |
| **C1orf112** | chr1 | 169830739 | 169830739 | G | T | rs2272920 |
| **CABP5** | chr19 | 48040649 | 48040649 | G | C | rs34862923 |
| **CACNA1D** | chr3 | 53722370 | 53722370 | A | G | . |
| **CADPS2** | chr7 | 122471513 | 122471513 | C | A | . |
| **CALCRL** | chr2 | 187352301 | 187352301 | C | A | . |
| **CAPN3** | chr15 | 42386197 | 42386197 | G | A | . |
| **CAPS2** | chr12 | 75298889 | 75298889 | C | T | rs780956187 |
| **CCDC157** | chr22 | 30366058 | 30366058 | G | A | rs773064591 |
| **CCDC28B** | chr1 | 32202008 | 32202008 | C | T | rs1407134 |
| **CD34** | chr1 | 207887797 | 207887797 | C | A | rs28362497 |
| **CDADC1** | chr13 | 49280548 | 49280548 | T | A | . |
| **CDH12** | chr5 | 21755800 | 21755800 | C | T | rs763451842 |
| **CDH2** | chr18 | 28009736 | 28009736 | T | A | . |
| **CDK7** | chr5 | 69254638 | 69254638 | T | C | . |
| **CELF2** | chr10 | 11275094 | 11275094 | C | T | . |
| **CELSR3** | chr3 | 48661578 | 48661578 | C | A | . |
| **CENPN** | chr16 | 81028273 | 81028273 | G | A | rs779941636 |
| **CEP41** | chr7 | 130440963 | 130440963 | A | C | rs782769549 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **CEP78** | chr9 | 78265412 | 78265412 | G | T | . |
| **CEP97** | chr3 | 101724690 | 101724690 | G | T | rs201771736 |
| **CHEK1** | chr11 | 125633204 | 125633204 | C | T | rs140276570 |
| **CILP2** | chr19 | 19543858 | 19543858 | G | A | . |
| **CLCA4** | chr1 | 86580034 | 86580034 | C | A | . |
| **CLCN1** | chr7 | 143330868 | 143330868 | G | T | . |
| **CLEC10A** | chr17 | 7075164 | 7075164 | C | T | rs375550026 |
| **CLIC4** | chr1 | 24839985 | 24839985 | G | T | . |
| **CLIP1** | chr12 | 122272966 | 122272966 | T | C | rs765475076 |
| **CLP1** | chr11 | 57661006 | 57661006 | G | T | . |
| **CNGA3** | chr2 | 98396864 | 98396864 | C | T | rs201747279 |
| **CNIH1** | chr14 | 54441317 | 54441317 | G | C | . |
| **CNN2** | chr19 | 1032601 | 1032601 | G | A | . |
| **COG6** | chr13 | 39659437 | 39659437 | G | T | . |
| **COL6A6** | chr3 | 130571132 | 130571132 | C | T | rs200963433 |
| **COPB2** | chr3 | 139379155 | 139379155 | C | T | rs79043251 |
| **COPS9** | chr2 | 240133971 | 240133971 | G | A | . |
| **CRACR2A** | chr12 | 3659605 | 3659605 | C | T | . |
| **CREBRF** | chr5 | 173090589 | 173090589 | C | T | . |
| **CRYBG2** | chr1 | 26338416 | 26338416 | A | C | rs151324745 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **CRYBG3** | chr3 | 97936895 | 97936895 | G | A | rs370818948 |
| **CSMD2** | chr1 | 33846925 | 33846925 | T | A | . |
| **CTNNA2** | chr2 | 79909661 | 79909661 | T | A | . |
| **CTR9** | chr11 | 10763830 | 10763830 | G | T | . |
| **CUX1** | chr7 | 102280082 | 102280082 | G | A | rs139293638 |
| **CYLD** | chr16 | 50792600 | 50792600 | C | T | . |
| **DCLK2** | chr4 | 150224528 | 150224528 | G | T | . |
| **DCN** | chr12 | 91151689 | 91151689 | C | A | . |
| **DDX31** | chr9 | 132612139 | 132612139 | G | T | . |
| **DDX46** | chr5 | 134794859 | 134794859 | A | C | . |
| **DGKH** | chr13 | 42127545 | 42127545 | G | A | rs150403121 |
| **DHRS9** | chr2 | 169083518 | 169083518 | G | A | rs146976196 |
| **DHTKD1** | chr10 | 12118774 | 12118774 | G | A | rs757252285 |
| **DHX40** | chr17 | 59570617 | 59570617 | G | T | . |
| **DKC1** | chrX | 154766284 | 154766284 | G | A | rs1036880108 |
| **DNAH11** | chr7 | 21816513 | 21816513 | C | A | rs573384750 |
| **DNAH11** | chr7 | 21816605 | 21816605 | C | T | . |
| **DNAH3** | chr16 | 20964700 | 20964700 | C | A | . |
| **DNAH6** | chr2 | 84624505 | 84624505 | G | A | rs529351050 |
| **DNAJC22** | chr12 | 49349230 | 49349230 | C | A | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **DNER** | chr2 | 229366928 | 229366928 | A | C | rs561051341 |
| **DOCK3** | chr3 | 51159259 | 51159259 | G | C | . |
| **DOCK5** | chr8 | 25372602 | 25372602 | G | A | . |
| **DOCK8** | chr9 | 406981 | 406981 | G | T | rs767147947 |
| **DPCD** | chr10 | 101600750 | 101600750 | G | C | rs149463737 |
| **DPP10** | chr2 | 115689911 | 115689911 | G | A | rs200502140 |
| **DPP3** | chr11 | 66493111 | 66493111 | G | T | . |
| **DQX1** | chr2 | 74522610 | 74522610 | C | T | rs755763286 |
| **DRG1** | chr22 | 31399705 | 31399705 | A | G | rs762210000 |
| **DSP** | chr6 | 7555816 | 7555816 | A | G | rs188516326 |
| **DTNB** | chr2 | 25432895 | 25432895 | C | A | . |
| **DTX2** | chr7 | 76482882 | 76482882 | C | T | rs148430404 |
| **DUSP11** | chr2 | 73769302 | 73769302 | T | G | . |
| **DUSP19** | chr2 | 183095464 | 183095464 | G | T | . |
| **DUSP2** | chr2 | 96145228 | 96145228 | C | T | . |
| **DYNC1I2** | chr2 | 171727870 | 171727870 | T | G | rs201982509 |
| **DYNC2H1** | chr11 | 103153334 | 103153334 | G | C | . |
| **DYSF** | chr2 | 71660618 | 71660618 | A | G | rs759505768 |
| **EFCAB5** | chr17 | 30090626 | 30090626 | G | A | . |
| **EIF4G3** | chr1 | 20981172 | 20981172 | G | A | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **ENDOG** | chr9 | 128819100 | 128819100 | G | A | . |
| **ENPP3** | chr6 | 131740348 | 131740348 | C | A | . |
| **EP300** | chr22 | 41169525 | 41169525 | G | T | rs1057519889 |
| **EPB41L3** | chr18 | 5428431 | 5428431 | C | A | . |
| **EPHA6** | chr3 | 97610853 | 97610853 | G | A | rs776617495 |
| **EPHA6** | chr3 | 97720330 | 97720330 | G | T | . |
| **ERAP2** | chr5 | 96883857 | 96883857 | C | T | rs3733905 |
| **ERI3** | chr1 | 44247944 | 44247944 | C | A | . |
| **ESAM** | chr11 | 124753662 | 124753662 | G | A | rs114481311 |
| **ETFA** | chr15 | 76286421 | 76286421 | G | A | rs1801591 |
| **ETFB** | chr19 | 51347045 | 51347045 | G | A | rs74735908 |
| **EVL** | chr14 | 100084712 | 100084712 | G | A | . |
| **EYA4** | chr6 | 133481576 | 133481576 | C | A | . |
| **F11** | chr4 | 186286493 | 186286493 | G | C | rs796615398 |
| **FAH** | chr15 | 80158159 | 80158159 | G | T | rs151264725 |
| **FAM161B** | chr14 | 73942692 | 73942692 | G | A | rs138585435 |
| **FAM184A** | chr6 | 118961921 | 118961921 | T | G | . |
| **FAM8A1** | chr6 | 17605016 | 17605016 | T | A | . |
| **FARP2** | chr2 | 241463341 | 241463341 | C | T | rs377640265 |
| **FICD** | chr12 | 108519140 | 108519140 | C | G | rs139753560 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **FLNC** | chr7 | 128854640 | 128854640 | G | A | . |
| **FMN2** | chr1 | 240330738 | 240330738 | G | T | . |
| **FN1** | chr2 | 215408378 | 215408378 | C | T | . |
| **FOXS1** | chr20 | 31845299 | 31845299 | C | T | rs371871349 |
| **FRMD6** | chr14 | 51708204 | 51708204 | G | T | . |
| **FRMPD4** | chrX | 12716884 | 12716884 | G | T | . |
| **FZD1** | chr7 | 91265206 | 91265206 | C | T | . |
| **GABRA4** | chr4 | 46979083 | 46979083 | A | G | . |
| **GABRB2** | chr5 | 161331121 | 161331121 | G | T | . |
| **GAREM1** | chr18 | 32287921 | 32287921 | C | G | . |
| **GCAT** | chr22 | 37815280 | 37815280 | G | A | rs200392738 |
| **GEMIN5** | chr5 | 154898571 | 154898571 | C | T | rs77711671 |
| **GHITM** | chr10 | 84150196 | 84150196 | G | A | . |
| **GINS4** | chr8 | 41541835 | 41541835 | G | A | rs141719341 |
| **GLRB** | chr4 | 157143846 | 157143846 | G | C | . |
| **GMPS** | chr3 | 155937651 | 155937651 | G | A | . |
| **GNAQ** | chr9 | 77922307 | 77922307 | T | G | rs773781296 |
| **GNL2** | chr1 | 37576439 | 37576439 | C | T | . |
| **GOLGA5** | chr14 | 92810322 | 92810322 | T | C | rs183339423 |
| **GOLT1A** | chr1 | 204201709 | 204201709 | C | A | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **GPR171** | chr3 | 151198557 | 151198557 | T | C | rs141833565 |
| **GPR55** | chr2 | 230910566 | 230910566 | G | A | rs772132028 |
| **GRIN3A** | chr9 | 101628322 | 101628322 | G | T | . |
| **GRIN3A** | chr9 | 101670326 | 101670326 | C | G | . |
| **GRM3** | chr7 | 86786366 | 86786366 | C | T | rs765354209 |
| **GRM6** | chr5 | 178990691 | 178990691 | C | A | rs371669870 |
| **HABP2** | chr10 | 113581984 | 113581984 | G | A | rs138864377 |
| **HCFC1R1** | chr16 | 3023308 | 3023308 | G | A | . |
| **HELQ** | chr4 | 83421627 | 83421627 | C | A | rs752417171 |
| **HERC6** | chr4 | 88383281 | 88383281 | C | A | . |
| **HEXA** | chr15 | 72345537 | 72345537 | C | T | rs145012038 |
| **HIF3A** | chr19 | 46305300 | 46305300 | G | T | . |
| **HMCN1** | chr1 | 186117572 | 186117572 | A | T | rs184081240 |
| **HOXB9** | chr17 | 48623109 | 48623109 | C | G | . |
| **HSD17B6** | chr12 | 56773968 | 56773968 | C | T | rs201098412 |
| **HSP90AA1** | chr14 | 102083084 | 102083084 | C | T | . |
| **HSPB2** | chr11 | 111913662 | 111913662 | C | T | . |
| **HSPD1** | chr2 | 197498682 | 197498682 | C | T | rs200514123 |
| **HSPH1** | chr13 | 31151740 | 31151740 | C | A | . |
| **HTR2C** | chrX | 114848129 | 114848129 | C | A | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **ICA1** | chr7 | 8218358 | 8218358 | C | T | rs148267518 |
| **IFT80** | chr3 | 160268495 | 160268495 | G | A | . |
| **IGFBP2** | chr2 | 216633651 | 216633651 | C | A | . |
| **IGSF9B** | chr11 | 133922617 | 133922617 | C | A | . |
| **ING3** | chr7 | 120969065 | 120969065 | A | T | . |
| **INPP5J** | chr22 | 31128603 | 31128603 | G | T | . |
| **IQCA1** | chr2 | 236338370 | 236338370 | G | A | rs186626813 |
| **IQUB** | chr7 | 123457531 | 123457531 | C | A | . |
| **IRAG2** | chr12 | 25101230 | 25101230 | T | C | rs531075630 |
| **ITGAE** | chr17 | 3759546 | 3759546 | A | G | rs375922702 |
| **ITGB5** | chr3 | 124809140 | 124809140 | A | T | rs781752274 |
| **ITK** | chr5 | 157248957 | 157248957 | C | T | rs34482255 |
| **JAKMIP2** | chr5 | 147644990 | 147644990 | G | T | . |
| **JDP2** | chr14 | 75469399 | 75469399 | G | A | . |
| **KALRN** | chr3 | 124562922 | 124562922 | G | A | . |
| **KCNA3** | chr1 | 110674379 | 110674379 | C | T | . |
| **KCNG4** | chr16 | 84222560 | 84222560 | G | T | . |
| **KDM1A** | chr1 | 23083317 | 23083317 | C | T | rs779514366 |
| **KIAA0930** | chr22 | 45203063 | 45203063 | C | A | . |
| **KIF19** | chr17 | 74351877 | 74351877 | A | G | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **KIF4A** | chrX | 70302021 | 70302021 | A | C | . |
| **KIT** | chr4 | 54736592 | 54736592 | G | T | . |
| **KLHDC7B** | chr22 | 50548924 | 50548924 | T | C | rs202093399 |
| **KLHL21** | chr1 | 6602651 | 6602651 | G | A | . |
| **KNTC1** | chr12 | 122597908 | 122597908 | G | C | rs186936079 |
| **KRTCAP3** | chr2 | 27442662 | 27442662 | G | A | . |
| **KSR2** | chr12 | 117476578 | 117476578 | C | T | rs183487509 |
| **LACTB** | chr15 | 63141708 | 63141708 | G | T | . |
| **LAMB1** | chr7 | 107959361 | 107959361 | C | T | rs35710474 |
| **LAMC1** | chr1 | 183110604 | 183110604 | A | G | rs765463384 |
| **LBX1** | chr10 | 101227281 | 101227281 | C | A | . |
| **LCK** | chr1 | 32285523 | 32285523 | A | C | . |
| **LHX9** | chr1 | 197920128 | 197920128 | T | G | . |
| **LIMK2** | chr22 | 31267879 | 31267879 | G | A | rs890335349 |
| **LNX2** | chr13 | 27581441 | 27581441 | C | A | . |
| **LONP2** | chr16 | 48348238 | 48348238 | G | A | rs372954881 |
| **LOXL1** | chr15 | 73949528 | 73949528 | C | A | . |
| **LPO** | chr17 | 58264849 | 58264849 | G | C | . |
| **LRP1** | chr12 | 57169296 | 57169296 | G | A | . |
| **LRRC1** | chr6 | 53795362 | 53795362 | C | T | rs755914646 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **LTK** | chr15 | 41505968 | 41505968 | C | G | . |
| **LY75;LY75-CD302** | chr2 | 159834145 | 159834145 | G | A | rs78446341 |
| **M1AP** | chr2 | 74607174 | 74607174 | T | A | rs117569670 |
| **MACF1** | chr1 | 39368154 | 39368154 | A | T | rs145646053 |
| **MAEA** | chr4 | 1322441 | 1322441 | C | T | rs11553129 |
| **MAP3K10** | chr19 | 40206080 | 40206080 | G | A | rs904167606 |
| **MAPK7** | chr17 | 19382867 | 19382867 | G | T | . |
| **MAPKAPK2** | chr1 | 206728803 | 206728803 | G | C | . |
| **MAPKAPK5** | chr12 | 111892970 | 111892970 | G | A | rs150495619 |
| **MARK3** | chr14 | 103491850 | 103491850 | C | T | rs181804827 |
| **MDGA1** | chr6 | 37654333 | 37654333 | C | T | . |
| **MDN1** | chr6 | 89751446 | 89751446 | C | T | rs151315441 |
| **MED13** | chr17 | 61982284 | 61982284 | C | T | rs764794724 |
| **METRN** | chr16 | 715893 | 715893 | C | A | . |
| **METTL26** | chr16 | 636169 | 636169 | T | C | . |
| **MGAT1** | chr5 | 180792412 | 180792412 | C | T | rs775360849 |
| **MICALL1** | chr22 | 37906529 | 37906529 | C | T | rs781388991 |
| **MINDY3** | chr10 | 15779071 | 15779071 | G | C | . |
| **MIS18BP1** | chr14 | 45217099 | 45217099 | T | A | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **MME** | chr3 | 155167006 | 155167006 | G | A | . |
| **MPDZ** | chr9 | 13221413 | 13221413 | C | T | . |
| **MPO** | chr17 | 58278090 | 58278090 | C | T | . |
| **MPP4** | chr2 | 201654906 | 201654906 | C | A | . |
| **MST1R** | chr3 | 49891297 | 49891297 | C | T | rs771218531 |
| **MTFR2** | chr6 | 136241501 | 136241501 | G | A | rs151177675 |
| **MTHFD1L** | chr6 | 150972005 | 150972005 | C | T | . |
| **MTNR1A** | chr4 | 186533881 | 186533881 | G | T | . |
| **MUSK** | chr9 | 110787743 | 110787743 | T | C | rs751296377 |
| **MUTYH** | chr1 | 45332242 | 45332242 | C | T | rs730881833 |
| **MVD** | chr16 | 88656233 | 88656233 | C | G | . |
| **MYCBP2** | chr13 | 77206778 | 77206778 | G | A | . |
| **MYH1** | chr17 | 10500677 | 10500677 | G | A | rs117616137 |
| **MYH2** | chr17 | 10526652 | 10526652 | C | A | . |
| **MYH4** | chr17 | 10452167 | 10452167 | C | A | . |
| **MYH4** | chr17 | 10461008 | 10461008 | A | G | . |
| **MYO10** | chr5 | 16781804 | 16781804 | C | T | . |
| **MYO7A** | chr11 | 77142773 | 77142773 | T | C | . |
| **MYO7A** | chr11 | 77181443 | 77181443 | C | T | rs373089701 |
| **MYO7B** | chr2 | 127622049 | 127622049 | G | A | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **MYO7B** | chr2 | 127635878 | 127635878 | C | T | rs181858311 |
| **NALCN** | chr13 | 101192026 | 101192026 | G | A | . |
| **NCOR2** | chr12 | 124336761 | 124336761 | C | T | rs775458405 |
| **NEB** | chr2 | 151688389 | 151688389 | T | C | rs77151072 |
| **NEXN** | chr1 | 77917984 | 77917984 | G | A | . |
| **NF1** | chr17 | 31235922 | 31235922 | A | G | rs1060500243 |
| **NFAT5** | chr16 | 69659882 | 69659882 | C | A | . |
| **NHSL1** | chr6 | 138424464 | 138424464 | C | T | . |
| **NID1** | chr1 | 236032409 | 236032409 | C | T | . |
| **NOTCH1** | chr9 | 136515621 | 136515621 | A | T | . |
| **NOTCH1** | chr9 | 136517800 | 136517800 | C | T | rs1057523819 |
| **NOTCH1** | chr9 | 136518695 | 136518695 | C | T | . |
| **NOTCH2** | chr1 | 119967546 | 119967546 | C | T | . |
| **NOX3** | chr6 | 155454894 | 155454894 | C | A | rs149127858 |
| **NPFFR1** | chr10 | 70255274 | 70255274 | G | A | . |
| **NR1D1;THRA** | chr17 | 40093276 | 40093276 | G | A | rs994746702 |
| **NRAP** | chr10 | 113606219 | 113606219 | C | T | rs764226845 |
| **NRP2** | chr2 | 205726063 | 205726063 | C | T | . |
| **NSL1** | chr1 | 212782413 | 212782413 | G | A | . |
| **NUDCD2** | chr5 | 163454013 | 163454013 | C | T | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **OGDH** | chr7 | 44645345 | 44645345 | C | T | rs377142006 |
| **OGDH** | chr7 | 44647661 | 44647661 | G | C | . |
| **OPA1** | chr3 | 193659547 | 193659547 | C | T | rs190235251 |
| **OXSM** | chr3 | 25794301 | 25794301 | C | G | . |
| **P2RX1** | chr17 | 3898987 | 3898987 | G | T | . |
| **PABPC1** | chr8 | 100705591 | 100705591 | A | G | rs80006036 |
| **PABPC1** | chr8 | 100705604 | 100705604 | G | C | rs75035099 |
| **PAH** | chr12 | 102855297 | 102855297 | T | C | rs199475617 |
| **PAICS** | chr4 | 56446775 | 56446775 | G | A | . |
| **PAK4** | chr19 | 39175331 | 39175331 | G | A | rs557625940 |
| **PANK3** | chr5 | 168561428 | 168561428 | T | A | rs77612793 |
| **PAPPA** | chr9 | 116227448 | 116227448 | G | A | rs757211142 |
| **PARP1** | chr1 | 226363128 | 226363128 | T | C | rs3219145 |
| **PARP3** | chr3 | 51944827 | 51944827 | C | G | . |
| **PCDH10** | chr4 | 133151443 | 133151443 | G | A | . |
| **PCDH15** | chr10 | 53903293 | 53903293 | C | T | rs149478475 |
| **PCDH20** | chr13 | 61413861 | 61413861 | C | T | . |
| **PCDHGC5** | chr5 | 141491593 | 141491593 | A | T | . |
| **PCSK5** | chr9 | 76071789 | 76071789 | C | A | . |
| **PDE8A** | chr15 | 85091077 | 85091077 | T | A | rs117384144 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **PDE8B** | chr5 | 77349474 | 77349474 | G | A | . |
| **PDE9A** | chr21 | 42762102 | 42762102 | A | C | . |
| **PDGFD** | chr11 | 103926988 | 103926988 | C | T | . |
| **PDPR** | chr16 | 70156728 | 70156728 | T | G | . |
| **PDSS1** | chr10 | 26742528 | 26742528 | G | A | rs759110650 |
| **PDSS2** | chr6 | 107212134 | 107212134 | C | T | rs796819612 |
| **PES1** | chr22 | 30581589 | 30581589 | G | A | rs142214789 |
| **PGK2** | chr6 | 49786991 | 49786991 | C | T | rs185374279 |
| **PIAS1** | chr15 | 68175711 | 68175711 | C | T | . |
| **PIK3CD** | chr1 | 9718830 | 9718830 | G | A | rs373591202 |
| **PLA2G1B** | chr12 | 120322275 | 120322275 | C | T | rs151139112 |
| **PLA2G4F** | chr15 | 42142138 | 42142138 | G | C | rs75560163 |
| **PLCB3** | chr11 | 64262485 | 64262485 | G | A | . |
| **PLCD4** | chr2 | 218622852 | 218622852 | T | C | rs567172551 |
| **PLK3** | chr1 | 44805566 | 44805566 | G | A | rs55654497 |
| **PLPP5** | chr8 | 38267296 | 38267296 | C | A | . |
| **PNPLA6** | chr19 | 7559128 | 7559128 | T | G | . |
| **POLR2D** | chr2 | 127852988 | 127852988 | G | A | . |
| **POR** | chr7 | 75985744 | 75985744 | C | T | rs782533830 |
| **PPP1R12A** | chr12 | 79778596 | 79778596 | C | T | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **PRDM8** | chr4 | 80201414 | 80201414 | G | C | . |
| **PREX1** | chr20 | 48649391 | 48649391 | C | A | . |
| **PRKCE** | chr2 | 45652345 | 45652345 | C | A | . |
| **PSMB10** | chr16 | 67935430 | 67935430 | G | C | rs200864375 |
| **PSMB4** | chr1 | 151400482 | 151400482 | C | T | . |
| **PTGIS** | chr20 | 49513085 | 49513085 | G | T | . |
| **PTPN13** | chr4 | 86803816 | 86803816 | C | T | . |
| **PTPRCAP** | chr11 | 67436148 | 67436148 | C | T | rs370062028 |
| **PTPRZ1** | chr7 | 122053995 | 122053995 | A | G | rs530492628 |
| **PXN** | chr12 | 120212499 | 120212499 | G | A | . |
| **RAD54B** | chr8 | 94400395 | 94400395 | G | A | rs367836964 |
| **RBL1** | chr20 | 37040201 | 37040201 | G | A | . |
| **RELN** | chr7 | 103522060 | 103522060 | C | T | . |
| **RELN** | chr7 | 103651665 | 103651665 | T | G | rs115734214 |
| **RET** | chr10 | 43119588 | 43119588 | G | T | . |
| **RGS22** | chr8 | 99999384 | 99999384 | G | A | rs3133711 |
| **RHOA** | chr3 | 49368487 | 49368487 | G | A | . |
| **RIMS2** | chr8 | 103766249 | 103766249 | G | T | . |
| **RIMS3** | chr1 | 40635973 | 40635973 | C | G | . |
| **RIPOR2** | chr6 | 24842899 | 24842899 | A | G | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **RIT2** | chr18 | 42923744 | 42923744 | C | G | . |
| **RNF113B** | chr13 | 98177177 | 98177177 | G | T | . |
| **RNF123** | chr3 | 49699104 | 49699104 | C | T | rs77308703 |
| **RNF44** | chr5 | 176530167 | 176530167 | T | C | . |
| **RO60** | chr1 | 193069486 | 193069486 | G | T | . |
| **ROS1** | chr6 | 117329369 | 117329369 | C | G | rs12664076 |
| **RPH3A** | chr12 | 112890972 | 112890972 | G | C | . |
| **RPS3A** | chr4 | 151102986 | 151102986 | A | C | rs139979828 |
| **RRAS** | chr19 | 49636887 | 49636887 | C | G | . |
| **RSBN1L** | chr7 | 77778351 | 77778351 | C | T | rs946521018 |
| **RTEL1** | chr20 | 63662859 | 63662859 | C | T | rs866785110 |
| **RUBCN** | chr3 | 197705129 | 197705129 | C | T | rs112632845 |
| **RUNX2** | chr6 | 45422815 | 45422815 | G | A | . |
| **RXFP2** | chr13 | 31793004 | 31793004 | G | T | . |
| **RYR2** | chr1 | 237634940 | 237634940 | T | G | . |
| **RYR3** | chr15 | 33550204 | 33550204 | G | C | . |
| **SACS** | chr13 | 23331499 | 23331499 | C | A | . |
| **SAMD4A** | chr14 | 54760459 | 54760459 | A | G | . |
| **SCFD1** | chr14 | 30630562 | 30630562 | A | G | . |
| **SCN8A** | chr12 | 51662908 | 51662908 | G | C | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **SCN8A** | chr12 | 51768884 | 51768884 | C | A | . |
| **SCN9A** | chr2 | 166272778 | 166272778 | G | A | . |
| **SCP2D1** | chr20 | 18814227 | 18814227 | G | T | . |
| **SCUBE3** | chr6 | 35231733 | 35231733 | G | A | . |
| **SEC23B** | chr20 | 18510909 | 18510909 | C | A | rs6045440 |
| **SF3B4** | chr1 | 149927260 | 149927260 | C | A | . |
| **SFI1** | chr22 | 31556999 | 31556999 | G | A | rs202143232 |
| **SFMBT1** | chr3 | 52906182 | 52906182 | A | T | . |
| **SGCE** | chr7 | 94623379 | 94623379 | G | A | rs557861177 |
| **SH2D3C** | chr9 | 127755120 | 127755120 | C | T | . |
| **SHKBP1** | chr19 | 40591056 | 40591056 | G | A | rs116820916 |
| **SHMT2** | chr12 | 57231762 | 57231762 | C | T | rs375584473 |
| **SI** | chr3 | 165016065 | 165016065 | C | G | . |
| **SI** | chr3 | 165019607 | 165019607 | G | C | rs202077921 |
| **SIPA1L2** | chr1 | 232483858 | 232483858 | C | G | . |
| **SIRT4** | chr12 | 120303934 | 120303934 | C | T | rs757053577 |
| **SLC16A13** | chr17 | 7038376 | 7038376 | C | T | rs201304096 |
| **SLC25A23** | chr19 | 6459564 | 6459564 | T | C | . |
| **SLC25A32** | chr8 | 103400460 | 103400460 | T | C | rs141856398 |
| **SLC26A1** | chr4 | 989837 | 989837 | C | T | rs3796623 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **SLC2A10** | chr20 | 46725191 | 46725191 | G | A | . |
| **SLC45A3** | chr1 | 205664517 | 205664517 | C | T | . |
| **SLC45A4** | chr8 | 141218217 | 141218217 | C | T | . |
| **SLC4A4** | chr4 | 71447653 | 71447653 | T | G | . |
| **SLC6A7** | chr5 | 150202628 | 150202628 | G | A | rs117381766 |
| **SLC9A6** | chrX | 136024366 | 136024366 | G | A | . |
| **SLC9A9** | chr3 | 143578629 | 143578629 | C | G | . |
| **SMAD4** | chr18 | 51059914 | 51059914 | C | A | . |
| **SMAD7** | chr18 | 48921442 | 48921442 | C | A | . |
| **SMARCA1** | chrX | 129468820 | 129468820 | G | T | . |
| **SMARCA4** | chr19 | 10991310 | 10991310 | G | T | . |
| **SMARCA4** | chr19 | 11013017 | 11013017 | G | T | . |
| **SMG1** | chr16 | 18836125 | 18836125 | T | C | . |
| **SNX25** | chr4 | 185363394 | 185363394 | G | A | . |
| **SORBS3** | chr8 | 22564469 | 22564469 | C | T | rs3758036 |
| **SPIRE1** | chr18 | 12452388 | 12452388 | G | A | rs568852514 |
| **SPRING1** | chr12 | 116737869 | 116737869 | C | G | . |
| **SPTA1** | chr1 | 158635935 | 158635935 | G | A | rs116959874 |
| **SREBF2** | chr22 | 41868716 | 41868716 | C | T | rs376619164 |
| **SSH2** | chr17 | 29648240 | 29648240 | C | A | . |

| | | | | | | |
|---|---|---|---|---|---|---|
| **ST6GALNAC5** | chr1 | 77044274 | 77044274 | G | C | . |
| **STARD6** | chr18 | 54331801 | 54331801 | C | T | rs765279796 |
| **STARD9** | chr15 | 42693273 | 42693273 | T | A | rs754927437 |
| **STAT3** | chr17 | 42333748 | 42333748 | C | T | . |
| **STK10** | chr5 | 172052954 | 172052954 | C | A | . |
| **STK26** | chrX | 132069583 | 132069583 | A | T | . |
| **STK31** | chr7 | 23729146 | 23729146 | T | G | . |
| **STYK1** | chr12 | 10629498 | 10629498 | G | A | rs34981955 |
| **SULT1A2** | chr16 | 28593476 | 28593476 | C | A | . |
| **SULT1C3** | chr2 | 108264928 | 108264928 | A | G | . |
| **SYNM** | chr15 | 99113697 | 99113697 | T | C | rs141391292 |
| **SYT17** | chr16 | 19183797 | 19183797 | C | T | rs571053295 |
| **TAS1R1** | chr1 | 6576602 | 6576602 | T | C | rs114597256 |
| **TBC1D32** | chr6 | 121161043 | 121161043 | C | T | . |
| **TBC1D8** | chr2 | 101050476 | 101050476 | G | A | rs199891498 |
| **TBL2** | chr7 | 73570706 | 73570706 | T | A | . |
| **TBX1** | chr22 | 19766863 | 19766863 | G | C | rs926713371 |
| **TEC** | chr4 | 48138937 | 48138937 | G | A | rs376411203 |
| **TECPR2** | chr14 | 102497671 | 102497671 | G | A | rs77170608 |
| **TEKT1** | chr17 | 6815298 | 6815298 | C | T | rs879790277 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **TEP1** | chr14 | 20385003 | 20385003 | C | T | rs145122151 |
| **TESPA1** | chr12 | 54963062 | 54963062 | G | C | . |
| **TET2** | chr4 | 105237058 | 105237058 | C | T | rs111678678 |
| **THADA** | chr2 | 43430286 | 43430286 | A | C | . |
| **TLR8** | chrX | 12920137 | 12920137 | T | G | . |
| **TMEFF2** | chr2 | 192191932 | 192191932 | T | G | . |
| **TMEM183A** | chr1 | 203007840 | 203007840 | T | G | . |
| **TMEM63B** | chr6 | 44150581 | 44150581 | G | C | . |
| **TMEM71** | chr8 | 132751915 | 132751915 | G | T | . |
| **TMEM72** | chr10 | 44934800 | 44934800 | C | T | rs527934388 |
| **TMUB1** | chr7 | 151081892 | 151081892 | A | G | . |
| **TOGARAM1** | chr14 | 44964361 | 44964361 | G | A | rs759654991 |
| **TP53** | chr17 | 7673802 | 7673802 | C | A | rs28934576 |
| **TP53** | chr17 | 7673806 | 7673806 | C | A | rs121912657 |
| **TP53** | chr17 | 7673823 | 7673823 | C | T | rs193920774 |
| **TP53** | chr17 | 7674872 | 7674872 | T | C | rs121912666 |
| **TRIP12** | chr2 | 229815131 | 229815131 | C | T | . |
| **TRIT1** | chr1 | 39850148 | 39850148 | C | A | . |
| **TSPAN14** | chr10 | 80516286 | 80516286 | T | A | . |
| **TTC12** | chr11 | 113366253 | 113366253 | G | A | rs138333675 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **TTC17** | chr11 | 43492078 | 43492078 | C | A | . |
| **TTN** | chr2 | 178539023 | 178539023 | C | T | rs4894028 |
| **TUBGCP3** | chr13 | 112559339 | 112559339 | G | A | rs375434522 |
| **TXLNG** | chrX | 16829616 | 16829616 | G | T | . |
| **UBQLN4** | chr1 | 156050311 | 156050311 | T | G | . |
| **UBR5** | chr8 | 102293783 | 102293783 | C | T | rs966636782 |
| **UBR5** | chr8 | 102342618 | 102342618 | C | G | . |
| **UNC13C** | chr15 | 54015404 | 54015404 | G | A | rs201822096 |
| **UNC45B** | chr17 | 35150068 | 35150068 | G | A | rs187578844 |
| **UNC45B** | chr17 | 35180654 | 35180654 | G | A | rs76329788 |
| **VAX1** | chr10 | 117136651 | 117136651 | C | A | . |
| **VCL** | chr10 | 74107316 | 74107316 | G | C | rs150385900 |
| **VIPR1** | chr3 | 42536241 | 42536241 | G | T | rs3733055 |
| **VOPP1** | chr7 | 55497685 | 55497685 | C | A | . |
| **VSX1** | chr20 | 25079442 | 25079442 | C | G | . |
| **VWF** | chr12 | 6018901 | 6018901 | G | A | rs61750100 |
| **WDR33** | chr2 | 127725144 | 127725144 | C | G | . |
| **WNT9A** | chr1 | 227925361 | 227925361 | G | A | . |
| **XDH** | chr2 | 31366042 | 31366042 | C | G | . |
| **XIRP2** | chr2 | 167246619 | 167246619 | G | A | rs181539061 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **YWHAB** | chr20 | 44906032 | 44906032 | C | T | rs374058745 |
| **ZBTB7B** | chr1 | 155014806 | 155014806 | G | T | . |
| **ZCCHC14** | chr16 | 87411864 | 87411864 | C | T | rs931933177 |
| **ZGRF1** | chr4 | 112541241 | 112541241 | G | C | rs368808540 |
| **ZNF704** | chr8 | 80641474 | 80641474 | C | A | . |
| **ZNF793** | chr19 | 37537409 | 37537409 | G | C | . |
| **ZPR1** | chr11 | 116778990 | 116778990 | C | G | . |
| **ZZEF1** | chr17 | 4074189 | 4074189 | G | A | rs201136550 |
| **FBXW8** | chr12 | 117028032 | 117028032 | C | T | rs199668216 |
| **FER** | chr5 | 108954858 | 108954858 | G | C | . |
| **FLNB** | chr3 | 58154836 | 58154836 | C | G | rs138327769 |
| **FLNC** | chr7 | 128837699 | 128837699 | G | T | . |

# DISCUSSION

Smoking has been highly associated with head and neck cancer (Lalrammawia et al., 2022). Our study showed significant risk in smoking, in both low or heavy quantities. The risk increases in a dose-dependent manner with the number of pack years, as well as the duration and frequency of cigarette smoking (Hashibe et al., 2007). This is consistent with our findings, wherein the OR increased three-fold for those with more than 70 pack years compared to those with 70 or fewer pack years. More than 70 carcinogens and heavy metals, such as mercury, cadmium, nickel, arsenic, chromium and lead are present in branded cigarettes (Khariwala et al., 2012; Raju et al., 1999; Dhaware et al., 2009; Ashraf, 2012; Janaydeh et al., 2019). Exposure to these heavy metals found in cigarettes have been associated with HNC. In a study among Tunisian population, higher concentration of nickel, cadmium, arsenic and chromium were found in HNC tumours from smokers when compared with non-smokers (Khlifi & Hamza-Chaffai, 2010). The mechanism of carcinogenesis by heavy metals involved induction of oxidative stress and inhibition of methylation, apoptosis and DNA repair (Khlifi & Hamza-Chaffai, 2010).

Moreover, tobacco filler in the local cigarette zozial has been found to contain significant amount of aluminium, manganese and silicon (Gomaa et al., 1993). In our study, majority of the patients were frequent smokers of zozial. These zozial also contain high concentrations of other heavy elements like arsenic, cobalt, copper, lead, iron, mercury and cadmium when compared with other common brands (Lalrammawia et al., 2022, Khariwala et al., 2012). Due to the low quality of the paper used to roll these zozial, smokers need to puff frequently to keep it lit, thereby increasing the inhalation of chlorine present in the bud, imposing a greater risk in developing HNC in the population (Laugesen et al., 2009). Tobacco plants used to produce Zozial are cultivated by local farmers as part of their annual farming practices. These plants naturally absorb heavy metals from the soil, which then accumulate in their leaves (Regassa & Chandravanshi, 2016). However, there are no estimation of the concentrations of the metallic compounds present in the soil or pesticides or fertilizers used where tobacco plants were grown. In our analysis, dipping (sahdah) and tuibur

showed weak associations with HNSCC. However, smokeless tobacco use has been strongly linked to oral cavity cancer (Siddiqi et al., 2015). Due to its widespread use, oral cancer is a significant burden in India (Wyss et al., 2016). The low odds ratio (OR) observed in our study might be attributed to the infrequent use of smokeless tobacco compared to smoking among both cases and controls in our small sample set. One of the risk factors of oral cancer is consumption of betel nut (areca nut), which contains carcinogens that can induce the production of Reactive Oxygen Species (Alsahafi et al., 2019; Wyss et al., 2016). Our study showed a high OR of 1.218 (CI 0.594–2.497) for patients consuming areca nut (76/100), however, non-significant. From the evidences in other studies, the impact of areca nut and smokeless tobacco are directed more towards oral cavity as it is the primary site of exposure (Lee et al., 2019). One of the limitations in the study was the small sample size, which restricts site-wise analysis hindering the study of direct impact of smokeless tobacco and areca nut in the oral cavity cancer cases alone.

Alcohol consumption has been linked with several cancers like liver, gastric and oesophageal cancer (Sung et al., 2021). Mechanism of carcinogenesis by alcohol have been fully explained (Brooks et al., 2014; Marziliano et al., 2020). Alcohol is converted to acetaldehyde, a carcinogenic compound that can disrupt the stability of DNA (ref). In a case-control study, alcohol intake was categorized by the number of drinks per day (One drink is comparable to 30 ml of spirits), the OR increases from 2.1 to 5.0 to 12.2 21.1 for drinkers of 3-4, 5-7, 8-11 and > 12 drinks/day, respectively with <2 drinks/day as reference (Altieri et al., 2004). Similarly, in a meta-analysis, which categorized alcohol intake into light (≤12.5 g per day), moderate (≤50 g per day), and heavy (>50 g per day), the risk of oral cavity and pharyngeal cancers increased with higher consumption showing estimated risk ratios of 1.13, 1.83, and 5.13, respectively (Bagnardi et al., 2015). Based on the frequency of drinks per day, a study in Taiwan reported an increased risk associated with higher frequency of alcohol consumption, categorized into monthly, weekly, and daily intake (Huang et al., 2017). Similar trends in these studies are also observed in our study where alcohol showed a significant risk factor in a dose dependent manner. The drawback is that the absence of precise processing hinders the option of estimating the amount of alcohol in these

‘Local’ drinks, which were more frequently consumed by the patients. Regardless of the quantity and/or duration of alcohol exposure, there are high chances that other harmful compounds might be present that increases the risk not only in head and neck cancer but also in other sites and diseases. Further studies on local drinks are imperative to confirm this assumption and raise public awareness.

The association between environmental factors like alcohol and smoking with HNSCC has been extensively investigated (Miranda-Galvis et al., 2021). Our study showed that apart from smoking and alcohol, first-degree FHC might be a risk factor for HNSCC in the population. First-degree FHC among oral cavity cancer patients have been reported to be associated with an OR of 1.9 (95% CI 1.2–2.8) in an ICARE study conducted in France (Radoï et al., 2013). The associations of smoking, alcohol and FHC have been reported in several studies. In a case-control study of 8967 cases conducted by International Head and Neck Cancer Epidemiology (INHANCE) consortium, an association of smoking, alcohol and FHC with an OR of 7.2 (95% CI 5.5–9.5) among patients with FHC who were smokers as well as alcohol consumers was found (Negri et al., 2009). A strong association (OR: 2.27, 95% CI: 1.26–4.10) between HNC in patients <45 years and FHC was also observed in a 25 case-control study (Toporcov et al., 2015). Evidence suggesting a potential familial component in head and neck cancer has been reported, based on factors related to tobacco and alcohol metabolism, cell cycle regulation and DNA repair pathways. A germline *CDKN2A* mutation, leading to a premature stop codon, was detected in a 48-year-old proband with hypopharyngeal cancer (Cabanillas et al., 2013). An increased mutational burden in three genes in FANC pathway namely *FANCL*, *FANCE* and *FANCD2* have been identified among 417 HNC patients (Chandrasekharappa et al., 2017). Polymorphisms in the *RAD51* and *XRCC3* genes have also been reported to increase the risk of head and neck cancer by 2.5-fold and 16-fold, respectively (Kayani et al., 2014). In a study conducted in the North Eastern India, *XRCC1* and *XRCC2* polymorphisms and tobacco consumption together was found to contribute to the susceptibility of HNSCC (Choudhury et al., 2014). Other metabolic enzymes such as *ALDH2*, *Cytochrome p450* and *GST* genes have been reported to increase the risk of HNC (Venugopal et al., 2017).

In Figures 7 (a) & (b) the frequency of FHC is observed in approximately half of the cases across all ranges of alcohol consumption and smoking durations, including non-drinkers and non-smokers. Additionally, the number of FHC cases is higher in the oral cavity and hypopharynx compared to other sites [Figures 8 (a) & (b)]. Smoking and alcohol consumption have been practiced for many generations and remain prevalent among both men and women in the state. In addition to these practices, the incidence of HNSCC is high among patients with a family history of cancer, suggesting a potential influence of family history on susceptibility to the disease, regardless of alcohol and smoking exposure. Exposure to risk and environmental factors over many generations in a small endogamous population may have led to genetic alterations that predispose the population to cancer, with a notable impact among patients with a family history of the disease. This hypothesis needs to be validated through genetic mutation studies within these families. Additionally, investigating polymorphisms in the metabolic pathways of tobacco and alcohol could clarify the associated risks and how they might influence the familial tendency toward head and neck cancer in this population.

A retrospective study was conducted to provide valuable insights into the various treatment modalities and factors influencing the 2-year survival outcomes of patients with HNSCC from Mizoram State Cancer Institute (MSCI). The study found that the 2-year OS rate was 78.1%, which was higher than the 2-year PFS rate of 57.4%. The analysis revealed several factors that impacted survival outcomes, including TLC, ANC, N2 nodal stage, and the specific cancer site, with laryngeal cancer being particularly notable. Of the 210 patients, 188 (89.5%) underwent a multi-modality treatment approach, while 86 (41.0%) primarily received CCRT, followed by 85 (40.5%) patients treated along with IC and 17 (8.1%) patients underwent surgery along with CCRT and/or IC. Single modality approach involving radiotherapy alone was given to 22 (10.5%) patients. CCRT showed survival advantages compared to other treatment modalities in both OS and PFS, though the difference was not statistically significant.

A similar study conducted among the Indonesian population by Irawan and colleagues observed a 2-year PFS rate comparable to ours (50%) (Irawan et al., 2022).

In our cohort, the 2-year OS rate was nearly equivalent to that of a Korean cohort, where the 2-year OS rate was reported at 79.8% (Zhang et al., 2015). However, a study conducted in northern India by Badola and colleagues at a tertiary cancer care centre reported a lower 2-year survival rate of 58.8% at 18 months of follow-up (Badola et al., 2023). The study results also showed that patients who underwent IC had less favourable PFS compared to those who did not receive IC. Additionally, despite a nearly equal number of patients receiving either CCRT alone or IC followed by CCRT, the OS and PFS rates were poorer for those who received IC. The purpose of IC is typically to reduce tumour size or increase tumour sensitivity to radiotherapy, suggesting that patients receiving IC should gain an advantage. However, the results indicate a different outcome, leaving the benefits of IC in the management of head and neck cancer uncertain (Lim et al., 2021). Several randomized trials have consistently shown no significant difference in outcomes between IC followed by CCRT and CCRT alone in patients with HNC (Haddad et al., 2013; Luo et al., 2022; Budach et al., 2016; Cohen et al., 2014; Hitt et al., 2021). The selection of a treatment plan relies on the patient's body weight, comorbidities, location, size and nodal involvement of the tumour. The primary goals of curative treatment are tumour reduction and organ preservation (Zhang et al., 2015; Hanna et al., 2013; Gau et al., 2019; Rana et al., 2020). The poor response to IC may be attributed to residual toxicity, as this treatment is often administered to patients with higher T and N staging (Zhang et al., 2015). Additionally, increased nodal involvement has consistently been shown to be a significant predictor of poor response. In our study cohort, 70.4% of patients had neck nodal involvement at the time of diagnosis, which could be a contributing factor to the poor response to treatment. Within 2 years, 85 patients were classified as poor responders (1 with stable disease and 84 with progressive disease), with 32 of these patients progressing to local recurrence, regional metastasis, or distant metastasis. Neck nodal involvement has been strongly associated with poor survival and recurrence (Cho et al., 2009; Xing et al., 2016). A randomized Phase III trial by Cohen and colleagues reported that induction chemotherapy did not improve OS compared to concurrent chemoradiotherapy alone in patients with N2 and N3 HNSCC (Cohen et al., 2014).

Our study cohort revealed that leukocytosis and neutrophilia may be significant predictors of PFS and OS. This is in consistent with other studies where leukocytosis was found to predict OS and PFS in HNSCC patients treated with concurrent cisplatin and radiotherapy (Millrud et al., 2012; Schernberg et al., 2018). Leukocytosis has also been linked to post surgery metastasis and tumour recurrence in OSCC (Schernberg et al., 2018; Chen et al., 2014; Roh et al., 2019; Gouw et al., 2018). Jensen and colleagues have also shown that neutrophilia and leukocytosis in pre-treated patients had poor responses to radiotherapy (Jensen et al., 2017). Leukocytosis and neutrophilia have been found to be predictors of poor PFS and OS in several cancers like oesophageal, anal and lung cancers (Schernberg et al., 2017;2018).

Smoking and alcohol consumption are well-established risk factors for HNSCC (Su et al., 2016). Several studies have shown that smoking decreases 2-year PFS in HNSCC patients (Lee et al., 2020, Espeli et al., 2012; Thompson et al., 2011). Alcohol consumption has been observed to negatively impact OS and increase the risk of mortality in patients, regardless of whether they continued drinking or quit (Ferraguti et al., 2022; Abrahão et al., 2020; Denissoff et al., 2022). However, alcohol and smoking did not significantly impact the PFS or OS in our study. A study done by Su and colleagues indicated that a history of betelnut consumption combined with smoking was linked to poor prognosis in patients with HNSCC (Su et al., 2016). Despite 81.4% of patients having a history of betel nut chewing, we did not observe a significant association between betel nut chewing and PFS and OS. Similarly, the consumption of smokeless tobacco in the form of 'tuibur,' a common practice, was not associated with prognosis in our study. Although having a family history of cancer is known to increase the risk of developing HNSCC, it did not affect the treatment response in our study. This finding aligns with the study by Getz and colleagues, which observed a similar hazard ratio between family history of cancer and survival (Getz et al., 2017).

This study has several limitations, including a small sample size that prevented adequate stratification by cancer sites or stages, thereby limiting the statistical power of the analysis. Additionally, the retrospective nature of this study restricted the collection of direct information on patients' quality of life, diagnosis, and

comprehensive reports on their overall well-being, including toxicity profiles, which could introduce confounding by indication. Moreover, data on the presence of human papillomavirus (HPV) or Epstein–Barr virus (EBV) were not available, as these tests are not routinely conducted in the state. Furthermore, some parameters had missing details that could not be traced. In addition, this is a preliminary and exploratory study that has many shortcomings such as the selection of variables for multivariate models based on univariate analysis weakens the statistical power (Sun et al., 1996; Heinze et al., 2017). The findings of this study are tentative and necessitate further in-depth investigation to reach more definitive conclusions. Despite these limitations, the study's methodology and objectives are applicable to data from clinical investigations in remote, resource-limited cancer care centers.

To our knowledge, this is the first survival analysis of HNSCC conducted in a region with high cancer prevalence within the country. The 2-year OS and PFS rates were found to be 78.1% and 57.4%, respectively. The study revealed that the multi-modality approach, particularly concurrent chemoradiotherapy (CCRT), provided a survival advantage over other treatment methods, including sequential therapy. Poor prognosis was linked to factors such as elevated TLC, high ANC, nodal involvement and laryngeal cancer site.

Functional enrichment analysis showed that in biological process, acute phase response (APR) pathway was significantly enriched in our study. The APR is a systemic reaction triggered by injury, infection or any disruption of homeostasis, such as neoplastic growth (Baumann & Gauldie, 1994; Moreland, 2004). Local inflammatory cells, including macrophages and neutrophils, release pro-inflammatory cytokines such as tumour necrosis factor-alpha (TNF-α), interleukin-1 (IL-1) and interleukin-6 (IL-6) into the bloodstream and stimulates hepatocytes in the liver to synthesize and release various APR (Ehlting et al., 2021). Based on the serum concentration during inflammation, APR can be classified into positive or negative (Gulhar et al., 2024). Positive APRs are upregulated during inflammation. These include C-reactive protein, procalcitonin, fibrinogen, ferritin hepcidin and serum amyloid A. Negative APRs are downregulated reactants which includes transferrin, retinol-binding proteins, albumin, prealbumin and antithrombin. Among the APRs,

serum amyloid A family proteins are the most prominent (Lee et al., 2021). There are four isoforms of SAA, of which SAA1 and SAA2 collectively referred to as SAA are the predominant isoforms mainly released during Acute Phase Response. SAA1 encodes a pre-protein of 122 amino acids having an 18 amino acid signal peptide, which upon cleavage results in a mature SAA1 having 104 amino acids (Sun et al., 2016). One of the mechanisms by which SAA1 promotes metastasis in cancer is by facilitating the formation of a pro-metastatic niche. SAA1 functions as a chemoattractant, promoting tissue infiltration by monocytes and neutrophils (Badolato et al., 1994). SAA1 facilitates the recruitment of myeloid cells to the liver, leading to cellular remodelling that establishes a pro-metastatic niche. This environment supports the seeding and colonization of disseminated tumour cells (DTCs) thus leading to metastasis (Lee et al., 2019; Chin & Wang, 2016). Another study by Niu et al., 2022 demonstrated that SAA1 induces the accumulation of immunosuppressive neutrophils cytokine through the activation of toll-like receptor 2 (TLR) mediated signalling pathways (Niu et al., 2021). SAA activates PI3/NF-kB signalling pathway through TLR2/MYD88-mediated pathway and also promotes neutrophil apoptosis resistance, thereby results in progression of cancer (Niu et al., 2021). Our study found that patients who had downregulated SAA1 proteins during treatment when compared with pre-treated samples had better OS and PFS rates. Prognostic implications of SAA have also been observed in many cancers. SAA1 protein has been found to be elevated in nasopharyngeal cancer and is a potential biomarker to monitor tumour relapse (Cho et al., 2004). High serum levels of SAA have also been linked to liver metastasis in non-small cell lung cancer and colorectal cancer (Lee et al., 2019). SAA was also associated with poor treatment outcomes in newly diagnosed advanced pancreatic cancer (Wattenberg et al., 2021). SAA levels also associated with outcomes of patients having advanced melanoma (Findeisen et al., 2009). The prognostic role of SAA1 in predicting metastasis in hepatocellular carcinoma (HCC) with worse OS in higher SAA1 expression patients (log-rank test, p-value $< 0.001$) was studied by Li et al. (2023).

In our study, we have also observed that patients having down-regulated β2-microglobulin (β2M) expressions had better PFS than patients with up-regulated or no

change in β2M expression in treated samples compared to treatment naïve samples. β2M is a 99 amino acid well known housekeeping protein with a molecular weight of 12-kDa (Nomura et al., 2014). β2M is a light chain subunit of (MHC) Class I antigen present on the surface of all nucleated cells. β2M plays a crucial role in regulating host immune recognition by CD8+ T lymphocytes as well as in immunoglobulin transport and iron metabolism. β2 -microglobulin (β2M) functions as a factor that promotes growth, angiogenesis, EMT, and bone metastasis, and acts as a prognostic indicator in solid tumour cells (Nomura et al., 2014). Consistent with our study, a significantly increased survival rate was observed in OSCC patients with decreased β2M expression in a Taiwanese population (Chen et al., 2008). In another study the β2M expression between metastatic and primary OSCC patients was compared and it suggested that β2M may contribute to tumour invasion and metastasis, as its expression was significantly higher in metastatic OSCC (Jiang et al., 2012). Numerous studies have shown that serum β2M levels are among the most important independent prognostic factors and survival predictors for certain tumours such as T cell lymphoma, multiple myeloma and renal cell carcinoma (Nomura et al., 2014). Elevated serum β2-microglobulin (β2-m) level is an independent predictor of poor outcome in OS and PFS in multiple myeloma patients undergoing chemotherapy and stem cell transplantation (Nomura et al., 2014).

We also observed that among the 78 significantly differentially expressed proteins, patients with changes in serum Haemoglobin Subunit Beta (HBB) expression during treatment—either upregulation or downregulation compared to pre-treatment levels—had worse progression-free survival than those whose expression remained unchanged. HBB is a globin protein, structurally conserved group of proteins that typically contain a heme group and plays an essential role in the formation in oxygen transportation in the blood (Bonaventura et al, 2013). HBB has been found to be expressed in other cells besides erythrocytes including cancer cells. The association between HBB and cancer has been reported multiple times; however, the molecular mechanism through which HBB contributes to tumour progression remains unclear. HBB expression is significantly elevated in invasive breast carcinoma compared to carcinoma in situ and is absent in normal breast epithelium. It has been shown to drive

proliferation, migration, metastasis and tumour-mediated angiogenesis in breast cancer cells (Ponzetti et al., 2017). Similarly, increased tumour progression has been reported in cancers such as prostate, lung, and cervical cancer (Zheng et al., 2017; Li et al., 2013). However, Kang et al. (2022) found that HBB suppresses proliferation in non-small cell lung carcinoma via the MAPK and JNK pathways, with its growth-inhibiting effects also observed in thyroid cancer and neuroblastoma (Kang et al., 2022; Maman et al., 2017; Onda et al., 2005).

This study identified 78 significantly differentially expressed proteins in HNSCC samples during treatment, using treatment-naive samples as baseline. Functional analysis revealed that the blood microparticle and acute phase response pathways were the most significantly enriched. Notably, changes in SAA1, β2M, and HBB proteins were significantly associated with patient outcomes, suggesting their potential as biomarkers for HNSCC in this population. Further validation in larger cohorts is needed to confirm these findings.

To our knowledge, this is a pilot study on hypopharyngeal cancer in this population. It aims to explore the genomic landscape of hypopharyngeal cancer through WES of tumour tissues and matched blood samples from ten patients. To date, there have been limited WES studies on hypopharyngeal cancer alone. In comparison to the few available studies, we observed that the most frequently mutated genes in our population differ from those in other cohorts. Four genes (*FSIP2, HMCN2, MUC3A* and *ZNF705E*) were the most frequently mutated genes present across all ten of our samples, while *TP53, BRCA2* and *MUC16* were the top mutated genes in the TCGA cohort of nine hypopharyngeal cancer patients. In contrast to our findings, Yao et al. reported *TTN, TP53* and *ANK3* as the most frequently mutated genes, while another study by Machnicki et al. identified *TP53, FAT1* and *NOTCH1* as the most mutated in another cohort (Yao et al., 2023; Machnicki et al., 2022). *FSIP2* gene encodes Fibrous sheath interacting protein 2 is an essential gene for spermatogenesis (Martinez et al., 2018). This gene has not been directly linked with HNC, but mutations in this gene have been associated with oesophageal cancer, gastric cancer, cutaneous melanoma and breast cancer (Tang et al., 2021; Wang et al., 2021; Ying et al., 2021; Lefebvre et al., 2016). Changes in its expression have been linked with renal cell carcinoma (Zhang

et al., 2020). Amplification of *FSIP2* gene and *FSIP2* - *ALK* fusion have been identified in germ cell tumours and lung adenocarcinoma, respectively (Litchfield et al., 2015; Zhao et al., 2021). Among the top 50 mutated genes in our cohort, *TP53* and *NOTCH1*, both well-known cancer-associated genes were mutated in 80% of the samples and were also among the top mutated genes in TCGA cohort. These two genes have been found to be highly mutated in all types of cancers including HNC (Manda et al., 2024; Stransky et al., 2011; Leemans et al., 2018).

The ECM receptor interaction pathway and motor protein pathway were the most significantly enriched in our study (p-value < 0.000). These pathways interact dynamically and are essential for cellular functions, playing a critical role in cancer biology. The interaction between cancer cells and the tumour microenvironment (TME) leads to phenomena such as ECM stiffness and remodelling (Huang et al., 2021; Yuan et al., 2023). Dysregulation of the ECM contributes to uncontrolled proliferation by activating multiple signalling pathways, including the PI3K-AKT pathway, which was also significantly enriched in our study. ECM stiffness accelerates cancer cell migration through signalling pathways like PI3K-AKT, which downstream activates the protein AP1, promoting cell migration. Additionally, ECM stiffness enhances vascularization, leading to angiogenesis (Huang et al., 2021; Winkler et al., 2020). Stiffened ECM also promotes radioresistance and reduce the efficacy of drugs leading to chemoresistance (Huang et al., 2021; Darvishi et al., 2022). PI3K-Akt signalling pathway is one of the downstream pathways of ECM receptor interaction pathway. It is a signal transduction pathway that when activated regulates cell cycle and apoptosis. In a WES analysis of 106 HNSCC patients, PI3K-Akt was the most mutated pathways and demonstrated that tumours having mutations in this pathway were likely to have higher rate of mutations in other known cancer genes (Lui et al., 2013). HPV infection is one of the risk factors of HNC including hypopharyngeal cancer (Patel et al., 2022). However, we do not have information on the presence of absence of HPV as routine tests are not done. Mutations in ATP-binding cassette (ABC) transporters contribute to tumourigenesis by disrupting the Reactive Oxygen Species (ROS) homeostasis leading to DNA damage, modulate angiogenesis, promoting cell migration, metastasis and have also been linked to chemotherapy

resistance (Duvivier et al., 2023; Sun et al., 2012). It is noteworthy that the coagulation and complement cascade pathway is significantly enriched in the exome data from ten hypopharyngeal cancer samples as well as the differentially expressed proteins in the study. The coagulation cascade involves a series of clotting factors activated to form blood clot in response to bleeding or inflammation caused by tissue injury (Palta et al., 2014). The complement system is a cascade of enzymes activated in defence against infection by activating local inflammatory responses (Janeway et al., 2001). The complement and coagulation cascades play a critical role in cancer progression by modulating immune responses and promoting tumour microenvironment, contribute to tumour growth and metastasis by suppressing anti-tumour immunity and facilitating processes such as cell proliferation, migration and inflammation (Lima et al., 2013; Revel et al., 2020). Complement activation within tumours, in particular, recruits immunosuppressive cells like myeloid-derived suppressor cells (MDSCs) and promotes angiogenesis, aiding tumour development (Zhang et al., 2019; Afshar-Kharghan, 2017). Furthermore, the interplay between these systems, as they share common pathways can support tumour survival and growth, highlighting their potential as therapeutic targets in cancer treatment (Pryzdial et al., 2022). Changes in SERPINE1 expressions have been associated with poor outcome and higher risk of metastasis in HNSCC patients (Pavón et al., 2015; 2016). In a PAN Cancer Analysis, Lawal et al demonstrated that C3 and C5 were highly mutated in different types of cancer including head and neck cancer. Anaphylatoxins C3a and C5a (products of C3 and C5) are key players in tumour-specific immunity and influencing clinical responses (Lawal et al., 2021).

We have observed that the tumour mutational burden of these 31 variants were more clustering in T7, T8, T9 and T10 where alcohol, smoking and tobacco consumption were more among these patients. From the thirty-one variants present in 4 or more samples out of 10 and predicted deleterious, the variant rs79747830 (NC_000007.13:g.33054388T>C) in gene *NT5C3A* was found in 7 out of 10 of our samples. This variant was previously reported in HNSCC patients from Pakistan (Ghias et al., 2019). A single nucleotide polymorphism was identified in four samples in the *KMT2C* gene, which encodes Histone-Lysine N-Methyltransferase 2C. A recent

systematic review investigating the role of the KMT2 methyltransferase family in HNC found that *KMT2D* was the most frequently mutated gene, followed by *KMT2C* (da Silva Santos et al., 2024). Mutations in one or both of these genes, along with other oncogenes contribute to Tumour progression and poor outcomes. Another study by Machnicki revealed that *KMT2C* mutations were more common in hypopharyngeal cancer compared to other head and neck cancer sites (Machnicki et al., 2022). They also demonstrated that the *KMT2C* gene may act as a tumour suppressor, as CRISPR-Cas9-induced *KMT2C* loss of function led to increased proliferation in FaDu cell lines. In another study similar to our study, *KMT2C* was found to be mutated in all ten samples that were screened (Yao et al., 2023).

To date, studies on the genomic landscape of hypopharyngeal cancer alone still remain scarce. In this study, we used WES to analyse ten hypopharyngeal tissues with matched blood samples to explore the genetic makeup of hypopharyngeal cancer in the Mizo population. The top ten most frequently altered genes were *FSIP2, HMCN2, MUC3A, ANF705E, ATP5F1A, CCDC187, FOXD4L4, LRP2, MUC12* and *MUC16*. Additionally, 850 variants were predicted to be pathogenic, with *NT5C3A* present in 70% of the samples. Pathway analysis revealed the ECM receptor interaction as the most significantly enriched pathway. Notably, the complement and coagulation cascade pathway were enriched in both this cohort and the proteomics set. Several genes not previously reported in other hypopharyngeal cancer cohorts were identified. Further screening and functional validation are needed to provide more detailed insights.

# SUMMARY

- This study aims to identify the risk factors and their association with survival outcome of HNSCC in Mizo population, identify potential prognostic protein biomarkers that are differentially expressed in serum of patients during treatment compared with treatment naive samples and explore the genomic landscape of Hypopharyngeal Cancer using Whole Exome Sequencing (WES).
- Multivariate Regression analysis showed male patients had significantly higher Odds Ratio (OR = 6.694, p-value = <0.05) compared to female patients. Patients above 45 years of age were likely to develop Head and Neck Squamous Cell Carcinoma (HNSCC) compared to patients below 45 years of age (OR = 3.979, p-value = < 0.05).
- Smoking and consumption of alcohol were also significantly associated with HNSCC compared to non-smokers and non-drinkers. The association increases as level of smoking and consumption of alcohol increases.
- Patients having first degree family history of cancer were found to be significantly associated with HNSCC as well (OR = 1.921, p-value = 0.037)
- Smoking has been highly associated with HNSCC, with increased risk in a dose-dependent manner. Majority of the patients in our study smoked local made Zozial which was found to contain high levels of heavy metals (Lalrammawia et al., 2022; Hashibe et al., 2007)
- Significant associations between Head and Neck Cancer and First-Degree Family History of Cancer in combination with smoking and alcohol consumption have also been reported (Radoï et al., 2013; Negri et al., 2009; Toporcov et al., 2015)
- The overall 2-year overall survival (OS) was 78.1% and Progression Free Survival (PFS) was 57.4% which are comparable to other studies (Irawan et al., 2022; Zhang et al., 2015; Badola et al., 2023).
- Patients who received Induction Chemotherapy (IC) had lower PFS rate (47.3%) than patients who did not received IC with a significant log-rank test.

- Patients who had Leukocytosis [Total Leukocyte Count (TLC) > 10 thou/cumm] had lower OS rate (58.4%) compared to patients with TLC ≤ 10 thou/cumm.
- Patients who had Neutrophilia [Absolute Neutrophil Count (ANC) > 7 thou/cumm had significantly lower OS and PFS rates (57% and 36.4%, respectively) with log-rank tests p-value 0.014 and 0.043, respectively.
- Multivariate Cox Proportional Hazard showed that among all the sites, Laryngeal cancer was a strong predictor of OS (HR = 5.165, p-value = 0.009) and PFS (HR = 2.844m p-value = 0.028). Nodal involvement (N2), Leukocytosis and Neutrophilia were also significant predictors of OS and PFS.
- Induction Chemotherapy did not improve OS compared to CCRT in patients with N2 and N3 involved HNSCC (Cohen et al., 2014).
- Leukocytosis is a predictor of OS and PFS in patients treated with concurrent Cisplatin and Radiotherapy and linked with tumour recurrence (Schernberg et al., 2018).
- Patients with Leukocytosis and Neutrophilia has been shown to have poor response to radiotherapy (Jensen et al., 2017). They are also found to be predictors for poor OS and PFS in other cancers (Schernberg et al., 2017;2018)
- Out of 134 differentially expressed proteins found in our samples, 78 were significantly expressed 304 times during treatment when compared with treatment naïve samples, of which 155 were down regulated and 149 proteins were upregulated.
- Patients having downregulated SAA1 protein showed better OS and PFS compared to patients with upregulated or baseline SAA1 expressions.
- Patients with downregulated ß2M expressions had better PFS compared to patients with upregulated or baseline ß2M expressions.
- Patients with consistent HBB expressions had better PFS compared to patients with change in HBB expressions.
- The functional enrichment analysis of the 78 differentially expressed proteins showed that Acute Phase Response Pathway was significantly enriched. SAA1 protein play a huge role in Acute Phase Response facilitating pro-metastatic

niche which helped in colonizing of disseminated tumour cells thus leading to metastasis (Lee et al., 2019; Chin & Wang, 2016).

- SAA1 protein also induces accumulation of immunosuppressed neutrophils through Toll-like receptor 2 (TLR) mediated pathway and promotes neutrophil apoptosis resistance through TLR2/MYD88 mediated pathway, both of which leads to progression of cancer (Niu et al., 2021).
- β2M and HBB protein expressions have been linked with metastatic cancer and have been reported to contribute to tumour invasion, migration, angiogenesis, EMT and bone metastasis but the molecular mechanisms involved have not yet been fully explored (Nomura et al., 2014; Jiang et al., 2012; Ponzetti et al., 2017; Kang et al., 2022).
- WES in ten hypopharyngeal cancers identified 62,403 exonic variants in 12,122 genes where the top 10 most frequently mutated genes are *FSIP2*, *HMCN2*, *MUC3A*, *ANF705E*, *ATP5F1A*, *CCDC187*, *FOXD4L4*, *LRP2*, *MUC12* and *MUC16*.
- The significantly enriched pathways which have been associated with cancer are the ECM receptor interaction, Human Papillomavirus infection, ABC transporters, Complement and Coagulation Cascade, Small Cell Lung Cancer and P13K-Akt Signalling Pathways.
- Out of 29,279 exonic non-synonymous across 8333 genes, about 850 variants in 609 genes were predicted to be deleterious by SIFT, Polyphen, Mutation Taster and CADD. Out of these 850 variants, 107 variants in 31 genes were found to be altered in ≥3/10 samples.
- Mutations in the *NT5C3A* gene were identified in 7 out of 10 hypopharyngeal cancer cases, followed by mutations in *MTMR4* and *AP1G2*, found in 5 out of 10 samples. Variants of *AZIN2, IRX6, KMT2C, NUDT12, POP5*, and *SHANK2* were present in 4 out of 10 samples. Notably, samples T7, T8, T9, and T10 exhibited a higher number of mutations compared to the others.
- There is very limited information on WES of hypopharyngeal cancer till today. The top most frequently mutated genes are inconsistent with top genes

identified in few available studies (TCGA, Yao et al., 2023; Machnicki et al, 2022).

- KMT2C gene mutations have been reported in Hypopharyngeal cancer, contributing to tumour progression when altered with other oncogenes. Its role in tumour suppression have also been demonstrated were loss of function led to increased proliferation in hypopharyngeal cell lines (Yao et al., 2023; Machnicki et al, 2022).
- This study identified that patients having first degree of family history of cancer, smoking and alcohol consumption were significantly associated with HNSCC. Survival analysis revealed that nodal involvement, leukocytosis and neutrophilia were predictors of OS and PFS. Also, patients who received IC had worse outcome compared to patients who did not receive.
- Changes in expressions of SAA1, ß2M and HBB proteins during treatment may serve as prognostic biomarkers of HNSCC.
- WES revealed that the top mutated genes differed from those identified in other hypopharyngeal cancer cohorts from different populations. Mutations in the *KMT2C* gene were found in our samples and have been reported in other hypopharyngeal cancer cohorts, highlighting the need for further investigation into its role in this cancer. We also identified mutations in genes not previously reported in hypopharyngeal cancer.

# APPENDIX 1. Ethical Clearance

**IEC, CIVIL HOSPITAL, AIZAWL.**
**COMMUNICATION TO THE PRINCIPAL INVESTIGATOR/GUIDE BY THE MEMBER SECRETARY, INSTITUTIONAL ETHICS COMMITTEE**

**No.B.12018/1/13-CH(A)/IEC/69** **Dated: 18th June, 2018**

**To,**

***Dr. N. Senthil Kumar***
*Professor and Head, Department of Biotechnology*
*Mizoram University*

**Subject:** Ethics Committee Approval for the referenced projects.

Reference: ***"Studies of a set of predictive and prognostic biomarkers in head and neck cancer"***

Dear, ***Dr. N. Senthil Kumar***

With reference to submission of document for review and approval to conduct the above mentioned study. The Ethics Committee has reviewed and approved the study documents as mentioned below:

1. Curriculum Vitae of non Civil Hospital, Aizawl Investigators
2. Brief description of proposal/summary
3. Copy of the Protocol/Project and questionnaire (if any)
4. Copy of Patient information sheet & Consent form in local language
5. Copy of Clinical trial agreement
6. Copy of PI undertaking

The following members of the Ethics Committee were present at the meeting held on date 14th June, 2018 at time 1:30 pm at Medical Superintendent Hall, Civil Hospital, Dawrpui, Aizawl, Mizoram - 796001.
The quorum met as per ICG-GCP and schedule Y guidelines as mentioned below:

| S. No | Name of the member | Qualification | Designation in the Ethics Committee | Gender |
|---|---|---|---|---|
| 1. | Dr. Lal Biakkima | Director H&ME (Rtd) | Chairman | Male |
| 2. | Dr. T. Lalzawmliana MD | HoD, Dept. of Biochemistry | Member | Male |
| 3. | Dr. Mary Muanpuii Ralte MD | HoD, Dept. of Anesthesiology | Basic Medical Science | Female |
| 4. | Rev R. Lalchangliana | Pastor | Member | Male |
| 5. | Dr. Zoengpari | Associate Professor, MZU | NGO/Social activist | Female |
| 6. | Pu Rosangzuala Ralte | Advocate (Legal Expert) | Member | Male |
| 7. | Dr. Saia Chenkual | HoD, Dept. of Surgery | Member | Male |

Please note that this Ethics Committee is constituted as per schedule Y, ICH-GCP, applicable local laws and regulatory requirement.

We approve the project to be conducted in its present form. Ethics Committee expects to be informed about:

1. Any SAE occurring in the course of the study
2. A copy of final individual center report

We hereby confirm that neither you nor your study members have participated in the voting/decision making procedure of the Ethics Committee.

Yours sincerely,

T. Zawmliana 15/6/18

-Sd-
**(DR. LAL BIAKKIMA)**
**Chairman**
**Institutional Ethics Committee**
**Civil Hospital, Dawrpui, Aizawl**
**Mizoram - 796001**

Chairman
Institutional Ethical Committee
Civil Hospital, Aizawl

**(DR T. LALZAWWMLIANA)**
**Member Secretary**
**Institutional Ethics Committee**
**Civil Hospital, Dawrpui, Aizawl**
**Mizoram - 796001**

Member Secretary
Institutional Ethical Committee
Civil Hospital, Aizawl

**Note :** Since Dr. C. Lalchhandama, Member Secretary is the Co-Principal Investigator for this study, this committee unanimously appointed Dr. T. Lalzawmliana (Member, IEC) as Member Secretary for this meeting.

# APPENDIX 2. Questionnaire

**QUESTIONNAIRE FOR EPIDEMIOLOGICAL, CLINICOPATHOLOGICAL AND MORPHOLOGICAL STUDIES OF HEAD AND NECK CANCER IN MIZO POPULATION**

---

Name (Hming): _____Age:_____

Corresponding address: _____

Permanent address: _____

Mob: _____ Mob (Alt): _____

Case (Primary site) _____

File numbers: _____

Block No _____ MSCI File _____ MZU No. _____

LIFESTYLE HABITS

A. TOBACCO IN THE FORM OF SMOKING (MEIZIAL)

| Type of smoke | Duration | Units per day |
|---|---|---|
| | | |
| | | |

1) How soon after you wake up do you smoke your first cigarette?
(I thawh veleh engtikah nge I zuk hmasak ber?)
- Within 5 minutes ☐
- 6 - 30 minutes ☐
- 31 - 60 minutes ☐
- After 60 minutes ☐

2) Do you find it difficult to refrain from smoking in places where it is forbidden ( e.g., in church, hospital, library etc.)?
(Meizuk phallohna hmunah I awm hian meizuk insum har I ti em?)
- Yes ☐
- No ☐

3) When would you hate most to give up smoking?
(Engtik lai ber nge meizial zuk insum har I tih ber?)
- Morning (Zingah) ☐
- Other (A dang) ☐

4) How many cigarettes per day do you smoke?
(Ni khatah meizial engzat nge I zuk thin?)
- 10 or less (10 aia tlem) ☐
- 11 - 20 ☐
- 21- 30 ☐
- 31 or more (31 aia tam) ☐

5) Do you smoke more frequently during the first hours after waking than during the rest of the day? (Nileng leh zana I meizial zuk aiin I thawh atanga chawei hma thlengin I zuk a tam zawk em?)
- Yes ☐
- No ☐

6) Do you smoke if you are so ill that you are in bed most of the day?
(Khum beta I damloh changin I zu tho em?)
- Yes ☐
- No ☐

B. TOBACCO IN THE FORM OF SNUFF/TOBACCO-SMOKE INFUSED WATER (HMUAM CHI)/INGESTED GUTKHA (EI CHI)

| Type | Duration/Quit already | Units per day |
|---|---|---|
| Sahdah/Khaini/Raja | | |
| Tuibur | | |
| Kuhva Hring | | |
| Zarda pan/Mitha pan | | |
| Shikhar (Gutkha products) etc | | |

1) How many pouches of snuff do you typically use each week? (Kar khatah fun engzat nge sahdah I hmuam tlangpui thin?)
- 1 or less ☐
- 2-4 ☐
- 5 or more ☐

How many litres of tuibur do you typically use each week? (Kar khatah tuibur engzat nge I hmuam tlangpui thin?)
- 1 Koinonia bottle (150ml) or less ☐
- 160ml to 300ml (=2x150 ml koinonia) ☐
- More than 300ml ☐

How many pouches of gutkha do you typically use each week? (Kar khatah fun engzat nge I ei tlangpui thin?)
- 1 or less ☐
- 2-4 ☐
- 5 or more ☐

2) How often do you pinch/dip in a week? (Eng anga zingin nge I hmuam thin karkhatah?)
- 1 day each week or less ☐
- 2-5 days each week ☐
- 6-7 days each week ☐

How often do you use tuibur in a week? (Eng anga zingin nge tuibur I hmuam thin karkhatah?)
- 1 day each week or less ☐
- 2-5 days each week ☐
- 6-7 days each week ☐

How often do you take gutkha in a week? (Eng anga zingin nge I ei thin kar khatah?)
- 1 day each week or less
- 2-5 days each week ☐
- 6-7 days each week ☐

3) Do you intentionally swallow tobacco juice? (I sahdah hmuam tui I lem thin em?)
- No ☐
- Yes ☐

Do you intentionally swallow tuibur? (I tuibur hmuam tui I lem thin em?)
- No ☐
- Yes ☐

Do you intentionally swallow gutkha juice? (I gutkha ei tui I lem thin em?)
- No ☐
- Yes ☐

4) Do you pinch when you are sick or have mouth sore? No ☐
(I damloh chang emaw I ka a nat emawin sahdah I hmuam tho em? Yes ☐

Do you use tuibur when you are sick or have mouth sore? No ☐
(I damloh chang emaw I ka a nat emawin tuibur I hmuam tho em? Yes ☐

Do you take when you are sick or have mouth sore? No ☐
(I damloh chang emaw I ka a nat emawin I ei thin tho em? Yes ☐

5) How soon after waking up do youpinch? After 30 minutes ☐
(I thawh atanga engtia reiah nge sahdah I hmuam thin?) Within 30 minutes ☐

How soon after waking up do you use tuibur? After 30 minutes ☐
(I thawh atanga engtia reiah nge tuibur I hmuam thin?) Within 30 minutes ☐

How soon after waking up do you take gutkha? After 30 minutes ☐
(I thawh atanga engtia reiah nge I ei thin?) Within 30 minutes ☐

6) Is it difficult for you to not use snuff where it is restricted? No ☐
(Sahdah hmuam theihlohna hmuna I awmin hmuam loh harsa I ti em? Yes ☐

Is it difficult for you to not use tuibur where it is restricted? No ☐
(Tuibur hmuam theihlohna hmuna I awmin hmuam loh harsa I ti em? Yes ☐

Is it difficult for you to not take gutkha where it is restricted? No ☐
(Ei theihlohna hmuna I awmin ei loh harsa I ti em?) Yes ☐

---

C. ALCOHOL CONSUMPTION

| Type of alcohol | Duration / Quit already | Units per day |
|---|---|---|
| | | |
| | | |

1) How often do you drink alcohol? 2 or less days a week (Karkhatahni 2 aiatlem) ☐
(Engtiazinginnge I in thin?) 3-4 days a week (Karkhatahni 3-4) ☐
More than 5 days (Ni 5 aia tam karkhatah) ☐

2) How much water to alcohol do you take in your glass mostly? 1/4 alcohol:water ☐
(Zu no khatinahtuiengzatnge I telhthin?) 2/4 alcohol:water ☐
3/4 alcohol:water ☐
100% alcohol ☐

3) How many pegs do you normally drink? 2-3 pegs ☐
(Peg engzatnge I in thin?) 4-5 pegs ☐
6-7 pegs ☐
More than 8 pegs ☐

4) Do you drink in the morning? No ☐
(Zingahzu I in thin em?) Yes ☐

5) Do you smoke while drinking? Yes ☐
(Zu I in pahin meizial I zu tel ngai em?) No ☐

## D.FOOD PREFERENCES:

| Do you consume? ( I ei ngai em?) | 0(Never) | 1(Little) 1 days in a week | 2(Average) 2-4 days in a week | 3(Heavy) 5-7 days in a week |
|---|---|---|---|---|
| Spicy food (Thilthak) | | | | |
| Red meat (Bawngsa, Vawksa) | | | | |
| Fish (Sangha) | | | | |
| Smoked Meat (Sa rep) | | | | |
| Fresh Vegetables (Thlai hring) | | | | |
| Smoked vegetables (Thlai rep) | | | | |
| Fried food (Chawhhmeh kan) | | | | |
| Fruits (Thei) | | | | |
| Soda nena siam bai | | | | |

Do you re-use oil for cooking/ frying? Yes ☐
(Chawhmeh kana tel hmantawh hnu in hmang nawn thin em?) No ☐

How much quantity(litres) of water do you drink in a day? Less than 1 litre ☐
Nikhatah tui (litre) engzat nge I in thin? 1 – 2 litres ☐
More than 2 litres ☐

## E. ENVIRONMENTAL/ LIFESTYLE FACTORS

1) Is there a cell phone tower near your house or workplace? Yes ☐ No ☐
(I chena /thawhna hmunah cell phone tower a awm hnai em?)

2) Are you practicing jhum? Yes ☐ No ☐
(Lo neiin I hal thin em?)

3) Does your work involve exposure to sunlight? Less than 1 hour (Darkar 1 aia tlem) ☐
(Ni sa I do ngai em?) 2-3 hours daily (Darkar 2-3) ☐
4-5 hours daily (Darkar 4-5) ☐
More than 6 hours (Darkar 6 aia rei) ☐

4) Do you use sunscreen? Yes ☐ No ☐
(Sunscreen I hmang ngai em?)

5) How often do you exercise?
(Exercise I la ngai em?)

Never/Ngai lo ☐
Not regularly/ A chang changin ☐
Regularly/Ngun takin ☐

5) Do you use Cosmetics? (Cosmetics I hmang ngai em?)

Regularly ☐
Occasionally ☐
Never ☐

6) Are you exposed to secondary smoking at home or at your workplace?(I chenna/thawhna hmunah I bul a miten meizial an zu em?)

Everyday (Nitin) ☐
Occassionally (A chang changin) ☐

7) Was your mother smoking when you she was pregnant with you?
(I nuin a pailai che in mei a zu thin em?)

Everyday [Nitin] ☐
Occassionally (Zu zeuhzeuh) ☐

8) Do you work in any of the following? (Heng hmunah hian I thawk em?)

| | How many years? (Kumengzat?) |
|---|---|
| Quarry | |
| Motor workshops | |
| Steel industries | |
| Fertilizer factory | |
| Coke/Coal tar production (kawngsiampawh) | |
| Pesticides factory | |
| Plastic factory | |
| Chemical/Dye factory | |
| Tuibur/Vaihlo factory | |
| Wood/Wood cutting/Carpentry (Thing zaina) | |

F. HEALTH AND FAMILY INFORMATION

1) Besides cancer, do you have any other physical/mental conditions before?
(Taksalamahharsatna/natna I neitawh thin em (cancer tihloh)?

Yes ☐ No ☐ If yes, please specify _____

2) Do you have tonsillitis?
(Tonsillitis I neiem?)

Yes ☐ No ☐

3)How often do you brush your teeth? (Engtia zingin nge ha I nawh?) _____

Do you have broken tooth with sharp edges?
(Ha khem/ha rual lo/hriam bik I nei em?)

Yes ☐ No ☐

**REMTIHNA *(CONSENT):***

Heng a chunga thute hi ka hriatpui a, ka biological sample hi zirchian atan pek ka remti e.

*The information provided above was given with my full consent and I do not have any objection in providing my biological sample for research purposes including NGS, Genomics and Proteomics, and also in Epidemiology, Demography and Clinicopathology. I have read and understood the consent information.*

Signature:

Hming (*Name)*:

Witness:

Research chungchang hi participant hnenah hian ka hrilhfiah a, zawhna an neihte ka theihtawpa thain ka chhang e.

Signature:

Date:

Hmun *(Plac*e):

# REFERENCES


Abrahão R, Perdomo S, Pinto LFR, et al. Predictors of survival after head and neck squamous cell carcinoma in South America: the InterCHANGE study. *JCO Glob Oncol.* 2020;6:486–499.

Adelstein D, Gillison ML, Pfister DG, et al. NCCN Guidelines Insights: Head and Neck Cancers, Version 2.2017. *J Natl Compr Canc Netw*. 2017;15(6):761-770.

Afshar-Kharghan V. The role of the complement system in cancer. J Clin Invest. 2017;127(3):780-789.

Allen C, Duffy S, Teknos T, et al. Nuclear factor-kappaB-related serum factors as longitudinal biomarkers of response and survival in advanced oropharyngeal carcinoma. *Clin Cancer Res*. 2007;13(11):3182-3190.

Alsahafi E, Begg K, Amelio I, et al. Clinical update on head and neck cancer: molecular biology and ongoing challenges. *Cell Death Dis*. 2019;10(8):540.

Altieri A, Bosetti C, Gallus S, et al. Wine, beer and spirits and risk of oral and pharyngeal cancer: a case-control study from Italy and Switzerland. *Oral Oncol*. 2004;40(9):904-909.

Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. 2017;67(2):93-99.

Andersson BÅ, Lewin F, Lundgren J, et al. Plasma tumour necrosis factor-α and C-reactive protein as biomarker for survival in head and neck squamous cell carcinoma. *J Cancer Res Clin Oncol*. 2014;140(3):515-519.

Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data.

Ashraf MW. Levels of heavy metals in popular cigarette brands and exposure to these metals via smoking. *ScientificWorldJournal*. 2012;2012:729430.

Astradsson T, Sellberg F, Ehrsson YT, Sandström K, Laurell G. Serum Proteomics in Patients with Head and Neck Cancer: Peripheral Blood Immune Response to Treatment. *Int J Mol Sci*. 2022;23(11):6304.

Badola A, Mehta P, Mehra S, Sood S. Epidemiology and survival analysis of head and neck cancer: results from comprehensive care center in North India. *Oral Oncology Reports*. 2023;6:2772–9060.

Badolato R, Wang JM, Murphy WJ, et al. Serum amyloid A is a chemoattractant: induction of migration, adhesion, and tissue infiltration of monocytes and polymorphonuclear leukocytes. *J Exp Med*. 1994;180(1):203-209.

Bagnardi V, Rota M, Botteri E, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. *Br J Cancer*. 2015;112(3):580-593.

Banerjee J, Pradhan R, Gupta A, et al. CDK4 in lung, and head and neck cancers in old age: evaluation as a biomarker. *Clin Transl Oncol*. 2017;19(5):571-578.

Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, Risk Factors, and Prevention of Head and Neck Squamous Cell Carcinoma. *Med Sci (Basel).* 2023;11(2):42.

Baumann H, Gauldie J. The acute phase response. *Immunol.* 1994;15;74–80.

Bengaluru, India. Report of national cancer registry programme (ICMR-NCDIR); 2020. https://www.ncdirindia.org/All_Reports/Report_2020/default.aspx.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-2120.

Bonaventura C, Henkens R, Alayash AI, Banerjee S, Crumbliss AL. Molecular controls of the oxygenation and redox reactions of hemoglobin. *Antioxid Redox Signal*. 2013;18(17):2298-2313.

Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229-263.

Brooks PJ, Zakhari S. Acetaldehyde and the genome: beyond nuclear DNA adducts and carcinogenesis. *Environ Mol Mutagen*. 2014;55(2):77-91.

Budach W, Bölke E, Kammers K, et al. Induction chemotherapy followed by concurrent radio-chemotherapy versus concurrent radio-chemotherapy alone as treatment of locally advanced squamous cell carcinoma of the head and neck (HNSCC): a metaanalysis of randomized trials. *Radiother Oncol.* 2016;118(2):238–243.

Cabanillas R, Astudillo A, Valle M, et al. Novel germline CDKN2A mutation associated with head and neck squamous cell carcinomas and melanomas. *Head Neck*. 2013;35(3):E80-E84.

Chandrasekharappa SC, Chinn SB, Donovan FX, et al. Assessing the spectrum of germline variation in Fanconi anemia genes among patients with head and neck carcinoma before age 50. *Cancer*. 2017;123(20):3943-3954.

Chang KP, Kao HK, Wu CC, et al. Pretreatment interleukin-6 serum levels are associated with patient survival for oral cavity squamous cell carcinoma. *Otolaryngol Head Neck Surg*. 2013;148(5):786-791.

Chen CH, Su CY, Chien CY, et al. Overexpression of beta2-microglobulin is associated with poor survival in patients with oral cavity squamous cell carcinoma and contributes to oral cancer cell migration and invasion. *Br J Cancer*. 2008;99(9):1453-1461.

Chen YW, Chen IL, Lin IC, Kao SY. Prognostic value of hypercalcaemia and leucocytosis in resected oral squamous cell carcinoma. *Br J Oral Maxillofac Surg*. 2014;52(5):425-431.

Chin AR, Wang SE. Cancer Tills the Premetastatic Field: Mechanistic Basis and Clinical Implications. *Clin Cancer Res*. 2016;22(15):3725-3733.

Cho AH, Shah S, Ampil F, Bhartur S, Nathan CO. N2 disease in patients with head and neck squamous cell cancer treated with chemoradiotherapy: is there a role for posttreatment neck dissection? *Arch Otolaryngol Head Neck Surg.* 2009;135(11):1112–1118.

Cho WC. Nasopharyngeal carcinoma: molecular biomarker discovery and progress. *Mol Cancer*. 2007;6:1.

Choudhury JH, Choudhury B, Kundu S, Ghosh SK. Combined effect of tobacco and DNA repair genes polymorphisms of XRCC1 and XRCC2 influence high risk of head and neck squamous cell carcinoma in northeast Indian population. *Med Oncol*. 2014;31(8):67.

Cohen EE, Karrison TG, Kocherginsky M, et al. Phase III randomized trial of induction chemotherapy in patients with N2 or N3 locally advanced head and neck cancer. *J Clin Oncol*. 2014;32(25):2735–2743.

da Silva Santos ME, de Carvalho Abreu AK, Martins da Silva FW. KMT2 (MLL) family of methyltransferases in head and neck squamous cell carcinoma: A systematic review. *Head Neck*. 2024;46(2):417-434.

Darvishi B, Eisavand MR, Majidzadeh-A K, Farahmand L. Matrix stiffening and acquired resistance to chemotherapy: concepts and clinical significance. *Br J Cancer*. 2022;126(9):1253-1263.

Das L, Murthy V, Varma AK. Comprehensive Analysis of Low Molecular Weight Serum Proteome Enrichment for Mass Spectrometric Studies. *ACS Omega*. 2020;5(44):28877-28888.

De Schutter H, Landuyt W, Verbeken E, Goethals L, Hermans R, Nuyts S. The prognostic value of the hypoxia markers CA IX and GLUT 1 and the cytokines VEGF and IL 6 in head and neck squamous cell carcinoma treated by radiotherapy +/- chemotherapy. *BMC Cancer*. 2005;5:42.

Decker AS, Pylaeva E, Brenzel A, et al. Prognostic Role of Blood NETosis in the Progression of Head and Neck Cancer. *Cells*. 2019;8(9):946.

Denissoff A, Huusko T, Ventelä S, Niemelä S, Routila J. Exposure to alcohol and overall survival in head and neck cancer: a regional cohort study. *Head Neck.* 2022;44(10):2109–2117.

Dhaware D, Deshpande A, Khandekar RN, Chowgule R. Determination of toxic metals in Indian smokeless tobacco products. *ScientificWorldJournal*. 2009;9:1140-1147.

Duvivier L, Gerard L, Diaz A, Gillet JP. Linking ABC transporters to the hallmarks of cancer. *Trends Cancer*. 2024;10(2):124-134.

Ehlting C, Wolf SD, Bode JG. Acute-phase protein synthesis: a key feature of innate immune functions of the liver. *Biol Chem*. 2021;402(9):1129-1145.

Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228-247.

Espeli V, Zucca E, Ghielmini M, et al. Weekly and 3-weekly cisplatin concurrent with intensity- modulated radiotherapy in locally advanced head and neck squamous cell cancer. *Oral Oncol.* 2012;48(3):266–271.

Ferraguti G, Terracina S, Petrella C, et al. Alcohol and head and neck cancer: updates on the role of oxidative stress, genetic, epigenetics, oral microbiota, antioxidants, and alkylating agents. *Antioxidants.* 2022;11(1):145.

Findeisen P, Zapatka M, Peccerella T, et al. Serum amyloid A as a prognostic marker in melanoma identified by proteomic profiling. *J Clin Oncol*. 2009;27(13):2199-2208.

Gau M, Karabajakian A, Reverdy T, Neidhardt EM, Fayette J. Induction chemotherapy in head and neck cancers: results and controversies. *Oral Oncol*. 2019;95:164–169.

Getz KR, Rozek LS, Peterson LA, et al. Family history of cancer and head and neck cancer survival. *Laryngoscope.* 2017;127(8):1816–1820.

Ghias K, Rehmani SS, Razzak SA, Mutational landscape of head and neck squamous cell carcinomas in a South Asian population. *Genet Mol Biol.* 2019;4242(3):526-542.

Gokhale AS, Haddad RI, Cavacini LA, et al. Serum concentrations of interleukin-8, vascular endothelial growth factor, and epidermal growth factor receptor in patients with squamous cell cancer of the head and neck. *Oral Oncol*. 2005;41(1):70-76.

Gomaa EA, Gray JI, Rabie S, Lopez-Bote C, Booren AM. Polycyclic aromatic hydrocarbons in smoked food products and commercial liquid smoke flavourings. *Food Addit Contam*. 1993;10(5):503-521.

Gouw ZAR, Paul de Boer J, Navran A, van den Brekel MWM, Sonke JJ, Al-Mamgani A. Baseline peripheral blood leukocytosis: Biological marker predicts outcome in oropharyngeal cancer, regardless of HPV-status. *Oral Oncol*. 2018;78:200-206.

Gross M, Meirovich A, Rachmut J, et al. The Diagnostic and Prognostic Value of sIL-2R as an Immune Biomarker in Head and Neck Cancers. *Anticancer Res*. 2016;36(8):4347-4352.

Gulhar R, Ashraf MA, Jialal I. Physiology, Acute Phase Reactants. [Updated 2023 Apr 24]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK519570/

Haddad R, O'Neill A, Rabinowits G, et al. Induction chemotherapy followed by concurrent chemoradiotherapy (sequential chemoradiotherapy) versus concurrent chemoradiotherapy alone in locally advanced head and neck cancer (PARADIGM): a randomised phase 3 trial. *Lancet Oncol*. 2013;14(3):257–264.

Hagiwara S, Sasaki E, Hasegawa Y, et al. Serum CD109 levels reflect the node metastasis status in head and neck squamous cell carcinoma. *Cancer Med*. 2021;10(4):1335-1346.

Hanna GJ, Haddad RI, Lorch JH. Induction chemotherapy for locoregionally advanced head and neck cancer: past, present, future? *Oncol.* 2013;18(3):288–293.

Hao W, Zhu Y, Zhou H. Prognostic value of interleukin-6 and interleukin-8 in laryngeal squamous cell cancer. *Med Oncol*. 2013;30(1):333.

Hashibe M, Brennan P, Benhamou S, et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium [published correction appears in J Natl Cancer Inst. 2008 Feb 6;100(3):225.

Heinze G, Dunkler D. Five myths about variable selection. *Transpl Int*. 2017;30(1):6–10.

Hitt R, Iglesias L, López-Pousa A, et al. Long-term outcomes of induction chemotherapy followed by chemoradiotherapy vs chemoradiotherapy alone as treatment of unresectable head and neck cancer: follow-up of the Spanish Head and Neck Cancer Group (TTCC) 2503 Trial. *Clin Transl Oncol*. 2021;23(4):764–772.

Huang CC, Hsiao JR, Lee WT, et al. Investigating the Association between Alcohol and Risk of Head and Neck Cancer in Taiwan. *Sci Rep*. 2017;7(1):9701.

Huang J, Zhang L, Wan D, et al. Extracellular matrix and its therapeutic potential for cancer treatment. *Signal Transduct Target Ther*. 2021;6(1):153.

Huang SH, O'Sullivan B. Overview of the 8th Edition TNM Classification for Head and Neck Cancer. *Curr Treat Options Oncol*. 2017;18(7):40.

Irawan C, Benbella LG, Rachman A, Mansjoer A. Factors that Influence 2-Year Progression-Free Survival Among Head and Neck Cancer Patients. *J Epidemiol Glob Health*. 2022;12(1):16-24.

Janaydeh M, Ismail A, Zulkifli SZ, Omar H. Toxic heavy metal (Pb and Cd) content in tobacco cigarette brands in Selangor state, Peninsular Malaysia. *Environ Monit Assess*. 2019;191(10):637.

Janeway CA Jr, Travers P, Walport M, et al. Immunobiology: The Immune System in Health and Disease. 5th edition. New York: Garland Science; 2001. The complement system and innate immunity.

Jensen GL, Blanchard P, Gunn GB, et al. Prognostic impact of leukocyte counts before and during radiotherapy for oropharyngeal cancer. *Clin Transl Radiat Oncol*. 2017;7:28–35.

Jethwa AR, Khariwala SS. Tobacco-related carcinogenesis in head and neck cancer. *Cancer Metastasis Rev*. 2017;36(3):411-423.

Jiang Q, Patima S, Ye DX, Pan HY, Zhang P, Zhang ZY. Upregulation of β2-microglobulin expression in progressive human oral squamous cell carcinoma. *Oncol Rep*. 2012;27(4):1058-1064.

Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Author Correction: Head and neck squamous cell carcinoma. *Nat Rev Dis Primers*. 2023;9(1):4.

Kang N, QiuWJ, Wang B, Tang DF, Shen XY. Role of hemoglobin alpha and hemoglobin beta in non-small-cell lung cancer based on bioinformatics analysis. *Mol Carcinog*. 2022;61(6):587-602.

Kayani MA, Khan S, Baig RM, Mahjabeen I. Association of RAD 51 135 G/C, 172 G/T and XRCC3 Thr241Met gene polymorphisms with increased risk of head and neck cancer. *Asian Pac J Cancer Prev*. 2014;15(23):10457-10462.

Khariwala SS, Hatsukami D, Hecht SS. Tobacco carcinogen metabolites and DNA adducts as biomarkers in head and neck cancer: potential screening tools and prognostic indicators. *Head Neck*. 2012;34(3):441-447.

Khlifi R, Hamza-Chaffai A. Head and neck cancer due to heavy metal exposure via tobacco smoking and professional exposure: a review. *Toxicol Appl Pharmacol*. 2010;248(2):71-88.

Lalrammawia K, Buragohain A, Kakki B, et al. Determination of Multi Elements in Tobacco Plant of Northeast India by Neutron Activation Analysis and Atomic Absorption Spectrometry. *Biol Trace Elem Res*. 2022;200(10):4534-4549.

Laugesen M, Epton M, Frampton CM, Glover M, Lea RA. Hand-rolled cigarette smoking patterns compared with factory-made cigarette smoking in New Zealand men. *BMC Public Health*. 2009;9:194.

Lawal B, Tseng SH, Olugbodi JO. Pan-Cancer Analysis of Immune Complement Signature C3/C5/C3AR1/C5AR1 in Association with Tumour Immune Evasion and Therapy Resistance. *Cancers (Basel).* 2021;13(16):4124.

Lee J, Beatty GL. Serum Amyloid A Proteins and Their Impact on Metastasis and Immune Biology in Cancer. *Cancers (Basel)*. 2021;13(13):3179.

Lee JW, Stone ML, Porrett PM, et al. Hepatocytes direct the formation of a pro-metastatic niche in the liver. *Nature*. 2019;567(7747):249-252.

Lee YA, Li S, Chen Y, et al. Tobacco smoking, alcohol drinking, betel quid chewing, and the risk of head and neck cancer in an East Asian population. *Head Neck*. 2019;41(1):92-102.

Lee YG, Kang EJ, Keam B, et al. Treatment strategy and outcomes in locally advanced head and neck squamous cell carcinoma: a nationwide retrospective cohort study (KCSG HN13-01). *BMC Cancer.* 2020;20(1):813.

Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer*. 2018;18(5):269-282.

Lefebvre C, Bachelot T, Filleron T, et al. Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med*. 2016;13(12):e1002201.

Li G, Shen Q, Xu H, et al. SAA1 identified as a potential prediction biomarker for metastasis of hepatocellular carcinoma *via* multi-omics approaches. *Front Oncol*. 2023;13:1138995.

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v1 [q-bio.GN].

Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.

Li X, Wu Z, Wang Y, Mei Q, Fu X, Han W. Characterization of adult α- and β-globin elevated by hydrogen peroxide in cervical cancer cells that play a cytoprotective role against oxidative insults. *PLoS One*. 2013;8(1):e54342.

Lim SH, Sun JM, Hong J, et al. Induction chemotherapy followed by concurrent chemoradiotherapy versus CCRT for locally advanced hypopharynx and base of tongue cancer. *Korean J Intern Med*. 2021;36(1):S217–S224.

Lima LG, Monteiro RQ. Activation of blood coagulation in cancer: implications for tumour progression. *Biosci Rep*. 2013;33(5):e00064.

Litchfield K, Summersgill B, Yost S, et al. Whole-exome sequencing reveals the mutational spectrum of testicular germ cell tumours. *Nat Commun*. 2015;6:5973.

López-Verdín S, Lavalle-Carrasco J, Carreón-Burciaga RG. Molecular Markers of Anticancer Drug Resistance in Head and Neck Squamous Cell Carcinoma: A Literature Review. *Cancers*. 2018;10(10):376.

Lui VW, Hedberg ML, Li H, et al. Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. *Cancer Discov*. 2013;3(7):761-9.

Luo X, Huang X, Luo J, et al. Induction TPF followed by concurrent chemoradiotherapy versus concurrent chemoradiotherapy alone in locally advanced hypopharyngeal cancer: a preliminary analysis of a randomized phase 2 trial. *BMC Cancer*. 2022;22:1235.

Machnicki MM, Rzepakowska A, Janowska JI, et al. Analysis of Mutational Profile of Hypopharyngeal and Laryngeal Head and Neck Squamous Cell Carcinomas Identifies *KMT2C* as a Potential Tumour Suppressor. *Front Oncol*. 2022;12:768954.

Maman S, Sagi-Assif O, Yuan W, et al. The Beta Subunit of Hemoglobin (HBB2/HBB) Suppresses Neuroblastoma Growth and Metastasis. *Cancer Res*. 2017;77(1):14-26.

Manda SS, Arfa N, Sharma N, et al. Genomic landscape of head and neck cancer in Asia: A comprehensive meta-analysis of 1016 samples. *Oral Oncology Reports*. 2024;11:100628

Martinez G, Kherraf ZE, Zouari R, et al. Whole-exome sequencing identifies mutations in FSIP2 as a recurrent cause of multiple morphological abnormalities of the sperm flagella. *Hum Reprod*. 2018;33(10):1973-1984.

Marziliano A, Teckie S, Diefenbach MA. Alcohol-related head and neck cancer: Summary of the literature. *Head Neck*. 2020;42(4):732-738.

Mastronikolis NS, Delides A, Kyrodimos E, et al. Insights into metastatic roadmap of head and neck cancer squamous cell carcinoma based on clinical, histopathological and molecular profiles. *Mol Biol Rep*. 2024;51(1):597.

Meyer F, Samson E, Douville P, Duchesne T, Liu G, Bairati I. Serum prognostic markers in head and neck cancer. *Clin Cancer Res*. 2010;16(3):1008-1015.

Millrud CR, Månsson Kvarnhammar A, Uddman R, Björnsson S, Riesbeck K, Cardell LO. The activation pattern of blood leukocytes in head and neck squamous cell carcinoma is correlated to survival. *PLoS One*. 2012;7(12):e51120.

Miranda-Galvis M, Loveless R, Kowalski LP, Teng Y. Impacts of Environmental Factors on Head and Neck Cancer Pathogenesis and Progression. *Cells*. 2021;10(2):389.

Moreland, L.W. Acute phase response. *Rheumatology and Immunology Therapy. Springer, Berlin, Heidelberg* 2004.

National Cancer Institute. Head and Neck Cancers Fact Sheet. Available at: https://www.cancer.gov/types/head-and-neck/head-neck-fact-sheet Accessed on August 2024

National Cancer Institute. NCI Dictionary of Cancer Terms. 2019. Available at: **https://www.cancer.gov/publications/dictionaries/cancer-terms** Accessed August 2024.

Negri E, Boffetta P, Berthiller J, et al. Family history of cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Int J Cancer*. 2009;124(2):394-401.

Niu X, Yin L, Yang X, et al. Serum amyloid A 1 induces suppressive neutrophils through the Toll-like receptor 2-mediated signaling pathway to promote progression of breast cancer. *Cancer Sci*. 2022;113(4):1140-1153.

Nomura T, Huang WC, Zhau HE, Josson S, Mimata H, Chung LW. β2-Microglobulin-mediated signaling as a target for cancer therapy. Anticancer Agents Med Chem. 2014;14(3):343-352.

Onda M, Akaishi J, Asaka S, et al. Decreased expression of haemoglobin beta (HBB) gene in anaplastic thyroid cancer and recovery of its expression inhibits cell growth. *Br J Cancer*. 2005;92(12):2216-2224.

Palta S, Saroa R, Palta A. Overview of the coagulation system. *Indian J Anaesth*. 2014;58(5):515-23.

Patel EJ, Oliver JR, Jacobson AS, et al. Human Papillomavirus in Patients With Hypopharyngeal Squamous Cell Carcinoma. *Otolaryngol Head Neck Surg*. 2022;166(1):109-117.

Pavón MA, Arroyo-Solera I, Céspedes MV, Casanova I, León X, Mangues R. uPA/uPAR and SERPINE1 in head and neck cancer: role in Tumour resistance, metastasis, prognosis and therapy. Oncotarget. 2016;7(35):57351-57366.

Pavón MA, Arroyo-Solera I, Téllez-Gabriel M. Enhanced cell migration and apoptosis resistance may underlie the association between high SERPINE1 expression and poor outcome in head and neck carcinoma patients. *Oncotarget*. 2015;6(30):29016-33.

Picon H, Guddati AK. Mechanisms of resistance in head and neck cancer. *Am J Cancer Res.* 2020;10(9):2742-2751.

Ponzetti M, Capulli M, Angelucci A, et al. Non-conventional role of haemoglobin beta in breast malignancy. *Br J Cancer*. 2017;117(7):994-1006.

Pryzdial ELG, Leatherdale A, Conway EM. Coagulation and complement: Key innate defense participants in a seamless web. *Front Immunol*. 2022;13:918775.

Radoï L, Paget-Bailly S, Cyr D, et al. Tobacco smoking, alcohol drinking and risk of oral cavity cancer by subsite: results of a French population-based case-control study, the ICARE study. *Eur J Cancer Prev*. 2013;22(3):268-276.

Raju NJ, Kofod M, Isenbeck-Schroter M, Muller G. Heavy metal content of Indian cigarettes. Toxicol *Environ Chem*. 1999;72:215–219.

Rana A, Rana P, Gupta M, Seam R, Gupta M. Conventional chemoradiation vs. induction chemotherapy followed by conventional chemoradiation for locally advanced head and neck cancer: a prospective, randomized study. *World Acad Sci J*. 2020;2:24.

Regassa G, Chandravanshi BS. Levels of heavy metals in the raw and processed Ethiopian tobacco leaves. *Springerplus*. 2016;5:232.

Revel M, Daugan MV, Sautés-Fridman C, Fridman WH, Roumenina LT. Complement System: Promoter or Suppressor of Cancer Progression? *Antibodies (Basel)*. 2020;9(4):57

Rodrigo JP, González MV, Lazo PS. Genetic alterations in squamous cell carcinomas of the hypopharynx with correlations to clinicopathological features. *Oral Oncol.* 2002;38(4):357-63.

Roh JL, Lee H, Choi SH, Nam SY, Kim SY. Tumour-related leukocytosis predictive of recurrence and survival in patients with oral cavity squamous cell carcinoma. *Oral Dis*. 2019;25(6):1511-1518.

Schernberg A, Blanchard P, Chargari C, et al. Leukocytosis, prognosis biomarker in locally advanced head and neck cancer patients after chemoradiotherapy. *Clin Transl Radiat Oncol*. 2018;12:8-15.

Schernberg A, Huguet F, Moureau-Zabotto L, et al. External validation of leukocytosis and neutrophilia as a prognostic marker in anal carcinoma treated with definitive chemoradiation. *Radiother Oncol*. 2017;124(1):110–117.

Schernberg A, Mezquita L, Boros A, et al. Neutrophilia as prognostic biomarker in locally advanced stage III lung cancer. *PLoS One*. 2018;13(10):e0204490.

Schernberg A, Moureau-Zabotto L, Rivin Del Campo E, et al. Leukocytosis and neutrophilia predict outcome in locally advanced esophageal cancer treated with definitive chemoradiation. *Oncotarget*. 2017;8(7):11579–11588.

Sharan RN, Mehrotra R, Choudhury Y, Asotra K. Association of betel nut with carcinogenesis: revisit with a clinical perspective. *PLoS One*. 2012;7(8):e42759.

Siddiqi K, Shah S, Abbas SM, et al. Global burden of disease due to smokeless tobacco consumption in adults: analysis of data from 113 countries. *BMC Med*. 2015;13:194.

Srivastava VK, Gara RK, Rastogi N, et al. Serum vascular endothelial growth factor-A (VEGF-A) as a biomarker in squamous cell carcinoma of head and neck patients undergoing chemoradiotherapy. *Asian Pac J Cancer Prev*. 2014;15(7):3261-3265.

Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science*. 2011;333(6046):1157-1160.

Su YY, Chien CY, Luo SD, et al. Betel nut chewing history is an independent prognosticator for smoking patients with locally advanced stage IV head and neck squamous cell carcinoma receiving induction chemotherapy with docetaxel, cisplatin, and fluorouracil. *World J Surg Oncol*. 2016;14:86.

Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol*. 1996;49(8):907–916.

Sun L, Ye RD. Serum amyloid A1: Structure, function and gene polymorphism. *Gene*. 2016;583(1):48-57.

Sun YL, Patel A, Kumar P, Chen ZS. Role of ABC transporters in cancer chemotherapy. *Chin J Cancer.* 2012;31(2):51-7.

Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71(3):209-249.

Tang P, Tan C, Pang Q, et al. Combination of 35-Gene Mutation Profile and Radiotherapy Dosimetry Predicts the Therapeutic Outcome of Definitive Chemoradiation in Patients With Esophageal Squamous Cell Carcinoma. *Front Oncol.* 2021 Aug 27;11:729418.

Terra X, Gómez D, García-Lorenzo J, et al. External validation of sTWEAK as a prognostic noninvasive biomarker for head and neck squamous cell carcinoma. *Head Neck*. 2016;38 Suppl 1:E1358-E1363.

Thompson TL, Pagedar NA, Karnell LH, Funk GF. Factors associated with mortality in 2-year survivors of head and neck cancer. *Arch Otolaryngol Head Neck Surg*. 2011;137(11):1100–1105.

Toporcov TN, Znaor A, Zhang ZF, et al. Risk factors for head and neck cancer in young adults: a pooled analysis in the INHANCE consortium. *Int J Epidemiol*. 2015;44(1):169-185.

Venugopal R, Bavle RM, Konda P, Muniswamappa S, Makarla S. Familial Cancers of Head and Neck Region. *J Clin Diagn Res*. 2017;11(6):ZE01-ZE06.

von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370(9596):1453-1457.

Wang H, Ding Y, Chen Y, et al. A novel genomic classification system of gastric cancer via integrating multidimensional genomic characteristics. *Gastric Cancer*. 2021;24(6):1227-1241.

Wattenberg MM, Herrera VM, Giannone MA, Gladney WL, Carpenter EL, Beatty GL. Systemic inflammation is a determinant of outcomes of CD40 agonist-based therapy in pancreatic cancer patients. *JCI Insight*. 2021;6(5):e145389.

Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun*. 2020;9;11(1):5120.

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/

Wu P, Wu H, Tang Y, et al. Whole-exome sequencing reveals novel mutations and epigenetic regulation in hypopharyngeal carcinoma. *Oncotarget*. 2017;8(49):85326-85340.

Wyss AB, Hashibe M, Lee YA, et al. Smokeless Tobacco Use and the Risk of Head and Neck Cancer: Pooled Analysis of US Studies in the INHANCE Consortium. *Am J Epidemiol*. 2016;184(10):703-716.

Xiao BJ, Sima XX, Chen G, Gulizeba H, Zhou T, Huang Y. Predictive and prognostic role of early apolipoprotein A-I alteration in recurrent or metastatic nasopharyngeal carcinoma patients treated with anti-PD-1 therapy. *Cancer Med*. 2023;12(16):16918-16928.

Xing Y, Zhang J, Lin H, et al. Relation between the level of lymph node metastasis and survival in locally advanced head and neck squamous cell carcinoma. *Cancer*. 2016;122(4):534–545.

Yamashita T, Shimada H, Tanaka S, et al. Serum midkine as a biomarker for malignancy, prognosis, and chemosensitivity in head and neck squamous cell carcinoma. Cancer Med. 2016;5(3):415-425.

Yao J, Ding Y, Liu X, et al. Application value of whole exome sequencing in screening and identifying novel mutations of hypopharyngeal cancer. *Sci Rep*. 2023;13(1):107.

Ying H, Lin A, Liang J, Zhang J, Luo P. Association Between FSIP2 Mutation and an Improved Efficacy of Immune Checkpoint Inhibitors in Patients With Skin Cutaneous Melanoma. *Front Mol Biosci*. 2021;8:629330.

Yuan Z, Li Y, Zhang S, et al. Extracellular matrix remodeling in Tumour progression and immune escape: from mechanisms to treatments. *Mol Cancer*. 2023;11;22(1):48.

Zhang L, Jiang N, Shi Y, Li S, Wang P, Zhao Y. Induction chemotherapy with concurrent chemoradiotherapy versus concurrent chemoradiotherapy for locally advanced squamous cell carcinoma of head and neck: a meta-analysis. *Sci Rep*. 2015;5:10798.

Zhang M, Wei S, Su L, Lv W, Hong J. Prognostic significance of pretreated serum lactate dehydrogenase level in nasopharyngeal carcinoma among Chinese population: A meta-analysis. *Medicine (Baltimore)*. 2016;95(35):e4494.

Zhang R, Liu Q, Li T, Liao Q, Zhao Y. Role of the complement system in the tumour microenvironment. *Cancer Cell Int*. 2019;19:300.

Zhang Y, Gu D. Prognostic Impact of Serum CRP Level in Head and Neck Squamous Cell Carcinoma. *Front Oncol*. 2022;12:889844.

Zhang Y, Zhu X, Qiao X, et al. FSIP2 can serve as a predictive biomarker for Clear Cell Renal Cell Carcinoma prognosis. *Int J Med Sci*. 2020;17(17):2819-2825.

Zhao G, Chen L, Xiao M, Yang S. Rare coexistence of three novel CDCA7-ALK, FSIP2-ALK, ALK-ERLEC1 fusions in a lung adenocarcinoma patient who responded to Crizotinib. *Lung Cancer*. 2021;152:189-192.

Zheng Y, Miyamoto DT, Wittner BS, et al. Expression of β-globin by cancer cells promotes cell survival during blood-borne dissemination. *Nat Commun*. 2017;8:14344.

Zhou Y, Tao L, Qiu J, et al. Tumour biomarkers for diagnosis, prognosis and targeted therapy. *Signal Transduct Target Ther*. 2024;9(1):132.

Zomawia E, Zami Z, Vanlallawma A, et al. Cancer awareness, diagnosis and treatment needs in Mizoram, India: evidence from 18 years trends (2003-2020). *Lancet Reg Health Southeast Asia*. 2023;17:100281.

# LIST OF ABBREVIATIONS

| Abbreviation | Full form |
|---|---|
| HNSCC | Head and neck squamous cell carcinoma |
| WES | Whole Exome Sequencing |
| AAR | Age-adjusted rate |
| ASIR | Age Standardized Incidence Rate |
| ASMR | Age Standardized Mortality Rate |
| APC | Annual Percent Change |
| TNM | Tumour, node, metastasis |
| ENE | Extranodal extension |
| EBV | Epstein Barr Virus |
| CCRT | Concurrent Chemoradiotherapy |
| IC | Induction Chemotherapy |
| RT | Radiotherapy |
| S | Surgery |
| OS | Overall survival |
| PFS | Progression free survival |
| TSNA | Tobacco Specific nitrosamines |
| PAH | Polycyclic Aromatic Hydrocarbons |
| NNN | N - nitrosonornicotine |
| PAH | Polycyclic Aromatic Hydrocarbons |
| DNA | Deoxyribonucleic acid |
| NCCN | National Comprehensive Cancer Network |
| TSNA | Tobacco Specific nitrosamines |
| HR | Hazard Ratio |
| OR | Odds Ratio |
| LOH | Loss of Heterozygosity |
| FHC | Family History of Cancer |
| CR | Complete Response |
| SD | Stable Disease |
| PD | Progressive Disease |
| PR | Partial Response |
| TLC | Total Leukocyte Count |
| ANC | Absolute Neutrophil Count |
| µl | Microliter |
| ml | Milliliter |
| Gy | Grey |
| thou/cumm | thousands of cells per cubic millimeter |

# BRIEF BIODATA

**ZOTHANZAMI**

Email: mamifanai.185@gmail.com
Phone no: 7085881632

**Career objective**: Contributing to society by deepening the understanding and advancement of cancer care, with a focus on cancer genomics and proteomics.

**Area of interest**: Cancer genomics, cancer proteomics, diagnostic and prognostic biomarkers, head and neck cancer

**Academic Qualifications**

| Year | Education | Institute | Division |
|---|---|---|---|
| 2009 | HSLC | Mizoram Board of School Education, Mizoram | First |
| 2011 | HSSLC | Mizoram Board of School Education, Mizoram | First |
| 2015 | B.Sc. (Biotechnology) | Ch. Charan Singh University, Meerut, UP | Second |
| 2017 | M.Sc. (Biotechnology) | Mizoram University, Mizoram | First |

**Awards**

- Awarded Gold medal in M.Sc Biotechnology, Mizoram University
- Awarded INSPIRE Fellowship by Department of Science and Technology (Application Reference Number - DST/INSPIRE/03/2018/002031, INSPIRE Fellowship Registration Number – IF180827).
- Awarded best poster presentation on “Impact of Doublet Induction Chemotherapy and Clinical Factors on HNSCC from a cancer hospital in Mizoram, Northeast India” at a National Seminar on Biotechnology for

Sustainable Biosphere organized by Dept. of Biotechnology, Mizoram University.

**Conference and paper presented**

- Poster presentation on "Screening for Human papillomavirus 45 associated with Cervical lesions among Mizo women" organized by National Conference on Recent Advances in Biotechnology, Dept of Biotechnology, Mizoram University (Dissertation work)

- Oral Presentation on "First-degree Family History of Cancer might be a potential risk factor for HNSCC in an isolated mizo tribal population in North East India" at an International Conference on BEAST organized by Dept. of Biotechnology, KS Rangasamy College of Arts & Science.

- Poster presentation on "Impact of Doublet Induction Chemotherapy and Clinical Factors on HNSCC from a cancer hospital in Mizoram, Northeast India" at a National Seminar on Biotechnology for Sustainable Biosphere organized by Dept. of Biotechnology, Mizoram University.

**Workshops and trainings attended**

- Basics of Molecular Biology techniques organized by Department of Biotechnology, Mizoram University.

- Statistical Methods in Biological Research organized by Bioinformatics Infrastructure Facility (BIF), Department of Biotechnology, Mizoram University

- Antibiotic Awareness and Infection Control Program organized by Advanced level State Biotech-Hub Facility, Dept of Biotechnology, Mizoram University

- Understanding Human Disease and Improving Human Health Using Genomics Driven Approaches organized by National Institute of Biomedical Genomics, Kalyani & Dept of Biotechnology, Mizoram University

- NER workshop on 'Gene Cloning, Protein Biochemistry, Structural Biology & Bioinformatics' organized by DBT Biotechnology/Bioinformatics Training Centre, Advanced Centre for Treatment, Research & Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, India.

- Training on Proteomics – Protein purification for ESI/MS-MS Mass Spectrometry from serum in Advanced Centre for Treatment, Research & Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, India.

- Advanced Workshop on 'Understanding Human Disease and Improving Human Health Using Genomics Driven Approaches organized by National Institute of Biomedical Genomics, Kalyani.

- Hands-on Training on "Cancer Genomics" organized by DBT – NER Biotechnology/ Bioinformatics Training Centre, Advanced Centre for Treatment, Research & Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, India.

**Papers Published**

- **Zami Z**, Pachuau L, Bawihtlung Z, Khenglawt L, Hlupuii L, Lalthanpuii C, Hruaii V, Lalhruaitluanga H, Kumar NS. Treatment regimens and survival among patients with head and neck squamous cell carcinoma from Mizo tribal population in northeast India - a single centre, retrospective cohort study. **Lancet Reg Health Southeast Asia.** 2024 Mar 1;24:100377..

- Pachuau L, **Zami Z**, Nunga T, Zodingliana R, Zoramthari R, Lalnuntluanga R, Sangi Z, Rinmawii L, Kumar NS, Lalhruaitluanga H. (2022) First-degree family history of cancer can be a potential risk factor among head and neck cancer patients in an isolated Mizo tribal population, northeast India. **Clinical Epidemiology and Global Health**

- Zomawia E, **Zami Z**, Vanlallawma A, Kumar NS, Zothanzama J, Tlau L, Chhakchhuak L, Pachuau L, Pautu JL, Hmangaihzuali EVL. Cancer awareness, diagnosis and treatment needs in Mizoram, India: evidence from 18 years trends (2003-2020). **Lancet Reg Health Southeast Asia**. 2023 Sep 21;17:100281.

- Sailo CV, **Zami Z**, Lalremruata R, Sanga Z, Fela V, Kharkongor F, Chhakchhuak L, Chhakchhuak Z, Laldinmawii G, Kumar D, Kumar NS. MGIT sensitivity testing and genotyping of drug resistant Mycobacterium tuberculosis isolates from Mizoram, Northeast India. **Indian Journal of Medical Microbiology**. 2022 Jul-Sep;40(3):347-353.

- Sailo CV, **Zami Z**, Ghatak S, Nemi L, Lalremmawia K, Pachuau L, Zomawia E, Siama Z, Kumar NS. (2022) Prevalence of High-Risk HPV Types in Women with Negative Cervical Cytology in a State of Northeast India with a High Burden of Cervical Cancer. **Indian Journal of Gynecologic Oncology** 20, 8

- Vanlallawma A, **Zami Z**, Pautu JL, Bawihtlung Z, Khenglawt L, Lallawmzuali D, Chhakchhuak L, Senthil Kumar N. Pediatric leukemia could be driven

predominantly by non-synonymous variants in mitochondrial complex V in Mizo population from Northeast India. **Mitochondrial DNA A** DNA Mapp Seq Anal. 2020 Aug;31(6):245-249.

- Sailo CV, Pandey P, Mukherjee S, **Zami Z**, Lalremruata R, Nemi L, Kumar NS. Pathogenic microbes contaminating mobile phones in hospital environment in Northeast India: incidence and antibiotic resistance. **Tropical Medicine and Health.** 2019 Dec 11;47:59.

**Papers communicated**

- **Z Zami**, Pachuau L, Bawihtlung Z, Khenglawt L, Hlupuii L, Lalthanpuii C, Hruaii V, Lalhruaitluanga H, Kumar NS. "Prognostic role of Acute Phase Response in Head and Neck Squamous Cell Carcinoma among Mizo population".
- **Zami Z**, Pachuau L, Nunga T, Zodingliana R, Zoramthari R, Lalnuntluanga R, Sangi Z, Rinmawii L, Lalhruaitluanga H, Kumar NS. "Mutational landscape on Hypopharyngeal Cancer from a tribal population in Northeast India"

Personal details:

| | |
|---|---|
| Date of Birth | : May 18, 1992 |
| Gender | : Female |
| Nationality | : Indian |
| Permanent Address | : Mission Vengthlang, Aizawl, Mizoram |

Declaration: I hereby declare that all the above information is true to the best of my knowledge and belief.

Name: Zothanzami
Date:

**PARTICULARS OF THE CANDIDATE**

NAME OF CANDIDATE: ZOTHANZAMI

DEGREE: Ph. D.

DEPARTMENT: BIOTECHNOLOGY

TITLE OF THESIS: Identification of Predictive and Prognostic Biomarkers of Head and Neck Cancer in Mizo Population

DATE OF ADMISSION: 08.08.2018

APPROVAL OF RESEARCH PROPOSAL:

1. DRC: 03.05.2019
2. BOS: 07.05.2019
3. SCHOOL BOARD: 17.05.2019

MZU REGISTRATION NO.: 1506904

Ph.D. REGISTRATION NO. & DATE: MZU/Ph.D./1274 OF 08.08.2018

EXTENSION: No.16-2/MZU(Acad)/21/33-318 OF 12.06.2023

Head

Department of Biotechnology

## INTRODUCTION

Head and neck cancers include tumours arising from various regions such as the oral cavity, nasopharynx, oropharynx, hypopharynx, larynx, salivary glands, and paranasal sinuses. Head and Neck Squamous Cell Carcinoma (HNSCC) accounts for about 90% of these cancers, originating in the epithelial cells of the oral cavity, larynx, hypopharynx, nasopharynx, oropharynx, and paranasal sinuses. The remaining 10% includes cancers of the salivary glands, melanomas and sarcomas (National Cancer Institute).

Head and Neck Squamous Cell Carcinoma (HNSCC) accounts for over 0.89 million cases and 0.45 million deaths globally (Bray et al., 2024). In India, there are over 0.14 million cases and 0.13 million deaths. Mizoram, Northeast India, has a high incidence of head and neck cancers, ranking 5$^{th}$ in men (AAR 45.6/100,000) and 8$^{th}$ in women (AAR 22.7/100,000) (ICMR NCDIR, 2020). From 2003 to 2020, HNSCC was the second most common cancer in men and the sixth in women, with an increasing trend in incidence and mortality over time (Zomawia et al., 2023).

HNSCC is a highly heterogeneous cancer, presenting with varied clinical features and outcomes, heavily influenced by the molecular profile of each tumour site. Despite advancements in cancer treatments, early-stage HNSCC is typically treated with definitive Radiotherapy (RT), while locally advanced cases require concurrent chemoradiotherapy (CCRT) (Adelstein et al., 2017). However, survival rates remain low, with 70-80% survival for early-stage disease, under 70% for advanced-stage, and less than 40% for metastatic HNSCC (Barsouk et al., 2023). Resistance to treatment, tumour relapse, and metastasis further reduce survival rates in late-stage patients (Picon & Guddati, 2020; López-Verdín et al., 2018).

The Mizo population has a long history of consuming tobacco, smoked food, alcohol, and areca nuts. Tobacco is used in various forms, including cigarettes, smokeless products like "tuibur," and local hand-rolled cigarettes known as "zozial." These products contain carcinogens like Tobacco-Specific Nitrosamines (TSNAs) and Polycyclic Aromatic Hydrocarbons (PAHs), with high levels of N-nitrosonornicotine (NNN) found in tuibur (Jethwa et al., 2017; Lalrammawia et al.,

2022). Smoked foods also contain PAHs. Areca nuts, chewed with slaked lime, release reactive carcinogens, while alcohol, a known risk factor for head and neck cancers, contributes significantly to pharyngeal, laryngeal, and oral cancers through acetaldehyde, a carcinogenic byproduct (Brooks et al., 2014).

HNSCC is strongly associated with these habits, showing variation in risk factors, clinical presentation, and prevalence by subsite. Treatment follows the National Comprehensive Cancer Network (NCCN) guidelines: early-stage HNSCC patients are treated with surgery (S) or radiotherapy (RT), while advanced stages require multi-modality treatment combining chemotherapy (CT), RT, and surgery (Adelstein et al., 2017). Despite advances like CCRT and induction chemotherapy (IC), survival outcomes remain poor, especially in late stages (Zhang et al., 2015; Lee et al., 2020).

Low survival rates in HNSCC have been linked to habits like smoking and betel nut chewing, along with advanced T and N staging. Elevated leukocyte and neutrophil counts are also associated with survival outcomes, as various studies highlight the role of leukocytosis and neutrophilia in response to treatment (Millrud et al., 2012; Schernberg et al., 2018; Chen et al., 2014).

Biomarkers play a critical role in cancer prognosis and treatment monitoring. Defined as biological molecules indicating normal or abnormal processes, biomarkers help track treatment response and disease progression. Blood-based tumour biomarkers, produced by tumours or the body's response to them, have greatly enhanced cancer screening, diagnosis, and treatment strategies (Zhou et al., 2024). Over time, serum protein biomarkers have advanced therapeutic options, particularly in cancer care. Candidate tumour protein biomarkers that have been found in serum samples for head and neck cancer includes Interkeukin - 2, Interleukin – 8, Interleukin – 6, C- reactive proteins, VEGF – A, CD109, Lactate dehydrogenase etc.

Hypopharyngeal cancer is a significant subtype of head and neck squamous cell carcinoma (HNSCC), especially prevalent in Mizoram, India, which has the second-highest age-adjusted incidence rate (APC) for this cancer among all PBCR

states in the country. Despite its prevalence, it remains one of the least studied subtypes of HNSCC, with limited information on its mutational landscape. Data from the TCGA, based on only nine samples, showed TP53 as the most frequently mutated gene (in 5 of 9 samples), followed by *BRCA2* and *MUC16* (each in 3 samples). Other mutations, including *PIK3CA, FAT4, EGFR, TENT5C, LRP1B, KMT2D* and *NUMA1*, were present in two samples each. In a study of hypopharyngeal cancers from ten Chinese patients, whole-exome sequencing (WES) revealed 8,113 mutations across 5,326 genes, with *KMT2C* mutated in all cases, and other frequently mutated genes like *MEGF8, ITPR1, DYSF*, and LRP1 identified in 6 out of 10 samples. Another Chinese cohort study of 23 hypopharyngeal cancer tissues revealed somatic mutations in genes such as *TP53, REC8, PRB4, PIK3CA, CDKN2A, NSD1* and *KLK3,* along with copy number variations in *ATF1, CDKN2A*, and *CDKN2B*. A study comparing 23 hypopharyngeal and 25 laryngeal cancer samples found frequent alterations in *TP53, FAT1, NOTCH1, KMT2C* and *CDKN2A*, and noted significant differences in *CASP8* and *HRAS* mutations compared to other HNSCC subsites. Amplifications of the 11q13 region, encompassing genes like *CCND1, FGF3, FGF4* and *EMS1*, were frequently reported, alongside mutations in *ERBB1* and *MYC* oncogenes. Loss of heterozygosity (LOH) was also observed in *TP53* and *NAT2* genes. Despite these findings, the genetic profile of hypopharyngeal cancer remains largely underexplored, particularly in specific populations like Mizoram. Further research is essential to better understand the molecular mechanisms driving this cancer subtype and identify potential prognostic and therapeutic biomarkers.

This study aims to investigate the epidemiology of HNSCC in Mizoram, focusing on risk factors such as tumour sites, family history, and unique lifestyle habits (smoked food, zozial, tuibur, and local alcohol). It also evaluates treatment outcomes and survival in HNSCC patients, analysing overall survival (OS) and progression-free survival (PFS) over two years. This study also aims to identify differentially expressed serum proteins as potential prognostic or predictive biomarkers for treatment response and outcomes. Lastly, the study aims to characterize the molecular landscape of hypopharyngeal cancer in this population.

## OBJECTIVES

1. To identify the epidemiological risk factors associated with the Head and Neck cancer and to correlate with the progression of the disease in the Mizo population.

2. To identify potential predictive and prognostic protein biomarker(s) in Head and Neck cancer.

3. To identify the genomic alterations involved in Hypopharyngeal Cancer.

## METHODOLOGY

The study included patients diagnosed with HNSCC at Civil Hospital Aizawl between 2017 and 2019. The following anatomical sites were covered: oral cavity (C00.0 - C06.2), nasopharynx (C11.0–11.9), oropharynx (C09.0 - C10.9), hypopharynx (C12 - C13.9), and larynx (C32.0 - C32.2). A total of 100 HNSCC patients and 200 age-matched healthy controls were surveyed using questionnaires regarding lifestyle habits, including alcohol consumption, smoked food, tobacco use (smoking, dipping, chewing gutkha), and family history of cancer.

Smoking was quantified in pack years, calculated by dividing cigarettes smoked per day by 10 and multiplying by the number of smoking years. Participants were classified into three groups: non-smokers, smokers with below-average pack years, and smokers with above-average pack years. Alcohol consumption was measured by multiplying the number of drinking days per week by the duration of alcohol use in years, categorizing participants as non-drinkers, below-average consumers, and above-average consumers.

Family history of cancer was documented to determine if participants had blood relatives with cancer. Participants were classified into three categories: Family History of Cancer (FHC), First-Degree Family History of Cancer (First-Degree FHC, including parents and siblings), and Second-Degree Family History of Cancer (Second-Degree FHC, including uncles/aunties, cousins, and grandparents).

A retrospective cohort study was conducted to analyse survival outcomes in patients with HNSCC diagnosed between 2017 and 2020 at the Mizoram State Cancer Institute (MSCI) in Mizoram, Northeast India. Data were extracted from medical records, and patients were followed up for two years. Out of 850 patients diagnosed with head and neck cancer during this period, 210 were selected based on specific inclusion and exclusion criteria. Patients having squamous cell carcinoma primarily from oral cavity, nasopharynx, oropharynx, hypopharynx or larynx were included. The selected patients were M0 (Metastasis) stage at the time of diagnosis, belonging to Mizo population and residing within Mizoram were selected. Exclusion criteria included patients not receiving treatment or not registered at MSCI, those referred to other institutions in other states, and those lost to follow-up or who left before treatment started. Patients who refused treatment or were deemed unfit for treatment were also excluded. The study was conducted following the STROBE guidelines for reporting observational studies.

The study extracted clinical and demographic data from medical records, including age, sex, primary tumour site, TNM classification, total leukocyte count, absolute neutrophil count, and treatment regimen. Lifestyle factors such as alcohol consumption, betelnut chewing, and tobacco use (smoking and smokeless) were recorded. Smokeless tobacco was categorized into snuffing (sahdah), tuibur, and gutkha products. Tumours were classified according to the International Classification of Diseases, 10th Revision, and TNM classification (8th edition). T and N classifications were used independently due to the heterogeneous cancer sites in the cohort. Patients were grouped based on treatment: Induction chemotherapy + concurrent chemoradiotherapy/radiotherapy, Concurrent chemoradiotherapy, Radiotherapy only, Surgery + adjuvant concurrent chemoradiotherapy/radiotherapy.

CT scans were performed pre-treatment and during follow-ups every 6 months for 2 years. Response was evaluated using RECIST v1.1 criteria with four categories: Complete response (CR), Partial response (PR), Stable disease (SD), and Progressive disease (PD). Overall Survival (OS) was defined as the time from treatment initiation to death, and Progression-Free Survival (PFS) as the time to PD, SD, or death. Leukocytosis was defined as a total leukocyte count (TLC) >10,000

thou/cumm, and neutrophilia as an absolute neutrophil count (ANC) >7,000 thou/cumm.

Descriptive analysis was performed for each clinical, lifestyle, demographic factors and treatment. Logistic Regression Analysis was conducted to calculate adjusted Odds Ratios (OR) with 95% Confidence Intervals (CI) to assess the risk factors associated with HNSCC compared to controls. A p-value > 0.05 was considered non-significant. Significant ORs from univariate analysis were treated as confounders in the multivariate analysis, which was adjusted for smoking, alcohol use, areca nut consumption, and family history of cancer. Univariate and multivariate analyses for OS and PFS were conducted using SPSS. Missing data were coded as unknown. Significant variables from univariate analysis were included in the multivariate model, with multicollinearity tested using a variance inflation factor (VIF) cut-off of 2. Survival analysis was performed using the Kaplan-Meier method and log-rank test in R Studio, with a p-value of <0.05 considered statistically significant.

Serum samples were collected from 20 patients at MSCI before treatment and two weeks (14 days) after treatment. The samples were stored in cryovials with a protease inhibitor cocktail at -20°C. Acetonitrile precipitation was done to deplete the highly abundant proteins and SDS-PAGE was run to check the depletion. The samples were then digested with MS grade trypsin (Sigma Aldrich) in the ratio 1:25 (1 µg of trypsin to 25 µg of protein), which were then subjected to mass spectrometry using nanoACQUITY UPLC® chromatographic system (Waters, Manchester, UK). Data acquisition and processing was done using MassLynx 4.1 SCN781 software and Progenesis QI for Proteomics V4.2 (Non Linear Dynamics, Waters) respectively.

Peptides with ≥2 total and at least one unique peptide were selected, excluding abundant proteins. A Volcano Plot (VolcaNoseR) identified differentially expressed proteins with a fold change of 2 and a q-value ≤ 0.05. GO and KEGG pathway analyses were performed using DAVID, and a PPI network was constructed with STRING. Protein clusters were identified using MCODE (Cytoscape).

Univariate Cox analysis for OS and PFS was conducted, followed by Kaplan-Meier analysis with the log-rank test.

Biopsy tissue samples and 2 ml of peripheral blood (in EDTA vials) were collected from 10 treatment-naïve hypopharyngeal cancer patients at Civil Hospital Aizawl, Mizoram alongwith lifestyle and clinical factors. Patients were followed up and treatment response were assessed. Genomic DNA were extracted and then subjected to Whole Exome Sequencing. The quality of the data was assessed using FASTQC, adapter and low-quality reads were trimmed using Trimmomatic software, data were aligned to reference genome using BWA-MEM2. Duplicates were removed and sorted using Samtools. Variant calling was done using Mutects and ANNOVAR was used for annotation of variants.

**RESULTS**

In our study, 40% were male and the average age was 54.66 years. Oral cavity was the most prevalent site followed by hypopharynx. There are 146 smokers (48.7%), 85 in the case and 61 in the control group. Zozial was the most frequently smoked cigarette. Also, 17% consumed alcohol where 18 are controls and 33 are cases. Regression analysis showed smokers above average smoking level (i.e. 70 pack years) showed higher risk compared to smokers below average smoking level. The analysis also indicated an increased risk with increase in alcohol consumption. A significant association was also observed between HNSCC and patients with FHC.

The 210 patients for survival analysis ranges from 21 to 84 years with a median of 55 years and 79% were male. The most common cancer site was the hypopharynx (31.9%), followed by the nasopharynx (22.9%), oral cavity (18.6%), oropharynx (14.8%), and larynx (11.9%). The most frequent T classification was T2 (41.0%), and N1 was the most common N classification (42.4%). Among patients, 78.1% smoked tobacco, and 51% consumed alcohol, with 51.4% showing leukocytosis and 51.0% neutrophilia.

Treatment modalities were categorized as follows: 40.5% received induction chemotherapy followed by concurrent chemoradiotherapy or radiotherapy (sequential chemoradiotherapy), 41.0% received CCRT alone, 10.5% received RT alone, and

8.1% underwent surgery followed by adjuvant CCRT or RT. Chemotherapy agents included cisplatin, carboplatin, paclitaxel, and docetaxel. Radical RT was administered to 87.6% of patients, adjuvant RT to 8.1%, and palliative RT to 4.3%. The treatment response showed 55.7% CR, 3.81% PR, 0.48% SD and 40% PD.

The 2-year OS rate for the 210 patients was 78.1% while the progression-free survival (PFS) rate was 57.4%. The lowest OS rate (70.4%) was observed in patients receiving radiotherapy (RT) alone, while those receiving induction chemotherapy (IC) + CCRT or IC + RT had the lowest PFS rate (47.3%), though these differences were not statistically significant. Patients in the IC group treated with cisplatin plus 5-fluorouracil had the poorest OS and PFS. A significant difference in PFS was found between the IC group and those not receiving IC ($p = 0.010$). TLC $\leq$10 thou/cumm was associated with a better survival rate of 81.3% ($p = 0.015$). Similarly, lower ANC were linked to better survival rates. Significant survival differences were also found across different N classifications ($p = 0.005$), with N2 patients having the worst PFS (39.8%). Among tumour locations, oral cavity patients had the worst OS (66.7%) and PFS (48.2%), while nasopharyngeal cancer patients had the highest OS rate (88.0%), and hypopharyngeal cancer had the best PFS rate (65.7%).

Univariate Cox Regression analysis identified T and N classifications, TLC and ANC as significant predictors of OS. Due to multicollinearity (VIF >2) between TLC and ANC, ANC was adjusted for T and N classifications and excluded from other multivariate models. Multivariate analysis showed that cancer site, N classification, TLC, and treatment type were significant predictors of OS. Laryngeal cancer (HR = 5.165, $p = 0.009$) was a strong predictor of poor survival, with N2 classification also showing increased hazard (HR = 3.835, $p = 0.020$). A TLC >10 thou/cumm was another significant predictor of poor OS. For PFS, N2 classification and ANC were identified as significant predictors in univariate analysis. After adjusting for TLC, multivariate analysis revealed that laryngeal cancer (HR = 2.844, $p = 0.028$), N2 classification (HR = 3.483, $p = 0.001$), leukocytosis (HR = 2.035, $p = 0.025$) and neutrophilia (HR = 1.946, $p = 0.033$) were significant predictors of poor PFS.

A total of 40 serum samples from 20 Head and Neck Cancer patients were analysed to compare protein expression before treatment and 2 weeks into treatment. A total of 134 differentially expressed proteins were identified. A Volcano Plot filtered 78 significant proteins based on a fold change cutoff of 2 and a q-value ≤ 0.05. Among these, 155 proteins were downregulated, and 149 proteins were upregulated. Gene Ontology (GO) analysis of the 78 differentially expressed proteins revealed enrichment in blood microparticles, extracellular exosomes, and the extracellular region under the Cellular Component category. In terms of Molecular Function, the proteins were associated with organic acid binding, haptoglobin binding, oxygen transporter activity, and peroxidase activity. For Biological Processes, they were primarily linked to the acute phase response and cellular oxidant detoxification. KEGG pathway analysis showed that the proteins were involved in pathways such as complement and coagulation cascades, cholesterol metabolism, and African trypanosomiasis.

Kaplan-Meier estimates and log-rank tests revealed that patients with downregulated SAA1 protein expression exhibited significantly better overall survival ($p = 0.010$) and progression-free survival ($p = 0.005$) compared to those with upregulated or unchanged expression. Similarly, patients with downregulated B2M expression showed better prognosis ($p = 0.047$). Additionally, patients with consistent HBB expression had improved progression-free survival ($p = 0.035$) compared to those with fluctuating levels. Cox proportional hazards analysis suggested that higher ASGH expression correlated with a worse prognosis, although this was not statistically significant in the log-rank test.

Whole Exome Sequencing (WES) was conducted on tumour tissues and matched blood DNA from ten hypopharyngeal cancer samples. The analysis revealed that C/A transversions and C/T transitions were more frequent than C/G, T/A transversions, and T/G transitions (Figure 24). A total of 56,88,669 variants were identified across all samples, with 1.10% of these being exonic variants, encompassing 62,403 alterations across 12,122 genes. Among these, 29,279 non-synonymous Single Nucleotide Variants (SNVs) were detected in 8,333 genes. The mutational burden was highest in sample T7.

The top four genes, FSIP2, HMCN2, MUC3A, and ZNF705E, were altered in all ten hypopharyngeal cancer samples, with most alterations being multi-hit events. ZNF705E was predominantly altered by nonsense mutations. Patients T9, T8, and T7 had 49 out of 50 gene alterations. ATP5F1A, CCDC187, and FOXD4L4 were altered in 9 out of 10 samples, except for T3, and genes like LRP2, MUC12, MUC16, NEB, and TTN were altered in 9 out of 10 samples. From 62,403 variants in 12,122 genes, those present in at least 4 samples were used for KEGG pathway analysis. Pathways like ECM receptor interaction, HPV, ABC transporters, lung cancer, PI3K-AKT, and complement and coagulation cascades were significantly enriched (Figures 30-32). Additionally, 850 deleterious variants in 609 genes were predicted using SIFT, Polyphen, Mutation Taster, and CADD. NT5C3A was the most frequently mutated gene (7/10 samples), followed by MTMR4 and AP1G2 (5/10). Samples T7, T8, T9, and T10 had the highest mutational burden, with 128 genes altered in 2 samples and 126 altered in a single sample.

**DISCUSSION**

Smoking has been strongly associated with HNSCC (Lalrammawia et al., 2022). The risk increases in a dose-dependent manner with the number of pack years, as well as the duration and frequency of cigarette smoking, which aligns with previous studies (Hashibe et al., 2007). Our findings show that the odds ratio (OR) for those with more than 70 pack years is three times higher (OR = 15.438, 95% CI 5.989–39.793) compared to those with 70 or fewer pack years (OR = 4.896, 95% CI 2.352–10.191). Cigarettes contain over 70 carcinogens and heavy metals, which are linked to an increased risk of HNC. A study among Tunisian smokers found higher concentrations of nickel, cadmium, arsenic, and chromium in HNC tumours compared to non-smokers (Khlifi & Hamza-Chaffai, 2010). The carcinogenic mechanisms of these metals involve oxidative stress, inhibition of DNA repair, and disruption of apoptosis and methylation processes (Khlifi & Hamza-Chaffai, 2010). Moreover, many patients were frequent smokers of Zozial, a local cigarette brand found to contain high levels of heavy elements such as aluminum, manganese, silicon, arsenic, cobalt, copper, lead, iron, mercury, and cadmium, compared to other brands (Lalrammawia et al., 2022; Khariwala et al., 2012).

Alcohol consumption has been strongly linked to several cancers, including liver, gastric, and oesophageal cancers (Sung et al., 2021). The carcinogenic mechanism involves alcohol being metabolized to acetaldehyde, which can disrupt DNA stability (Brooks et al., 2014; Marziliano et al., 2020). Studies have shown a dose-dependent increase in cancer risk with alcohol intake. In a case-control study, the odds ratio (OR) for alcohol-related cancers increased from 2.1 to 21.1 for individuals consuming 3-4 to more than 12 drinks per day (Altieri et al., 2004). A meta-analysis also found increased risk ratios for oral cavity and pharyngeal cancers with higher alcohol consumption, ranging from 1.13 for light drinkers to 5.13 for heavy drinkers (Bagnardi et al., 2015). In our study, alcohol consumption also emerged as a significant risk factor, with the risk increasing in a dose-dependent manner. However, a limitation of our study is the lack of precise data on the alcohol content of local drinks, which were more frequently consumed by patients. These local drinks may contain other harmful compounds that contribute to the increased risk of head and neck cancer and other diseases. Further research into these local beverages is crucial to confirm these findings and raise public awareness.

Our study identified that a first-degree family history of cancer (FHC) is a significant risk factor for head and neck squamous cell carcinoma (HNSCC), with an adjusted odds ratio (OR) of 1.921 (95% CI 1.040–3.547). This is consistent with findings from the ICARE study, which reported an OR of 1.9 among oral cavity cancer patients with FHC, and the INHANCE consortium, which found a much stronger association (OR 7.2) in patients with both FHC and smoking/alcohol consumption. Familial factors linked to head and neck cancer include genes involved in tobacco and alcohol metabolism, DNA repair, and cell cycle regulation.

A study conducted in Indonesia by Irawan et al. observed a 2-year progression-free survival (PFS) rate of 50%, similar to our cohort's rate. Our 2-year overall survival (OS) rate of 78.1% aligns closely with that reported in a Korean study (79.8%). However, a study from northern India by Badola et al. reported a lower 2-year survival rate of 58.8%. In our analysis, patients who received induction chemotherapy (IC) had poorer PFS compared to those who did not. Despite similar numbers of patients receiving concurrent chemoradiotherapy (CCRT) alone or IC

followed by CCRT, the OS and PFS rates were worse for those receiving IC. This raises questions about the effectiveness of IC, which is generally aimed at reducing tumour size and increasing sensitivity to radiotherapy. Several randomized trials have shown no significant difference in outcomes between IC followed by CCRT and CCRT alone. Factors such as body weight, comorbidities, tumour location, size, and nodal involvement are important in treatment planning.

Our study revealed that 70.4% of patients had neck nodal involvement at diagnosis, a significant predictor of poor response to treatment. This may explain the poor outcomes observed in our cohort, as 85 patients were classified as poor responders, with 32 progressing to local recurrence or metastasis. Nodal involvement has been strongly associated with poorer survival and recurrence in several studies. Additionally, a Phase III trial by Cohen et al. found that IC did not improve OS compared to CCRT alone in patients with N2 and N3 HNSCC.

Our study suggests that leukocytosis and neutrophilia may be significant predictors of PFS and OS in patients with HNSCC. This is consistent with previous studies, where leukocytosis was found to predict both OS and PFS in patients treated with concurrent cisplatin and radiotherapy (Millrud et al., 2012; Schernberg et al., 2018). Leukocytosis has also been linked to post-surgery metastasis and tumour recurrence in oral squamous cell carcinoma (OSCC) (Schernberg et al., 2018; Chen et al., 2014; Roh et al., 2019; Gouw et al., 2018). Additionally, studies have shown that neutrophilia and leukocytosis in pre-treated patients are associated with poor responses to radiotherapy (Jensen et al., 2017). These findings align with evidence from other cancers, including oesophageal, anal, and lung cancers, where leukocytosis and neutrophilia were predictors of poor PFS and OS (Schernberg et al., 2017, 2018).

This study has several limitations, including a small sample size, which restricted adequate stratification by cancer sites or stages, limiting the statistical power of the analysis. Additionally, the retrospective design hindered the collection of direct information on patients' quality of life, diagnosis, and comprehensive reports on their well-being, including toxicity profiles, potentially introducing

confounding by indication. The lack of data on human papillomavirus (HPV) or Epstein–Barr virus (EBV) infection is another limitation, as these tests are not routinely conducted in the region. Some parameters also had missing details that could not be traced. Moreover, the exploratory nature of this study, along with the selection of variables for multivariate models based on univariate analysis, weakens the statistical power (Sun et al., 1996; Heinze et al., 2017).

Despite these limitations, this study offers valuable insights into HNSCC survival in a region with a high cancer prevalence. The 2-year OS and PFS rates were 78.1% and 57.4%, respectively. The findings highlight that the multi-modality approach, particularly CCRT, provides a survival advantage over other treatment methods, including sequential therapy. Poor prognosis was linked to elevated TLC, high ANC, nodal involvement, and laryngeal cancer site. These findings underscore the need for further investigation in this context to confirm and expand upon the results.

Functional enrichment analysis revealed that the Acute Phase Response (APR) pathway was significantly enriched in our study. The acute phase response is a systemic reaction triggered by injury, infection, or neoplastic growth, involving the release of pro-inflammatory cytokines like TNF-α, IL-1, and IL-6 from local inflammatory cells such as macrophages and neutrophils (Baumann & Gauldie, 1994; Moreland, 2004). These cytokines stimulate hepatocytes to synthesize and release various Acute Phase Reactants (APRs) into the bloodstream (Ehlting et al., 2021). APRs are classified into positive and negative, based on their serum concentration during inflammation (Gulhar et al., 2024).

Positive APRs, including C-reactive protein, procalcitonin, fibrinogen, ferritin, hepcidin, and serum amyloid A, are upregulated during inflammation, while negative APRs like transferrin, albumin, and prealbumin are downregulated. Among the positive APRs, serum amyloid A (SAA) proteins, particularly SAA1 and SAA2, are prominent and mainly released during the acute phase (Lee et al., 2021). SAA1 promotes cancer metastasis by creating a pro-metastatic niche, facilitating myeloid cell infiltration into tissues, which supports the colonization of disseminated tumour

cells (DTCs) (Badolato et al., 1994; Lee et al., 2019). SAA1 also activates immunosuppressive neutrophils through the TLR2-mediated signalling pathway, leading to cancer progression (Niu et al., 2021).

In our study, patients with downregulated SAA1 protein levels during treatment had better survival and progression-free survival rates. SAA1 has been identified as a prognostic biomarker in various cancers. Elevated SAA1 levels have been linked to poor outcomes in nasopharyngeal cancer, liver metastasis in non-small cell lung cancer, advanced pancreatic cancer, and melanoma (Cho et al., 2004; Lee et al., 2019; Wattenberg et al., 2021; Findeisen et al., 2009). In Hepatocellular Carcinoma (HCC), higher SAA1 expression was associated with worse overall survival (Li et al., 2023). These findings highlighted the potential of SAA1 as a marker for metastasis and poor prognosis in multiple cancers.

In our study, we observed that patients with downregulated β2-microglobulin (β2M) expression had better progression-free survival (PFS) compared to those with upregulated or unchanged β2M expression in treated samples versus treatment-naïve samples. β2M, a 99-amino acid protein with a molecular weight of 12-kDa, is a crucial component of the Major Histocompatibility Complex (MHC) Class I antigen/Human Leukocyte Antigen (HLA) present on nucleated cells' surfaces. It plays a vital role in immune recognition by CD8+ T lymphocytes, as well as in immunoglobulin transport and iron metabolism (Nomura et al., 2014).

β2M has been implicated in promoting growth, angiogenesis, epithelial-mesenchymal transition (EMT), and bone metastasis, serving as a prognostic indicator in solid tumours. Consistent with our findings, a Taiwanese study reported significantly better survival in oral squamous cell carcinoma (OSCC) patients with decreased β2M expression (Chen et al., 2008). Another study comparing metastatic and primary OSCC patients found higher β2M expression in metastatic cases, suggesting its role in tumour invasion and metastasis (Jiang et al., 2012).

Elevated β2M levels have been established as a poor prognostic marker in various cancers. For example, in multiple myeloma, high serum β2M levels independently predict worse overall survival and progression-free survival,

especially in patients undergoing chemotherapy and stem cell transplantation (Nomura et al., 2014). This further supports the notion that β2M could be a significant biomarker in predicting poor treatment outcomes in HNSCC.

Our study observed that changes in serum Haemoglobin Subunit Beta (HBB) expression during treatment—whether upregulation or downregulation—were associated with worse progression-free survival (PFS). HBB is a well-known globin protein primarily responsible for oxygen transport in red blood cells. Interestingly, HBB expression has been reported in other cells, including cancer cells. Increased HBB expression has been linked to tumour progression in breast cancer, where it promotes proliferation, migration, and angiogenesis (Ponzetti et al., 2017). Similarly, elevated HBB levels have been observed in prostate, lung, and cervical cancers, contributing to metastasis and poor outcomes (Zheng et al., 2017; Li et al., 2013). However, in some studies, HBB has been shown to inhibit cancer cell proliferation in non-small cell lung carcinoma, indicating a complex role in tumour biology (Kang et al., 2022; Maman et al., 2017). Our findings suggest that altered HBB expression during treatment may serve as a potential biomarker for worse PFS in patients with HNSCC, though the exact mechanisms require further exploration.

This study identified 78 proteins with significant differential expression in HNSCC samples during treatment, using pre-treatment samples as a reference. Functional analysis revealed that the blood microparticle and acute phase response pathways were the most significantly enriched. Importantly, alterations in SAA1, β2M, and HBB proteins were strongly linked to patient outcomes, suggesting their potential as biomarkers for HNSCC. Larger cohort studies are needed to validate these results.

This pilot study explores the genomic landscape of hypopharyngeal cancer in this population through whole exome sequencing (WES) of tumour tissues and matched blood samples from ten patients. To date, there have been few WES studies focused solely on hypopharyngeal cancer. Our findings indicate that the most frequently mutated genes in our cohort differ from those reported in other studies. The four most commonly mutated genes in our samples were *FSIP2, HMCN2,*

*MUC3A, and ZNF705E*. In contrast, *TP53, BRCA2,* and *MUC16* were the top mutated genes identified in the TCGA cohort of nine hypopharyngeal cancer patients. Additionally, Yao et al. reported *TTN, TP53*, and *ANK3* as the most frequently mutated genes, while Machnicki et al. found *TP53, FAT1*, and *NOTCH1* to be the most mutated in another cohort (Yao et al., 2023; Machnicki et al., 2022).

The *FSIP2* gene encodes fibrous sheath interacting protein 2, which is essential for spermatogenesis (Martinez et al., 2018). Although not directly linked to head and neck cancer, mutations in *FSIP2* have been associated with esophageal cancer, gastric cancer, cutaneous melanoma, and breast cancer (Tang et al., 2021; Wang et al., 2021; Ying et al., 2021; Lefebvre et al., 2016). Changes in its expression have also been linked to renal cell carcinoma (Zhang et al., 2020). Moreover, amplifications of the *FSIP2* gene and *FSIP2-ALK* fusions have been identified in germ cell tumours and lung adenocarcinoma, respectively (Litchfield et al., 2015; Zhao et al., 2021).

Among the top 50 mutated genes in our cohort, *TP53* and *NOTCH1* were mutated in 80% of samples, consistent with their status as top mutated genes in the TCGA cohort. These genes are known for their high mutation rates across various cancers, including head and neck cancer (Manda et al., 2024; Stransky et al., 2011; Leemans et al., 2018).

The ECM receptor interaction and motor protein pathways were the most significantly enriched in our study (p-value $< 0.000000$), highlighting their importance in cancer biology. ECM stiffness, through PI3K-AKT signalling, enhances cancer cell migration, angiogenesis, and can contribute to radio- and chemoresistance (Huang et al., 2021). PI3K-Akt mutations were common in our cohort and linked to higher mutation rates in other cancer genes (Lui et al., 2013).

The coagulation and complement cascade pathways were also significantly enriched, playing critical roles in immune modulation, tumour growth, and metastasis (Palta et al., 2014; Janeway et al., 2001). Complement activation recruits immunosuppressive cells and aids tumour survival (Zhang et al., 2019). Additionally, changes in SERPINE1 expression and mutations in C3 and C5 were linked to poor

prognosis and immune modulation in HNSCC (Pavón et al., 2015; Lawal et al., 2021).

We observed a higher clustering of tumour mutational burden in patients from T7, T8, T9, and T10, where alcohol, smoking, and tobacco consumption were more prevalent. Among the 31 deleterious variants found in 4 or more samples, the variant rs79747830 (NC_000007.13.33054388T>C) in *NT5C3A* was present in 7 of the 10 samples, previously reported in HNSCC patients from Pakistan (Ghias et al., 2019). A single nucleotide polymorphism in the *KMT2C* gene, associated with histone methylation, was identified in four samples. *KMT2C* mutations have been linked to tumour progression and poor outcomes in head and neck cancer, with some studies suggesting its role as a tumour suppressor (Machnicki et al., 2022).

Our study, the first genomic analysis of hypopharyngeal cancer in the Mizo population, found the top ten most frequently altered genes to be *FSIP2, HMCN2, MUC3A, ANF705E, ATP5F1A, CCDC187, FOXD4L4, LRP2, MUC12* and *MUC16*. Out of 850 variants predicted to be pathogenic, *NT5C3A* was present in 70% of the samples. Pathway analysis revealed significant enrichment in the ECM receptor interaction and complement/coagulation cascade pathways, which were also enriched in the proteomics dataset. Many genes identified in this cohort have not been reported in other hypopharyngeal cancer studies. Further screening and functional validation are needed to deepen our understanding.

SUMMARY

- This study aims to identify the risk factors and their association with survival outcome of HNSCC in Mizo population, identify potential prognostic protein biomarkers that are differentially expressed in serum of patients during treatment compared with treatment naive samples and explore the genomic landscape of Hypopharyngeal Cancer using Whole Exome Sequencing (WES).
- Multivariate Regression analysis showed male patients had significantly higher Odds Ratio (OR = 6.694, p-value = <0.05) compared to female patients. Patients above 45 years of age were likely to develop Head and Neck Squamous Cell Carcinoma (HNSCC) compared to patients below 45 years of age (OR = 3.979, p-value = < 0.05).
- Smoking and consumption of alcohol were also significantly associated with HNSCC compared to non-smokers and non-drinkers. The association increases as level of smoking and consumption of alcohol increases.
- Patients having first degree family history of cancer were found to be significantly associated with HNSCC as well (OR = 1.921, p-value = 0.037)
- Smoking has been highly associated with HNSCC, with increased risk in a dose-dependent manner. Majority of the patients in our study smoked local made Zozial which was found to contain high levels of heavy metals (Lalrammawia et al., 2022; Hashibe et al., 2007)
- Significant associations between Head and Neck Cancer and First-Degree Family History of Cancer in combination with smoking and alcohol consumption have also been reported (Radoï et al., 2013; Negri et al., 2009; Toporcov et al., 2015)
- The overall 2-year overall survival (OS) was 78.1% and Progression Free Survival (PFS) was 57.4% which are comparable to other studies (Irawan et al., 2022; Zhang et al., 2015; Badola et al., 2023).
- Patients who received Induction Chemotherapy (IC) had lower PFS rate (47.3%) than patients who did not received IC with a significant log-rank test.

- Patients who had Leukocytosis [Total Leukocyte Count (TLC) > 10 thou/cumm] had lower OS rate (58.4%) compared to patients with TLC ≤ 10 thou/cumm.
- Patients who had Neutrophilia [Absolute Neutrophil Count (ANC) > 7 thou/cumm had significantly lower OS and PFS rates (57% and 36.4%, respectively) with log-rank tests p-value 0.014 and 0.043, respectively.
- Multivariate Cox Proportional Hazard showed that among all the sites, Laryngeal cancer was a strong predictor of OS (HR = 5.165, p-value = 0.009) and PFS (HR = 2.844m p-value = 0.028). Nodal involvement (N2), Leukocytosis and Neutrophilia were also significant predictors of OS and PFS.
- Induction Chemotherapy did not improve OS compared to CCRT in patients with N2 and N3 involved HNSCC (Cohen et al., 2014).
- Leukocytosis is a predictor of OS and PFS in patients treated with concurrent Cisplatin and Radiotherapy and linked with tumour recurrence (Schernberg et al., 2018).
- Patients with Leukocytosis and Neutrophilia has been shown to have poor response to radiotherapy (Jensen et al., 2017). They are also found to be predictors for poor OS and PFS in other cancers (Schernberg et al., 2017;2018)
- Out of 134 differentially expressed proteins found in our samples, 78 were significantly expressed 304 times during treatment when compared with treatment naïve samples, of which 155 were down regulated and 149 proteins were upregulated.
- Patients having downregulated SAA1 protein showed better OS and PFS compared to patients with upregulated or baseline SAA1 expressions.
- Patients with downregulated ß2M expressions had better PFS compared to patients with upregulated or baseline ß2M expressions.
- Patients with consistent HBB expressions had better PFS compared to patients with change in HBB expressions.

- The functional enrichment analysis of the 78 differentially expressed proteins showed that Acute Phase Response Pathway was significantly enriched. SAA1 protein play a huge role in Acute Phase Response facilitating pro-metastatic niche which helped in colonizing of disseminated tumour cells thus leading to metastasis (Lee et al., 2019; Chin & Wang, 2016).
- SAA1 protein also induces accumulation of immunosuppressed neutrophils through Toll-like receptor 2 (TLR) mediated pathway and promotes neutrophil apoptosis resistance through TLR2/MYD88 mediated pathway, both of which leads to progression of cancer (Niu et al., 2021).
- β2M and HBB protein expressions have been linked with metastatic cancer and have been reported to contribute to tumour invasion, migration, angiogenesis, EMT and bone metastasis but the molecular mechanisms involved have not yet been fully explored (Nomura et al., 2014; Jiang et al., 2012; Ponzetti et al., 2017; Kang et al., 2022).
- WES in ten hypopharyngeal cancers identified 62,403 exonic variants in 12,122 genes where the top 10 most frequently mutated genes are *FSIP2*, *HMCN2, MUC3A, ANF705E, ATP5F1A, CCDC187, FOXD4L4, LRP2, MUC12* and *MUC16*.
- The significantly enriched pathways which have been associated with cancer are the ECM receptor interaction, Human Papillomavirus infection, ABC transporters, Complement and Coagulation Cascade, Small Cell Lung Cancer and P13K-Akt Signalling Pathways.
- Out of 29,279 exonic non-synonymous across 8333 genes, about 850 variants in 609 genes were predicted to be deleterious by SIFT, Polyphen, Mutation Taster and CADD. Out of these 850 variants, 107 variants in 31 genes were found to be altered in ≥3/10 samples.
- Mutations in the NT5C3A gene were identified in 7 out of 10 hypopharyngeal cancer cases, followed by mutations in *MTMR4* and *AP1G2*, found in 5 out of 10 samples. Variants of *AZIN2, IRX6, KMT2C, NUDT12, POP5*, and *SHANK2* were present in 4 out of 10 samples. Notably, samples T7, T8, T9, and T10 exhibited a higher number of mutations compared to the others.

- There is very limited information on WES of hypopharyngeal cancer till today. The top most frequently mutated genes are inconsistent with top genes identified in few available studies (TCGA, Yao et al., 2023; Machnicki et al, 2022).
- KMT2C gene mutations have been reported in Hypopharyngeal cancer, contributing to tumour progression when altered with other oncogenes. Its role in tumour suppression have also been demonstrated were loss of function led to increased proliferation in hypopharyngeal cell lines (Yao et al., 2023; Machnicki et al, 2022).
- This study identified that patients having first degree of family history of cancer, smoking and alcohol consumption were significantly associated with HNSCC. Survival analysis revealed that nodal involvement, leukocytosis and neutrophilia were predictors of OS and PFS. Also, patients who received IC had worse outcome compared to patients who did not receive.
- Changes in expressions of SAA1, ß2M and HBB proteins during treatment may serve as prognostic biomarkers of HNSCC.
- WES revealed that the top mutated genes differed from those identified in other hypopharyngeal cancer cohorts from different populations. Mutations in the KMT2C gene were found in our samples and have been reported in other hypopharyngeal cancer cohorts, highlighting the need for further investigation into its role in this cancer. We also identified mutations in genes not previously reported in hypopharyngeal cancer.

## REFERENCES


Adelstein D, Gillison ML, Pfister DG, et al. NCCN Guidelines Insights: Head and Neck Cancers, Version 2.2017. *J Natl Compr Canc Netw*. 2017;15(6):761-770.

Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, Risk Factors, and Prevention of Head and Neck Squamous Cell Carcinoma. *Med Sci (Basel).* 2023;11(2):42.

López-Verdín S, Lavalle-Carrasco J, Carreón-Burciaga RG. Molecular Markers of Anticancer Drug Resistance in Head and Neck Squamous Cell Carcinoma: A Literature Review. *Cancers*. 2018;10(10):376.

Roh JL, Lee H, Choi SH, Nam SY, Kim SY. Tumor-related leukocytosis predictive of recurrence and survival in patients with oral cavity squamous cell carcinoma. *Oral Dis*. 2019;25(6):1511-1518.

Radoï L, Paget-Bailly S, Cyr D, et al. Tobacco smoking, alcohol drinking and risk of oral cavity cancer by subsite: results of a French population-based case-control study, the ICARE study. *Eur J Cancer Prev*. 2013;22(3):268-276.

Sharan RN, Mehrotra R, Choudhury Y, Asotra K. Association of betel nut with carcinogenesis: revisit with a clinical perspective. *PLoS One*. 2012;7(8):e42759.

Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol*. 1996;49(8):907–916.

Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science*. 2011;333(6046):1157-1160.

Schernberg A, Moureau-Zabotto L, Rivin Del Campo E, et al. Leukocytosis and neutrophilia predict outcome in locally advanced esophageal cancer treated with definitive chemoradiation. *Oncotarget*. 2017;8(7):11579–11588.

Toporcov TN, Znaor A, Zhang ZF, et al. Risk factors for head and neck cancer in young adults: a pooled analysis in the INHANCE consortium. *Int J Epidemiol*. 2015;44(1):169-185.

Tang P, Tan C, Pang Q, et al. Combination of 35-Gene Mutation Profile and Radiotherapy Dosimetry Predicts the Therapeutic Outcome of Definitive Chemoradiation in Patients With Esophageal Squamous Cell Carcinoma. *Front Oncol.* 2021 Aug 27;11:729418.

Picon H, Guddati AK. Mechanisms of resistance in head and neck cancer. *Am J Cancer Res.* 2020;10(9):2742-2751.

National Cancer Institute. Head and Neck Cancers Fact Sheet. Available at: https://www.cancer.gov/types/head-and-neck/head-neck-fact-sheet Accessed on August 2024

Venugopal R, Bavle RM, Konda P, Muniswamappa S, Makarla S. Familial Cancers of Head and Neck Region. *J Clin Diagn Res*. 2017;11(6):ZE01-ZE06.

Wattenberg MM, Herrera VM, Giannone MA, Gladney WL, Carpenter EL, Beatty GL. Systemic inflammation is a determinant of outcomes of CD40 agonist-based therapy in pancreatic cancer patients. *JCI Insight*. 2021;6(5):e145389.

Zomawia E, Zami Z, Vanlallawma A, et al. Cancer awareness, diagnosis and treatment needs in Mizoram, India: evidence from 18 years trends (2003-2020). *Lancet Reg Health Southeast Asia*. 2023;17:100281.