



Original Research Article

Genomic landscape of head and neck cancer in Asia: A comprehensive meta-analysis of 1016 samples



Srikanth S. Manda ^{a,b}, Nafisa Arfa ^c, Neha Sharma ^{a,b}, Aparna R. Parikh ^d, Thomas J. Roberts ^d, Sewanti Limaye ^e, Venkataraman Ramachandran ^{a,b}, Kumar Prabhash ^{g,h}, Moni A. Kuriakose ^{a,b}, Prashant Kumar ^{a,b,f,*}

^a Karkinos Foundation, Mumbai, Maharashtra, 400086, India

^b Karkinos Healthcare Pvt Ltd, Navi Mumbai, Maharashtra, 400705, India

^c Department of Chemical Engineering, Indian Institute of Technology, Guwahati, Assam, 781039, India

^d Cancer Center, Massachusetts General Hospital, Boston, MA, 02114, USA

^e Department of Medical and Precision Oncology, Sir HN Reliance Foundation Hospital and Research Centre, Mumbai, 400004, Maharashtra, India

^f Centre of Excellence for Cancer – Gangwal School of Medical Sciences and Technology, Indian Institute of Technology Kanpur, Kanpur, Uttar Pradesh, 208016, India

^g Homi Bhabha National Institute, Training School Complex, Anushakti Nagar, Mumbai, 400094, Maharashtra, India

^h Department of Medical Oncology, Tata Memorial Hospital, Mumbai, 400012, Maharashtra, India

ARTICLE INFO

Keywords:

Head and neck cancer

HNSCC

Asia

WES

NGS

Mutations

Cancer drivers

ABSTRACT

Head and neck cancer (HNC) is a diverse group of malignancies arising in the mucosal linings of the oral cavity, pharynx, and larynx, influenced by factors such as tobacco use, alcohol consumption, and human papillomavirus (HPV) infection. This study conducts a comprehensive meta-analysis of the mutational landscape of HNC across Asian cohorts, encompassing India, Korea, Japan, China, Singapore, and Saudi Arabia. The analysis highlights distinct genetic profiles influenced by environmental exposures, lifestyle habits, and genetic predispositions. Notably, the RAF family proteins, enriched in both Indian and Chinese cohorts, present potential therapeutic targets for RAF inhibitors like Vemurafenib. Additionally, specific mutations like MET in Singaporean patients can be effectively addressed with drugs like Crizotinib, leading to rapid responses in HNSCC. Smokers exhibited high frequencies of CASP8 and FAT1 mutations. Novel driver genes, including RYR2 and ANK2, emerged with significant mutational frequencies in smokers. The RAS signaling pathway was identified as a prominent driver in HNC, contrasting with the globally prevalent PIK3CA/MTOR pathway. This study also underscores the high prevalence of HRAS mutations in Indian and Saudi cohorts. The study emphasizes the necessity for region-specific data to understand the unique molecular differences and develop effective therapies. The identification of NBEA and ANK2 as potential novel driver genes in HNC highlights new avenues for research and targeted therapeutic interventions tailored to the genetic profiles of Asian HNC patients.

1. Introduction

Head and Neck Squamous Cell Carcinoma (HNSCC), accounting for more than 90 % of head and neck (HNC) malignancies, is one of the most prevalent cancers in Asian countries. It is the 6th most common human cancer comprising of a genetically heterogeneous collection of malignancies [1]. According to GLOBOCAN, approximately 70, 000 new cases of HNSCC are added annually, with more than 350, 000 deaths per year with a projection that by 2030, there would be 1.08 million new cases of HNSCC yearly, or a 30 % increase in incidence [2]. Oral cancer

incidence is particularly seen to be higher in Southeast Asia and Asia-Pacific regions often associated with chewing of areca nut, with or without tobacco. Tobacco accounts for more than 80 % of all HNSCC cases in India [1,3]. It is observed that carcinomas of the larynx and oral cavity are usually associated with smoking, alcoholism (or both), however pharyngeal cancer is often associated with human papillomavirus (HPV) infection, particularly the HPV-16 subtype [4,5].

The predominant genetic changes in HNSCC include loss-of-function mutations in genes such as p53, p16 and activation of oncogenes like epidermal growth factor receptor (EGFR) [6]. However, the etiology of

* Corresponding author. Karkinos Foundation, India.

E-mail address: prashant.kumar@karkinosfoundation.org (P. Kumar).

HNSCC varies significantly across different geographical regions and genetic backgrounds, additionally lack of timely access to healthcare facilities further complicates the identification of universal prognostic markers of HNSCC in developing nations [1,7]. These variations underscore the necessity for more focused studies in diverse populations to better understand the disease's multifaceted nature. Regional differences in HNSCC etiology are influenced by environmental exposures, lifestyle choices, and genetic predispositions. Comprehensive efforts considering the unique genetic and environmental contexts of different regions are essential to uncover the mechanisms driving HNSCC and improve patient outcomes.

To address this pertinent issue, we have attempted to understand the mutational landscape of HNSCC across Asia by collating datasets from studies on various ethnic populations. In the current study, we have performed a meta-analysis of published genomics studies, comprising both targeted and whole exome datasets from 1016 patient samples across six Asian countries. This study serves as a compendium of mutations across the Asian cohorts and can be used to guide future research efforts, inform clinical strategies, and facilitate the development of targeted therapeutic interventions tailored to the unique genetic profiles of HNC patients in these regions.

2. Methods

2.1. Study selection criteria

The studies with genomic mutation datasets were collated through an extensive literature review using the databases, i.e., PubMed and Google Scholar. The keywords employed for the database search

included "Oral cancer", "Head and neck", "Tongue cancer", "Oropharyngeal cancer", "Nasopharyngeal Cancer", "Hypopharyngeal Cancer", "HNSCC", "OSCC", "Asia", "Sequencing", "NGS", "Supplementary", "WES". The studies included articles published up to June 2022. The articles were screened further to exclude the studies that met at least one of the following criteria: missing patient information, i.e., no sample ID, studies on cell line data, unavailable mutation datasets, insufficient variant information, studies focused on single gene mutation, specific viral or bacterial infection based HNSCC studies, lack of in-depth study or studies with different ethnicities patients. Finally, the inclusion criteria for the studies used in analysis involve (1) information regarding genes across the samples; and (2) the corresponding mutation frequencies at gene and variant levels.

Other information pertaining to cancer genes census, driver and pathway analysis are provided in supplementary data.

3. Results

3.1. Study search and patient characteristics

In the current investigation, a total of 38 studies were collated through PubMed and Google scholar database search (Supplementary Table 1). Among them, 20 studies were excluded from the analysis based on several reasons as mentioned in Fig. 1A. After strict filtering of data, 18 studies including 10 whole exome sequencing (WES) and 8 targeted genome sequencing (TGS) datasets across six countries were considered for further analysis. Based on the collected data, highest number of samples were available for Chinese population (310 TGS; 231 WES) followed by India (149 TGS; 100 WES), Singapore, Japan, Korea

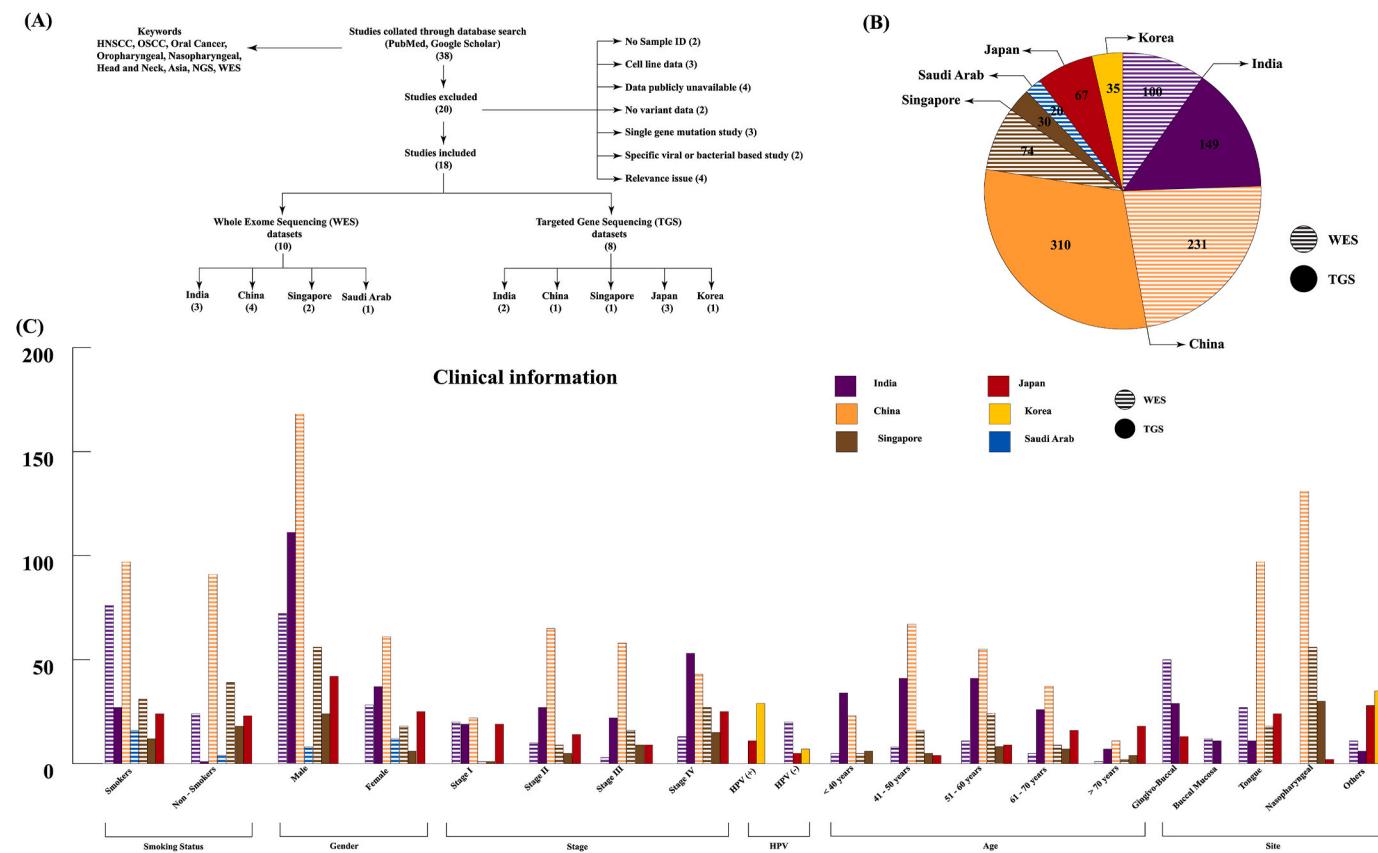


Fig. 1. (A) Workflow employed in the current study. Total number of samples at each stage are depicted in numbers along with the filtering criteria. (B) Pie Chart depicts the sample distribution across all the countries involved in the study. Whole exome sequencing (WES) data is shown as solid colour with black strip while the targeted gene sequencing (TGS) data is shown in solid colours. (C) Bar graphs showing the clinicopathological characteristics of the cohort. The vertical axis represents the number of samples while the horizontal axis shows the various clinical information, i.e., smoking status, gender, stage, HPV, age, and site. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

and Saudi Arabia (SA). Sample distribution is depicted in Fig. 1B. Clinico-pathological features of all the samples were grouped into various categories (Fig. 1C). HNSCC had a greater effect on smokers than non-smokers, with exception of China, where the difference was negligible. Male to female ratio of 2.5 was seen in overall Asian data (457 males, 181 females). In general, male predominance is seen in HNSCC in Asian countries. Similarly, higher number of smokers were present in the Asian cohort (271 smokers vs 182 non-smokers). This gender disparity may be attributed to higher consumption of tobacco and alcohol in men which are significant risk factors for HNSCC. In Asian countries, HNSCC is detected mostly at late and advanced stages as shown in Fig. 1C. A total of 81 cases (~7 %) were samples from stage I as against majority of the samples from later stages. Tongue and nasopharyngeal sites were the most affected tumor sites followed by gingiva-buccal and buccal mucosa.

3.2. Mutational profile of HNC in Asian cohorts

Whole-exome data collation from all the HNC studies from the four different countries showed at least one mutation in 14,665 genes (Supplementary Table 2). Among these genes, 1,746 were found to be recurrently mutated across the cohort. Further, a total of 598 cancer gene census (CGC) genes were altered across HNC samples, of which 153 showed recurrence. Fig. 2 depicts the top 50 CGC genes along with their mutation frequencies. Top highly mutated genes in HNC include TP53 (35.8 %), MUC16 (12.9 %), NOTCH1 (11.1 %), CDKN2A (8.7 %), and CASP8 (8.5 %). We further compared the frequency of the mutations in the top genes in HNC from the Asian population to the TCGA HNC cohorts (Table 1). We observed a striking reduction in the gene mutation

Table 1

Table showing top genes and their mutational frequencies in Asian and TCGA cohorts from WES datasets. The difference in mutational frequencies is shown in percentage.

Gene	Freq (Current) %	TCGA %	Difference
TP53	35.01	71.2	36.19
MUC16	12.59	20.1	7.51
NOTCH1	10.98	17.7	6.72
FAT1	9.382	22.2	12.82
CDKN2A	8.696	21.5	12.80
CSMD3	7.551	20.5	12.95
LRP1B	5.95	17.8	11.85
KMT2D	5.263	16.1	10.84
PIK3CA	5.263	18.5	13.24
NSD1	2.746	11.6	8.85
FAM135B	1.88	10	8.12

frequency of TP53 in the Asian population to 35.8 % as compared to 71.2 % in TCGA datasets. Apart from this, there was a significantly lower mutation frequency for other hallmark genes of HNSCC in the Asian population, including NOTCH1 (11.1 %), PIK3CA (5.4 %), FAT1 (9.4 %), and CDKN2A (8.7 %). The mutational frequencies were consistently lower for other frequently mutated genes in TCGA study including CSMD3 (7.8 %), MUC16 (12.9 %), KMT2D (5.4 %), and NSD1 (2.8 %) (Fig. 2).

While inspecting the mutation data for genes that are significantly more mutated in the four countries as compared to TCGA, we found genes unique to each country exhibiting considerable variation. CASP8 is mutated at frequencies 24 % in India and 30 % in SA which is significantly greater than the TCGA frequency of 10 %. CASP8 mutation

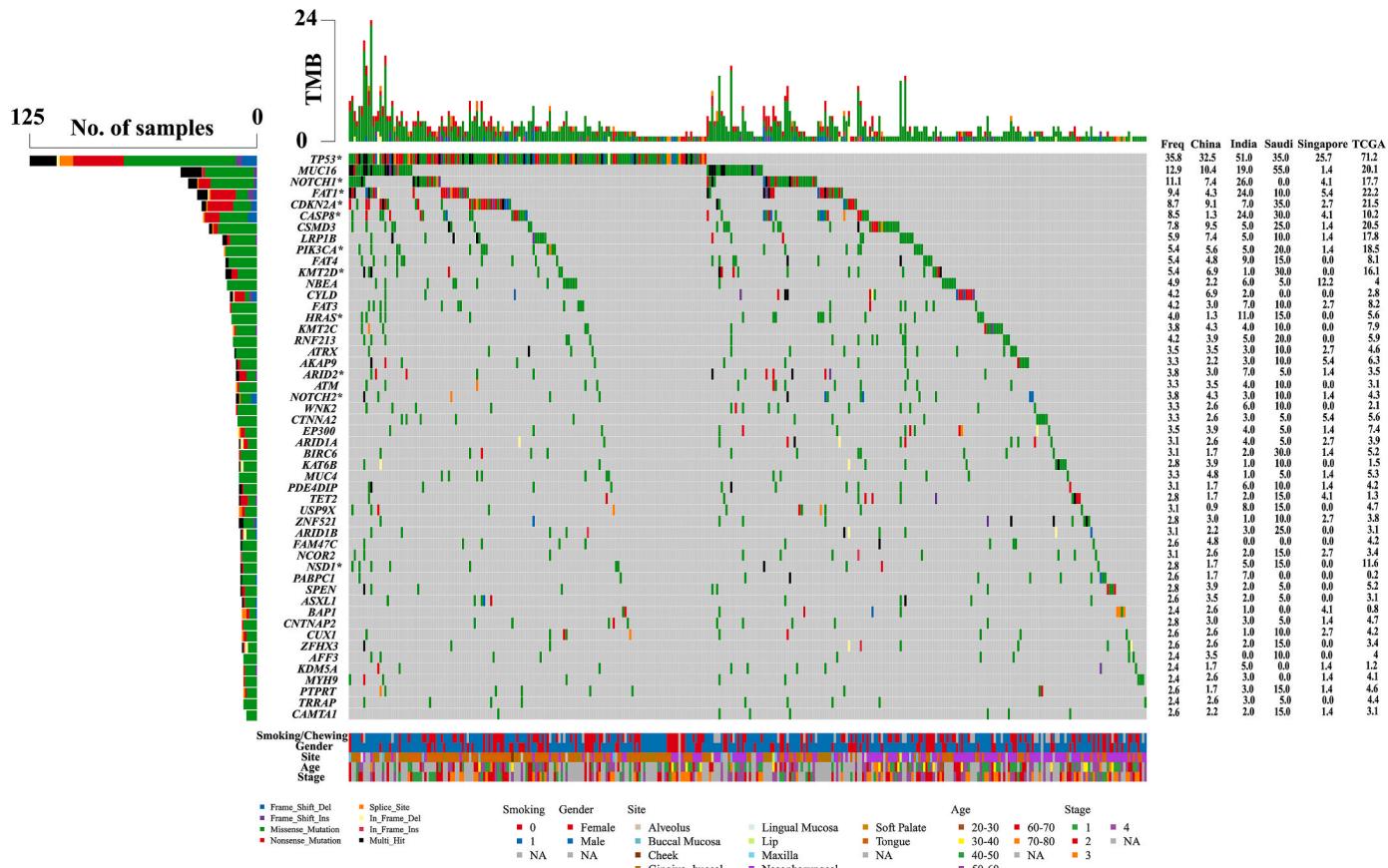


Fig. 2. Rainfall plot of the mutation profile for 50 most frequently occurring cancer census genes in the cohort comprising whole exome sequencing samples (India, China, SA, and Singapore). The Asterix represents the hallmark genes of HNSCC according to the cancer Genome Atlas (TCGA) study. The numeric data in the right represents the percentage frequency of the mutations across the study cohorts: all (Freq), Indian cohort (India), Chinese cohort (China), Saudi Arab cohort (Saudi), Singapore cohort (Singapore) and TCGA cohort (TCGA).

is negligible in China (1 %), while it is observed at 4 % in Singapore. NOTCH1 mutation is observed at 26 % in India, while TCGA reports 17.7 %. NOTCH1 frequencies were observed at 7.4 % in China, 4 % in Singapore and 5 % in SA cohorts. HRAS and PABPC1 mutations have frequencies of approximately 11 % and 7 % in India, compared to 5.6 % and 0.2 % in the TCGA dataset. HRAS is observed at a frequency of 15 % in South Asia, whereas in China, it appears in only 1.3 % of samples. Notably, no HRAS mutations are found in samples from Singapore. In the Chinese population, CYLD mutation was found to be approximately 6.2 % as against 2.8 % in the TCGA cohort. Similarly, ZNF479 and KAT6B occur at frequencies of 2 % and 4 % in the Chinese cohort as compared to 0.015 % and 1.5 % in the TCGA cohort. Furthermore, ZEB1, FIP1L1, and MSH6 mutations were at higher frequencies (~2 %) in China, compared to 0.015 % in the TCGA datasets.

MUC16 is highly mutated in the SA cohort with a frequency of 55 % compared to 20 % in the TCGA cohort. It is observed at frequencies as low as 1.35 % in Singapore and around 10 % in Chinese cohort. CDKN2A is also significantly mutated in the SA cohort (35 %) compared to TCGA (21.5 %), whereas its frequency is around 8 % in Indian and Chinese cohorts and 2 % only in Singapore cohort. ARID1B and PTPN13 were observed at 25 % in SA cohort in contrast to 3.1 % and 2.4 %, respectively, in TCGA. These genes are observed at 2 % in Indian cohort and 3 % in Chinese cohort, with no report in the Singapore population in this study. BIRC6 is mutated in 30 % of the SA cohort, while it occurs with a frequency of 5.2 % in TCGA and around 1.7 % in the other three countries. Furthermore, NCOR1 occurs at a frequency of 20 % in SA samples in comparison to 3.3 % in TCGA, around 1.8 % in China and Singapore cohorts respectively, while it is not present in the Indian cohort.

Singapore cohort exhibits NBEA mutation frequency of around 12 %, which is a notable difference compared to the TCGA frequency of 4 %. NBEA is mutated with frequencies 6 %, 5 %, and 2.1 % in India, SA, and China, respectively. Further, TET2, BAP1, and ARHGAP5 occur at frequencies around 4 %, which is only 1 % in the TCGA cohort. TET2 and ARHGAP5 are found at relatively higher frequencies in SA (15 %) compared to a frequency of around 2 % in India and China.

Some other notable genes where we observed significant variation in frequencies among the four countries include PIK3CA, KMT2D and CSMD3. PIK3CA is found to be mutated significantly in SA samples (20 %), while it occurs at a frequency of 5 % in both India and China. In the Singaporean cohort, the mutation frequency of PIK3CA is 1.35 % which is significantly different compared to other Asian countries. Similarly, KMT2D is mutated at a frequency of 30 %, while it is observed at 1 % in India, 7 % in China, and 0 % in Singapore. CSMD3 occurs with a frequency of 25 % in SA, while in India, China, and Singapore, the mutation frequencies are 5 %, 9.5 %, and 1.35 %, respectively. FAT1 is mutated at 24 % in India, while it is found at 4.3 % in China, 10 % in SA, and 5 % in Singapore. Interestingly, another gene family member FAT4 is reported at 15 % in the SA, but no FAT4 mutation exists in Singapore (or not yet reported). USP9X is reported at 8 % in India and 15 % in SA while it is detected at 0.86 % in China and 0 % in Singapore. SA cohort has 10 % mutation frequency for ATRX, AKAP9, NOTCH2, ZNF521, CUX1, POLG, ALK, PREX2, ANK1, ERBB2, SETDB1, DROSHA which are present around 2.1 % in India, 2.2 % in China and 2.3 % in Singapore. The genes NCOR2, TET2, PTPRT, CAMTA1, BRCA2, ATR, EGFR, and CLTCL are observed at a frequency of 15 % in the SA cohort, while in China, India, and Singapore cohorts, they are present around 1.67 %, 2 %, and 1.9 % respectively.

3.3. Dysregulated pathways in HNC

Pathway enrichment analysis was performed for the top 500 recurrently mutated genes identified using MSigDB in 425 WES samples. Fig. 3A and Supplementary Table 3 depicts key pathways identified in the cancer hallmark gene set. Apart from the common oncogenic pathways such as mitotic spindle and epithelial-mesenchymal transition, the

topmost pathway that got enriched was the KRAS signaling pathway in 173 patients, followed by NOTCH signaling (166 patients), TP53 (164 patients), and the Wnt-beta catenin signaling pathway (100 patients). Based on the above results, we initiated a detailed analysis of RAS signaling pathway genes from 4 countries and their frequencies using StringDB, which provided a comprehensive outlook of the functional protein association network for the Asian populations as represented in Fig. 3B. Three RAS signaling pathway genes, i.e., EGFR, ERBB2, and ALK, were mutated in all four countries. Few genes were exclusively found to be mutated in a specific country. For example, MET expression was only found in Singapore, whereas ERRFI1, PDGFRA, FLT3, IGF1R, and NTRK1 were exclusive to the Chinese population. Indian cohort showed higher frequency of mutations in MAPK, ARAF1, RASA1, BRAF, and RAF1 while, SA cohort showed mutations in ROS, RET, ALK, EGFR, CBL, and FGFR1. Mutations were found in all the regulators of the MAPK pathway, i.e., NF1, FGFR2, FGFR3, and ERBB3 in the Chinese cohort. Chinese and Indian cohorts have prevalence of the ERBB receptor tyrosine kinase family and the RAF family proteins. Among the RAS genes, HRAS mutations were found mostly in SA and Indian cohorts, followed by Chinese, while NRAS and KRAS were found only in Chinese and Singapore populations. Majority of the HRAS mutations are present in the oral cavity and salivary gland tumors, while NRAS and KRAS are more frequently mutated in nasopharyngeal and sinonasal tumors, respectively.

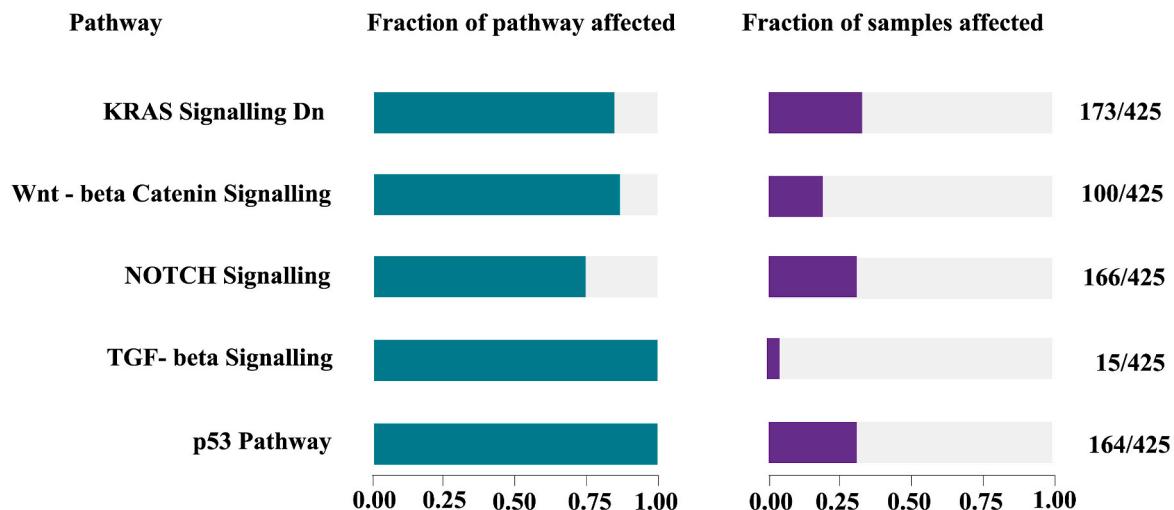
3.4. Comparative analysis of mutation patterns across smokers and non-smokers

We performed differential analysis to identify mutational profiles in smokers and non-smokers. Out of 425 samples, smoking status was known for 376 patients. Of these 219 were smokers and 157 were non-smokers. Comparative analysis identified a total of 26 genes that were statistically significant ($p < 0.05$, Log Rank Test). Higher number of mutations were observed in smokers in 21 genes whereas 5 genes showed higher number of mutations in non-smokers. (Fig. 4A). CASP8 and FAT1 were the top mutated genes in smokers (present in 31/219 samples), whereas LRP1, CFTR, ZNF493, ATRX, and ZNF99 were more frequently mutated in non-smokers. A closer look at the statistically significant genes from earlier analysis showed the TP53 had similar mutation frequency in both smokers and non-smokers while CASP8 and FAT1 mutations were seen in approximately 14 % of the smokers with predominantly non-sense and silent mutations (Fig. 4B). Similarly, NOTCH1, CDKN2A and MUC16 were at frequencies greater than 10 % in the smokers. Interestingly, HRAS was significantly mutated in smokers (7 %) compared to the non-smokers (1 %). Other notable gene is RYR2 (8 %) which exhibited higher mutation frequency in non-smokers as compared to smokers (Fig. 4B). The two subgroups of smokers and non-smokers were further subjected to pathway enrichment analysis (Fig. 4C). KRAS pathway was the predominant pathway in smokers, whereas the NOTCH signaling pathway was dysregulated in non-smokers. Other pathways include p53 and PI3K/AKT/mTOR signaling in the smoker cohort, and Wnt-beta and TGF-beta signaling cascades in non-smokers.

3.5. Mutually exclusive and co-occurring genes

We analysed the genomic data to understand the co-occurrence pattern among top mutated genes across all countries as well as within individual countries (Fig. 5). TP53 was mutually exclusive to CYLD (p -value: <0.05) gene across all the countries. This is in concordance with TCGA dataset where the gene pair is known to be mutually exclusive (p value = 0.039). Other interesting statistically significant concurrent pairs identified includes FAT1 and ARID2; MUC16 and KMT2D; MUC16 and FAT4; CDKN2A and TP53; CASP8 and HRAS; CASP8 and FAT1. CASP8 is known to co-occur with HRAS and FAT1 in TCGA data as well (p -value<0.001). In addition to these, four sets of co-occurring

(A)



(B)

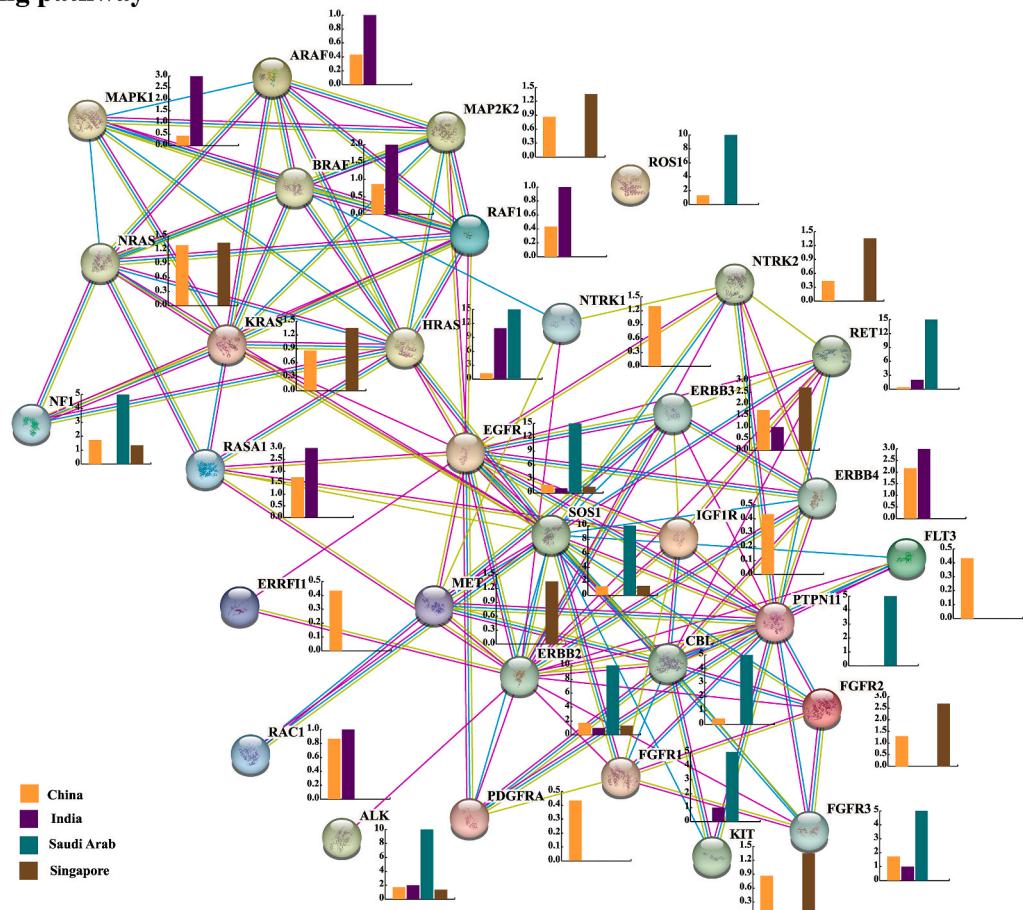
RAS signalling pathway

Fig. 3. A) Pathways enriched in Molecular Signature Database (MSigDB) from the top 500 recurrently mutated genes across all WES samples. The green bars represent pathways affected and purple bars shows the fraction of total genes from the respective pathway that are mutated in HNSCC. B) The Rat sarcoma virus (RAS) signaling network from StringDB: each node represents the genes in the pathway. The nodes are connected to each other based on available evidence from StringDB. The stronger the lines, more confident the interaction. The bar graphs next to the nodes represents the mutation frequencies of the respective genes across four countries. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

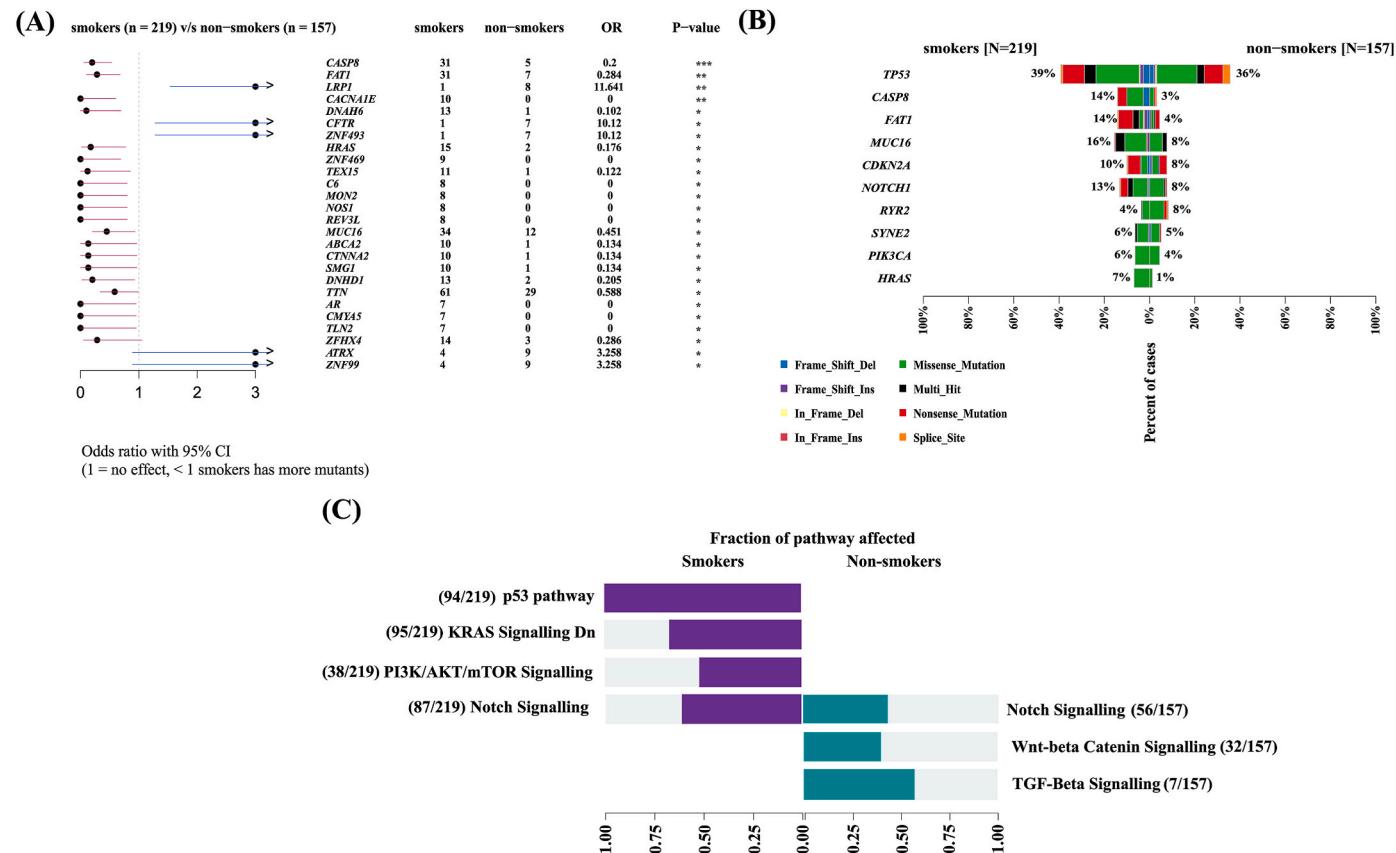


Fig. 4. (A) Forest plot depicting differences in mutated genes for smoker and non-smoker cohort. The odds ratios (OR) along with significant p-values of difference are shown. median hazard is shown as black circles (B) Cobar plot showing frequency differences in top mutated genes in both cohorts (smokers vs non-smokers). (C) Bar graph shows the enriched pathways in the top 500 recurrently mutated genes in smokers and non-smokers (MSigdb) as well as fraction of the pathway affected in terms of number of genes associated with the pathway. The fraction in brackets represent the samples affected out of the total number of samples in that category.

mutations were identified: LRP1B and ARID2, ASXL and CDKN2A, TP53 and CDKN2A, EP300 and ZNF521 across the Asian cohorts. Interestingly, ASXL1 and CDKN2A do not tend to co-occur in the TCGA data ($p\text{-value} = 0.716$ for co-occurrence). In the Indian dataset, the HRAS gene demonstrated mutual exclusivity with TP53, which was also observed in the TCGA ($p\text{-value} < 0.001$) head and neck datasets. In case of the SA cohort, no gene pair mutations were identified to be exclusive to each other, but 3 genes co-occurred with HRAS gene, namely CDKN2A, TP53 and CASP8. It is intriguing to note that TP53 exhibited mutual exclusivity with HRAS in both the Indian cohort and the SA cohorts only. Furthermore, few other co-occurring gene sets were also identified: POLD1 and ARID1B, NCOR1 and BRCA2, FAT1 and CLTCL1; ROBO2 and CLTCL1. Of these, NCOR1/BRCA2 and FAT1/CLTCL1 were observed as less likely to co-occur in TCGA cohort ($p = 0.106$) whereas, POLD1/ARID1B and ROBO2/CLTCL1 were also identified as co-occurring pairs in TCGA. None of the genes showed exclusivity or co-occurrence in the Singapore cohort, possibly due to limited sample sizes.

3.6. Identification of putative novel driver genes in Asian cohort

The driver gene analysis identified 893 potential drivers mapping to the CGCs. The mutational profiles of top 50 potential driver genes in the entire cohort are shown in [Supplementary Fig. 1](#). Top 10 potential drivers include namely: TP53, NOTCH1, FAT1, CDKN2A, CASP8, RYR2, KMT2D, CYLD, MAGEC1, and ABCA13. Further, we investigated the prognostic ability of the driver genes in the TCGA cohort. Several of the genes including TP53, CASP8 and NOTCH1 were well known in HNSCC. We further looked to those set of genes which were not reported in HNSCC as drivers but reported in other cancers. As shown in

[Supplementary Fig. 2](#), six putative novel driver genes, namely RYR2, PKHD1L1, ANK2, NBPF1, MMP16 and CPAMD8 were identified from the current analysis. The survival data for these genes were acquired from TCGA for this study to determine their prognostic ability in HNSCC. Patients with RYR2 mutations showed significantly lower survival in HNC patients as compared to unaltered group ($p\text{-value} = 0.0004$, Log-Rank test). Similarly, patients with mutations in PKHD1L1, ANK2, NBPF1, MMP16 and CPAMD8 showed poor prognosis ([Supplementary Fig. 2](#)). These genes could serve as potential therapeutic targets for HNC in Asian cohorts.

4. Discussion

We conducted a meta-analysis of published genomic studies on HNC from six Asian countries to identify similarities and differences in mutational profiles between Asian and non-Asian cohorts, and to find potential novel therapeutic targets specific to the Asian population. A total of 1016 samples from 38 studies were included. Our study found TP53 as the most prominently mutated tumor suppressor gene in HNSCC, with a lower frequency in Asian cohorts compared to TCGA. This lower rate of TP53 mutations is consistent with multiple Asian studies and might be attributed to environmental exposures, genetic predispositions, or distinct molecular pathways [8–14]. CDKN2A, another tumor suppressor gene associated with HNSCC, also showed a lower frequency, aligning with the pattern of TP53 mutations. CDKN2A, encoding cell cycle checkpoint inhibitor proteins p16 and p14, is associated with high tumor mutational burden (TMB) in HNC patients. These two genes are often co-occurring and are linked to the early stages of HNSCC formation [15,16]. CASP8, a tumor suppressor gene that triggers

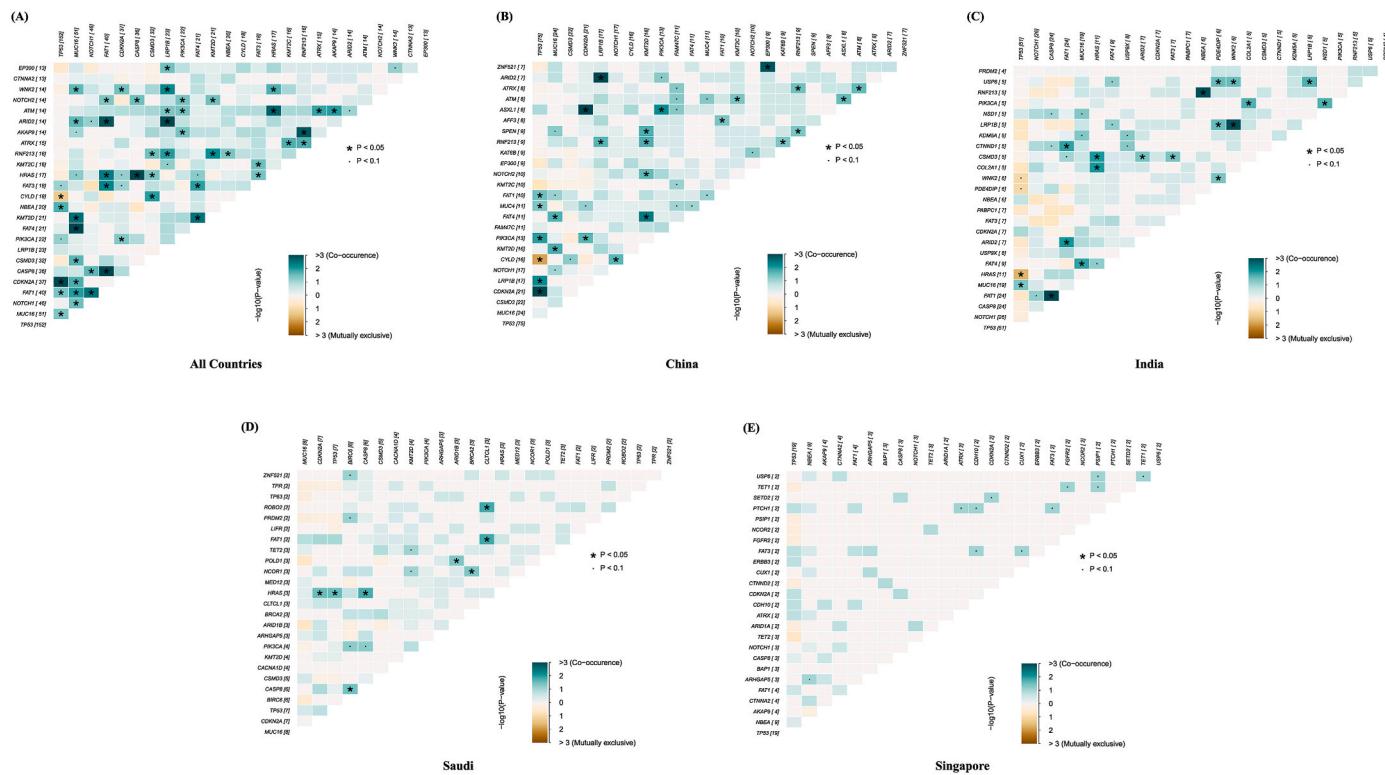


Fig. 5. Heat map showing the co-occurring and mutually exclusive genes among the top 25 recurrently mutated genes (A) across all four cohorts in current study (B) Chinese cohort (C) Indian cohort (D) SA cohort and (E) Singaporean cohort.

cell death via apoptotic pathways, was mutated more frequently in Indian and South Asian cohorts than in TCGA, and is associated with poor prognosis in HNSCC [17–20]. NOTCH1, which enhances tumorigenesis and disease progression [21,22], had higher frequencies in the Indian cohort, which is also reported in some independent studies [23,24]. Other genes like KMT2D and NSD1 had lower mutation frequencies in Asians.

Additionally, prognostic biomarkers such as CASP8 and FAT1 also showed lower frequency of mutations in Chinese cohorts. This warrants the need for more studies in the Asian cohorts to determine the true frequencies. At the same time, CYLD mutations were observed mutually exclusive to TP53 with higher frequencies in the Chinese cohort. CYLD regulates several signaling pathways and is associated with better survival in HPV-positive HNSCC [25–27]. Another interesting gene, MUC16 which has been linked to the proliferation and spread of cancer cells through the inhibition of natural killer cells and activation of JAK2-induced molecular pathways showed lower frequencies, especially in China and Singapore datasets, and relatively higher frequency in SA cohorts (50 %). MUC16 mutations are frequently seen in melanoma and some other cancers like hepatocellular, lung and gastric cancer [28,29]. Its role in HNSCC needs to be further explored, especially in Asian cohorts. Further, NBEA was found to be the second most frequently mutated gene in the Singapore cohort. NBEA is located within common fragile sites (CFS) in the genome [30] and potential driver mutations have been reported in renal, colorectal and mucinous salivary adenocarcinoma [30–32] but no reports in HNC to our knowledge. This could serve as a potential novel therapeutic target.

Pathway-level analysis revealed the KRAS signaling pathway as the most predominant driver in Asian HNC, contrasting with the PIK3CA/MTOR pathway found in over 90 % of HNSCC cases worldwide [14]. RAS mutations, particularly HRAS, were prevalent in Indian and SA cohorts, often linked to regional risk factors like betel quid and bidi smoking [16,33–35]. Recent studies by our group and other independent groups have shown prevalent HRAS mutations in HNSCC [36,37].

Both Indian and Chinese cohorts showed enrichment in RAF family proteins, suggesting potential targeting with RAF inhibitors. Furthermore, both the Indian and Chinese cohorts showed enrichment in RAF family proteins as well, which can potentially be targeted using RAF inhibitors [38–40]. Mutations in the regulators and activators of the MAPK pathway commonly found in Chinese cohort can be a potential target for therapies employing ERBB3 inhibitors, Cetuximab addition or EGFR inhibitors [41]. Further, the exclusive prevalence of MET mutations in Singaporean population can be targeted by Crizotinib, which is known for its rapid response to such alterations in HNSCC [42]. Mutational differences in smokers and non-smokers depicted the high enrichment of the KRAS signaling pathway, with smokers exhibiting high frequencies of CASP8 and FAT1 mutations. CASP8 gene is shown to be associated with smoking status in HNSCC in a TCGA cohort-based study [43]. Additionally, the smoker cohort exhibited significantly greater frequency of mutations in RYR2 gene, identified as a novel potential driver gene in our analysis for HNSCC. ANK2, a cytoskeletal protein, showed the least survival response in driver gene analysis. CPCRs like ANK2 are known to have a greater mutational frequency within the smoker subcategory of HNSCC [43]. Hence, there is a need to undertake systematic studies to understand these subgroups of HNSCC in detail.

Our meta-analysis revealed distinct mutational landscape of Asian cohorts. NBEA, RYR2 and ANK2 emerged as some of the potential novel driver genes, highlighting avenues for further research. However, limitations such as lack of WES datasets from all regions and small sample sizes underscore the necessity for more comprehensive, region-specific studies.

CRediT authorship contribution statement

Srikanth S. Manda: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Nafisa Arfa:** Data curation, Formal analysis,

Methodology, Visualization. Neha Sharma: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. Aparna R. Parikh: Conceptualization, Writing – review & editing. Thomas J. Roberts: Conceptualization, Writing – review & editing. Sewanti Limaye: Conceptualization, Writing – review & editing. Venkataraman Ramachandran: Resources, Software. Kumar Prabhash: Conceptualization, Methodology, Supervision. Moni A. Kuriakose: Supervision. Prashant Kumar: Conceptualization, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.oor.2024.100628>.

References

- [1] Johnson DE, Burtness B, Leemans CR, Lui VVY, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. *Nat Rev Dis Prim* 2020;6:92. <https://doi.org/10.1038/s41572-020-00224-3>.
- [2] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A Cancer J Clin* 2021;71:209–49. <https://doi.org/10.3322/caac.21660>.
- [3] Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. *Br Dent J* 2022;233:780–6. <https://doi.org/10.1038/s41415-022-5166-x>.
- [4] Solomon B, Young RJ, Rischin D. Head and neck squamous cell carcinoma: genomics and emerging biomarkers for immunomodulatory cancer treatments. *Semin Cancer Biol* 2018;52:228–40. <https://doi.org/10.1016/j.semancer.2018.01.008>.
- [5] Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DJ, Nguyen-Tân PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010;363:24–35. <https://doi.org/10.1056/NEJMoa0912217>.
- [6] Alsahafi E, Begg K, Amelio I, Raulf N, Lucarelli P, Sauter T, et al. Clinical update on head and neck cancer: molecular biology and ongoing challenges. *Cell Death Dis* 2019;10:540. <https://doi.org/10.1038/s41419-019-1769-9>.
- [7] Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, risk factors, and prevention of head and neck squamous cell carcinoma. *Med Sci* 2023;11:42. <https://doi.org/10.3390/medsci11020042>.
- [8] Klinakis A, Rampias T. TP53 mutational landscape of metastatic head and neck cancer reveals patterns of mutation selection. *EBioMedicine* 2020;58:102905. <https://doi.org/10.1016/j.ebiom.2020.102905>.
- [9] The Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015;517:576–82. <https://doi.org/10.1038/nature14129>.
- [10] Riaz N, Morris LG, Lee W, Chan TA. Unraveling the molecular genetics of head and neck cancer through genome-wide approaches. *Genes Dis* 2014;1:75–86. <https://doi.org/10.1016/j.gendis.2014.07.002>.
- [11] Wu H, Eckhardt CM, Baccarelli AA. Molecular mechanisms of environmental exposures and human disease. *Nat Rev Genet* 2023;24:332–44. <https://doi.org/10.1038/s41576-022-00569-3>.
- [12] Tsang CM, Lui VVY, Bruce JP, Pugh TJ, Lo KW. Translational genomics of nasopharyngeal cancer. *Semin Cancer Biol* 2020;61:84–100. <https://doi.org/10.1016/j.semancer.2019.09.006>.
- [13] Bruce JP, Yip K, Bratman SV, Ito E, Liu F-F. Nasopharyngeal cancer: molecular landscape. *J Clin Oncol* 2015;33:3346–55. <https://doi.org/10.1200/JCO.2015.60.7846>.
- [14] Li Q, Tie Y, Alu A, Ma X, Shi H. Targeted therapy for head and neck cancer: signaling pathways and clinical studies. *Signal Transduct Targeted Ther* 2023;8:31. <https://doi.org/10.1038/s41392-022-01297-0>.
- [15] Deneka AY, Baca Y, Serebriiskii IG, Nicolas E, Parker MI, Nguyen TT, et al. Association of TP53 and CDKN2A mutation profile with tumor mutation burden in head and neck cancer. *Clin Cancer Res* 2022;28:1925–37. <https://doi.org/10.1158/1078-0432.CCR-21-4316>.
- [16] Desai SS, K RR, Jain A, Bawa PS, Dutta P, Atre G, et al. Multidimensional mutational profiling of the Indian HNSCC sub-population provides IRAK1, a novel driver gene and potential druggable target. *Front Oncol* 2021;11:723162. <https://doi.org/10.3389/fonc.2021.723162>.
- [17] Usman S, Jamal A, Teh M-T, Waseem A. Major molecular signaling pathways in oral cancer associated with therapeutic resistance. *Front Oral Health* 2021;1:603160. <https://doi.org/10.3389/froh.2020.603160>.
- [18] Stupack DG. Caspase-8 as a therapeutic target in cancer. *Cancer Lett* 2013;332:133–40. <https://doi.org/10.1016/j.canlet.2010.07.022>.
- [19] Ando M, Kawazu M, Ueno T, Fukumura K, Yamato A, Soda M, et al. Cancer-associated missense mutations of caspase-8 activate nuclear factor- κ B signaling. *Cancer Sci* 2013;104:1002–8. <https://doi.org/10.1111/cas.12191>.
- [20] Olsson M, Zhivotovsky B. Caspases and cancer. *Cell Death Differ* 2011;18:1441–9. <https://doi.org/10.1038/cedd.2011.30>.
- [21] Shah PA, Huang C, Li Q, Kazi SA, Byers LA, Wang J, et al. NOTCH1 signaling in head and neck squamous cell carcinoma. *Cells* 2020;9:2677. <https://doi.org/10.3390/cells9122677>.
- [22] Lee SH, Do SI, Lee HJ, Kang HJ, Koo BS, Lim YC. Notch1 signaling contributes to stemness in head and neck squamous cell carcinoma. *Lab Invest* 2016;96:508–16. <https://doi.org/10.1038/labinvest.2015.163>.
- [23] Agrawal N, Frederick MJ, Pickering CR, Bettogowda C, Chang K, Li RJ, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* 2011;333:1154–7. <https://doi.org/10.1126/science.1206923>.
- [24] Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science* 2011;333:1157–60. <https://doi.org/10.1126/science.1208130>.
- [25] Hajek M, Sewell A, Kaech S, Burtness B, Yarbrough WG, Isaeva N. TRAF3/CYLD mutations identify a distinct subset of human papillomavirus-associated head and neck squamous cell carcinoma. *Cancer* 2017;123:1778–90. <https://doi.org/10.1002/cncr.30570>.
- [26] Williams EA, Montesinos M, Alexander BM, Ramkisson SH, Elvin JA, Ross JS, et al. CYLD mutation characterizes a subset of HPV-positive head and neck squamous cell carcinomas with distinctive genomics and frequent cylindromatosis-like histologic features. *Mod Pathol* 2021;34:358–70. <https://doi.org/10.1038/s41379-020-00672-y>.
- [27] Mathis BJ, Lai Y, Qu C, Janicki JS, Cui T. CYLD-mediated signaling and diseases. *Curr Drug Targets* 2015;16:284–94. <https://doi.org/10.2174/138945011566141024152421>.
- [28] Aithal A, Rauth S, Kshirsagar P, Shah A, Lakshmanan I, Junker WM, et al. MUC16 as a novel target for cancer therapy. *Expert Opin Ther Targets* 2018;22:675–86. <https://doi.org/10.1080/14728222.2018.1498845>.
- [29] Zhang X-Y, Hong L-L, Ling Z. MUC16: clinical targets with great potential. *Clin Exp Med* 2024;24:101. <https://doi.org/10.1007/s10238-024-01365-5>.
- [30] O’Neal J, Gao F, Hassan A, Monahan R, Barrios S, Lee I, et al. Neurobeachin (NBEA) is a target of recurrent interstitial deletions at 13q13 in patients with MGUS and multiple myeloma. *Exp Hematol* 2009;37:234–44. <https://doi.org/10.1016/j.exphem.2008.10.014>.
- [31] Panaccione A, Zhang Y, Mi Y, Mitani Y, Yan G, Prasad ML, et al. Chromosomal abnormalities and molecular landscape of metastasizing mucinous salivary adenocarcinoma. *Oral Oncol* 2017;66:38–45. <https://doi.org/10.1016/j.joraloncology.2016.12.011>.
- [32] Liu Z, Zhang Y, Dang Q, Wu K, Jiao D, Li Z, et al. Genomic alteration characterization in colorectal cancer identifies a prognostic and metastasis biomarker: fam83A/Ido1. *Front Oncol* 2021;11:632430. <https://doi.org/10.3389/fonc.2021.632430>.
- [33] Novopolsky O, Jagadeeshan S, Regev O, Menashe I, Elkabets M. Worldwide prevalence and clinical characteristics of RAS mutations in head and neck cancer: a systematic review and meta-analysis. *Front Oncol* 2022;12:838911. <https://doi.org/10.3389/fonc.2022.838911>.
- [34] Chauhan R, Trivedi V, Rani R, Singh U. A study of head and neck cancer patients with reference to tobacco use, gender, and subsite distribution. *South Asian J Cancer* 2022;11:46–51. <https://doi.org/10.1055/s-0041-1740601>.
- [35] Bagal S, Budukh A, Singh Thakur J, Dora T, Qayyumi B, Khanna D, et al. Head and neck cancer burden in India: an analysis from published data of 37 population-based cancer registries. *Eccancermedicalscience* 2023;17. <https://doi.org/10.3332/ecancer.2023.1603>.
- [36] Gauthaman A, Moorthy A. Prevalence of K-ras codon 12 mutations in Indian patients with head and neck cancer. *Indian J Clin Biochem* 2021;36:370–4. <https://doi.org/10.1007/s12291-020-00882-w>.
- [37] Sambath J, Noronha V, Manda SS, Mishra R, Chandran P, Patil V, et al. Whole exome sequencing uncovers HRAS mutations as potential mediators of resistance to metronomic chemotherapy. *Gene* 2024;893:147952. <https://doi.org/10.1016/j.gene.2023.147952>.
- [38] Kim A, Cohen MS. The discovery of vemurafenib for the treatment of BRAF-mutated metastatic melanoma. *Expert Opin Drug Discov* 2016;11:907–16. <https://doi.org/10.1080/17460441.2016.1201057>.
- [39] Lang M, Longerich T, Anamaterou C. Targeted therapy with vemurafenib in BRAF (V600E)-mutated anaplastic thyroid cancer. *Thyroid Res* 2023;16:5. <https://doi.org/10.1186/s13044-023-00147-7>.
- [40] Dao BD, Lingvay I, Sailors J, Landay M, Shapiro G. First-line use of vemurafenib to enable thyroidectomy and radioactive iodine ablation for BRAF-positive metastatic papillary thyroid carcinoma: a case report. *J Investig Med High Impact Case Rep* 2015;3:2324709615603723. <https://doi.org/10.1177/2324709615603723>.

- [41] Bahar ME, Kim HJ, Kim DR. Targeting the RAS/RAF/MAPK pathway for cancer therapy: from mechanism to clinical studies. *Signal Transduct Targeted Ther* 2023; 8:455. <https://doi.org/10.1038/s41392-023-01705-z>.
- [42] Chu LP, Franck D, Parachoniak CA, Gregg JP, Moore MG, Farwell DG, et al. MET genomic alterations in head and neck squamous cell carcinoma (HNSCC): rapid response to Crizotinib in a patient with HNSCC with a novel MET R1004G mutation. *Oncol* 2019;24:1305–8. <https://doi.org/10.1634/theoncologist.2019-0230>.
- [43] Ghasemi F, Prokopec SD, MacNeil D, Mundi N, Gameiro SF, Howlett C, et al. Mutational analysis of head and neck squamous cell carcinoma stratified by smoking status. *JCI Insight* 2019;4:e123443. <https://doi.org/10.1172/jci.insight.123443>.