Squamous Cell Carcinoma of the Oral Tongue in Young Non-Smokers Is Genomically Similar to Tumors in Older Smokers

Curtis R. Pickering1, Jiexin Zhang2, David M. Neskey1, Mei Zhao1, Samar A. Jasser1, Jiping Wang1, Alexandra Ward1, C. Jillian Tsai3, Marcus V. Ortega Alves1, Jane H. Zhou4, Jennifer Drummond6, Adel K. El-Naggar4, Richard Gibbs6,7, John N. Weinstein2,5, David A. Wheeler6, Jing Wang2, Mitchell J. Frederick1, and Jeffrey N. Myers1

Abstract

Purpose: Epidemiologic studies have identified an increasing incidence of squamous cell carcinoma of the oral tongue (SCCOT) in younger patients.

Experimental Design: DNA isolated from tongue tumors of young (<45 years, non-smokers) and old (>45 years) patients at was subjected to whole-exome sequencing and copy-number analysis. These data were compared with data from similar patients in the TCGA (The Cancer Genome Atlas) project.

Results: In this study, we found that gene-specific mutation and copy-number alteration frequencies were similar between young and old patients with SCCOT in two independent cohorts. Likewise, the types of base changes observed in the young cohort were similar to those in the old cohort even though they differed in smoking history. TCGA data also demonstrate that the genomic effects of smoking are tumor site–specific, and we find that smoking has only a minor impact on the types of mutations observed in SCCOT.

Conclusions: Overall, tumors from young patients with SCCOT appear genomically similar to those of older patients with SCCOT, and the cause for the increasing incidence of young SCCOT remains unknown. These data indicate that the functional impact of smoking on carcinogenesis in SCCOT is still poorly understood. Clin Cancer Res; 20(14); 3842–8. ©2014 AACR.
patients (6, 7). Mutation data can be found in Supplementary Table S3.

A similar cohort of patients was identified from within the TCGA HNSCC project. YT patients were chosen based on oral tongue site, age less than 46 and lifelong non-smoker or reformed smoker with unknown pack-year history. OT patients were chosen based on oral tongue site and age older than 45. Patient characteristics are shown in Supplementary Table S1. Mutation and copy-number data were obtained from the TCGA pan-cancer project (8) and the Synapse website (www.synapse.org; doi:10.7303/syn300013; ref. 9).

The statistical significance of mutation and copy-number frequencies was determined by the Fisher exact test. Mutation-type frequencies were determined for each patient, excluding indels and multiple-base mutations. Group frequencies were determined by averaging the individual patient frequencies: this prevented skewing of the data by patients with large numbers of mutations. Mutation-type frequencies were compared by using the $t$ test on arcsine transformed data.

Results

DNA from 16 YT and 28 OT patients treated at the MD Anderson Cancer Center (MDA) was subjected to whole-exome sequencing (Supplementary Table S1). A median of 29 and 83 mutations was identified in the YT and OT tumors, respectively. The elevated number of mutations in the OT is not surprising, because mutation number is known to be related to both age and smoking (10). This finding was validated in The Cancer Genome Atlas (TCGA) cohort, in which the median number of mutations increased from 63 in YT to 112 in OT. Supplementary Fig. S1 shows the relationships between mutation number and age. We next asked whether any specific genes were mutated at a different frequency in the YT cohort. Because of the low sample size, we limited our initial search to the most significantly mutated genes identified by MutSig in the entire TCGA HNSCC project (HNSC).

$TP53$ (unadjusted $P = 0.015$) showed a slight increase in mutation frequency in the MDA YT cohort, but it was not statistically significant when adjusted for multiple testing (Table 1; Fig. 1). Comparable analysis was then performed on the TCGA cohort. The $TP53$ mutation frequency was also elevated in the TCGA YT patients (Table 1; Fig. 1). To increase statistical power the two cohorts were combined. Three genes showed trends toward statistical significance; $FAT1$, $TP53$, and $PIK3CA$ (Table 1; Fig. 1). However, none of those genes showed a statistically significant difference between the combined YT and OT patient cohorts. The trend of increased $TP53$ mutations in YT is provocative because the YT lacks exposure to cigarette smoke, which has been associated with $TP53$ mutations. $FAT1$ and $PIK3CA$ showed a lower mutation frequency in the YT cohort. Mutation frequencies for HPV$^+$ tumors and the entire TCGA cohort are shown for comparison (Table 1). An analysis of mutation frequencies in all genes in the combined cohorts was also performed, but no genes were found to be significantly different. Additional subset analysis for really YTs (<30 years), OT smokers, and OT non-smokers are shown in Supplementary Table S2.

Whole-genome copy-number analysis was also performed. We compared the number and size of copy-number

<table>
<thead>
<tr>
<th>Table 1. Mutation frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Cohort (N)</strong></td>
</tr>
<tr>
<td><strong>Median number</strong></td>
</tr>
<tr>
<td><strong>mutations</strong></td>
</tr>
<tr>
<td><strong>FAT1</strong></td>
</tr>
<tr>
<td><strong>TP53</strong></td>
</tr>
<tr>
<td><strong>PIK3CA</strong></td>
</tr>
<tr>
<td><strong>CDKN2A</strong></td>
</tr>
<tr>
<td><strong>CASP8</strong></td>
</tr>
<tr>
<td><strong>AJUBA</strong></td>
</tr>
<tr>
<td><strong>NOTCH1</strong></td>
</tr>
<tr>
<td><strong>MLL2</strong></td>
</tr>
<tr>
<td><strong>NSD1</strong></td>
</tr>
<tr>
<td><strong>HLA-A</strong></td>
</tr>
<tr>
<td><strong>TGFB1</strong></td>
</tr>
</tbody>
</table>
alterations (CNA) between the YT and OT cohorts, identifying an average of 129 CNAs in YT samples and 72 in the OT samples from the MDA cohort (Table 2), and 62 and 64 CNAs in the YT and OT samples from the TCGA cohort, respectively (Table 2). These differences were not statistically significant, although the YT had a smaller mean segment length for copy-number gains ($P = 0.00817$; Table 2) in the MDA cohort. There was no difference in segment length of losses in between YT and OT in either cohort (Table 2). The mean copy number of gains or losses was also not different for gains or losses in either cohort (Table 2). Finally, no specific genomic regions were found to be significantly different between the YT and OT patients in either the MDA cohort or the TCGA cohort. Analysis of some genes frequently gained or lost in HNSCC revealed modest differences in frequency between YT and OT (Fig. 1). For example, gains in FADD and PIK3CA were less frequent in YT, but the difference was not significant (Fig. 1). Overall, the CNAs were very similar between the YT and OT cohorts, and the regions of copy-number change were similar to those reported previously (6).

Because smoking is known to leave its mark on the genome by causing certain types of mutations, we compared mutation types in the YT and OT patients. Taking into account directional redundancy, six types of mutations can be distinguished. The frequencies of these six types of mutations have been shown to vary across tumor types (8, 11), but we found no significant difference in that respect for YT and OT patients in either the MDA or TCGA cohort (Fig. 2A). The profiles resembled that of all head and neck tumors in the TCGA project. The profile, however, was distinct from that of HPV$^+$ tumors or laryngeal tumors (Fig. 2A). HPV$^+$ tumors show an increase in C>T mutations ($P < 0.0001$) and decreases in C>A, A>T (both $P < 0.0001$), and A>G ($P = 0.0074$) mutations when compared with HPV$^-$ HNSC tumors. Laryngeal tumors show a decrease in C>T mutations ($P < 0.0001$) and increases in C>A and A>T mutations (both $P < 0.0001$) when compared with non-laryngeal tumors (Fig. 2A). It was expected that OT tumors would exhibit a mutation signature related to cigarette smoking when compared with the YT tumors from non-smokers. The similarity between YT and OT mutation signatures could indicate either the presence of a smoking signature in the YT tumors or a lack of a smoking signature in the OT tumors. To address those alternative possibilities, we investigated the smoking signatures in other tumor sites from the TCGA project.

The mutational events linked to smoking are traditionally reported as an increase in C>A mutations and a decrease in C>T mutations. TCGA collected smoking history information for lung adenocarcinoma (LUAD), bladder urothelial carcinoma (BLCA), and HNSC. Each of these tumor types was analyzed for the presence of a smoking mutation signature. In LUAD tumors C>A mutations increased 23.7 percentage points in smokers when compared with non-smokers, and C>T mutation decreased.

![Figure 1](image-url)  
**Figure 1.** Frequency of common genomic alterations in YT and OT. The frequency of each event in the MDA cohort is shown by a bar to the left of center and the frequency in the TCGA cohort is shown by a bar to the right of center.

### Table 2. Copy-number alterations

<table>
<thead>
<tr>
<th></th>
<th>MDA</th>
<th>TCGA</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>YT</td>
<td>OT</td>
<td>P</td>
</tr>
<tr>
<td>Mean number</td>
<td>All CNAs</td>
<td>129.00</td>
<td>71.56</td>
</tr>
<tr>
<td></td>
<td>Gains</td>
<td>57.94</td>
<td>36.52</td>
</tr>
<tr>
<td></td>
<td>Losses</td>
<td>71.06</td>
<td>35.04</td>
</tr>
<tr>
<td>Mean copy number</td>
<td>Gains</td>
<td>3.45</td>
<td>3.09</td>
</tr>
<tr>
<td></td>
<td>Losses</td>
<td>11.41</td>
<td>1.41</td>
</tr>
<tr>
<td>Mean segment length (Mb)</td>
<td>Gains</td>
<td>7.37</td>
<td>20.50</td>
</tr>
<tr>
<td></td>
<td>Losses</td>
<td>11.39</td>
<td>15.24</td>
</tr>
</tbody>
</table>

$^a$MDA values are normalized absolute copy numbers.

$^b$TCGA values are log$_2$ centered at 0.
We next examined the impact of smoking on the mutation profile of HNSC. The changes in HNSC were generally similar to those in LUAD but of lower magnitude. We found a decrease in C>T of 13.2 percentage points and an increase in C>A of 7.3 percentage points that were consistent with the changes in LUAD (Fig. 2B and C). HNSC also demonstrated increases in A>G and A>T of 3.8 and 3.1 percentage points, respectively (Fig. 2B and C and Supplementary S2A).

Having defined the smoking signature, we could then ask whether YT and OT tumors exhibit such a signature. For this analysis, we combined the YT and OT tumors into one cohort (YT/OT) and compared it with the HNSC smoking and non-smoking cohorts. The YT/OT cohort had a lower C>A frequency and a higher C>T frequency when compared with the HNSC smoking cohort (Fig. 2C), indicating that the YT/OT tumors lack the most definitive characteristics of the smoking signature. However, the YT/OT also demonstrated increased A>G when compared with the HNSC non-smoking tumors and an intermediate level of A>T when compared with both smoking and non-smoking cohorts (Fig. 2C). Similar results were obtained when the YT and OT cohorts were separately compared with the non-smoking and smoking cohorts, although with lower statistical significance (data not shown) or when smoking and non-smoking cohorts within the OT cohort were compared with each other or YT (Supplementary Fig. S2B). Overall, the mutation profile of the YT/OT tumors seems most similar to the profile of non-smokers.

Because both YT and OT tumors lack a smoking signature, we next searched for the source of the smoking signature in HNSC. We found that tumors from the laryngeal subsite exhibit a smoking signature when compared with tumors from all other sites (Fig. 2A and C and Supplementary S2A and S2B) with an increase in C>A mutations and a decrease in C>T mutations. In addition, the smoking signature within the oral tongue subsite is less recognizable than the overall HNSC smoking signature (Fig. 2C). Those observations suggest that the smoking signature in HNSC is largely driven by the laryngeal subsite.

Discussion

Because the incidence of YT cancer has reportedly been increasing (1–3), we hypothesized that YT and OT tumors would prove to be genomically distinct. We found, to the contrary, that the two are genomically quite similar in their CNAs, mutations and types of mutations. Those conclusions were consistent for two independent patient cohorts. Although this study is the largest genomic analysis of YT tumors, it is still limited by the small sample size. The rarity of these tumors has also prevented ideal matching of the YT and OT cohorts, which differ in a few parameters (Supplementary Table S1), and resulted in a YT cohort that is not enriched for women as some have reported (2, 4).

The mutation frequencies of FAT1 and TP53 showed trends toward statistical significance, but would require larger cohorts for validation (estimated 170 and 189 total
patients, respectively). \textit{FAT1} is inactivated in a large proportion of oral cavity tumors, either by mutation or copy-number deletion (6). \textit{FAT1} may be related to differentiation, migration, or Hippo signaling in HNSCC (6, 12–14), but its role is largely still unknown. The lower frequency of \textit{FAT1} mutations in YT could suggest that its inactivation is not necessary for tumor formation or that it is somehow related to smoking status. However, because older smokers are expected to have more background mutations it is possible that the difference is an artifact of the increased frequency of mutations. Because \textit{TP53} mutations are more frequent in YT patients, this result is not likely to be an artifact. \textit{TP53} mutations are likely to be impactful on the tumor progression in YT, but a causative factor leading to increased \textit{TP53} mutations in YT is still unknown.

We also noted that OT tumors do not exhibit a mutation signature associated with smoking. The mutation signature from smoking seems to be site-specific, with laryngeal SCC exhibiting the strongest smoking signature among the HNSCC subsites. SCCOT did not exhibit a robust mutational smoking signature, in that regard echoing the TCGA findings in urothelial bladder cancer (15). Overall, the mutational signature that has been defined by analysis of lung tumors does not seem to reflect the effect of smoking on other tissue sites. Because smoking is clearly an epidemiologic risk factor for SCCOT (5), the mechanism of carcinogenesis may differ among sites. Cigarette smoke has been shown to be both a tumor initiator and a tumor promoter (16, 17). Only the role of smoke as a tumor initiator should leave a mutational signature on the DNA. Therefore, it is clear that smoke is a tumor initiator in lung cancer and laryngeal SCC. However, it may act primarily as a tumor promoter in SCCOT and BLCA. It is also possible that epigenetic alterations are involved because we have not yet examined the DNA methylation status of these tumors.

Alcohol consumption is a known risk factor for HNSCC. In some patients alcohol is likely to contribute as an initiator. Ethanol is not mutagenic, but its metabolite, acetaldehyde, can lead to mutations, primarily C>T (18). Because C>T is the most common mutation in many tumor types (8, 11), other factors must also contribute to their frequency. Analysis of the TCGA HNSC data for an alcohol mutation signature resulted in a less dramatic version of the smoking signature with decreases in C>T and increases in C>A in heavy drinkers (data not shown). That observation contrasts with the expected increase in C>T that could be caused by acetaldehyde. Because many heavy drinkers are also heavy smokers, it is likely that the smoking signature masks any alcohol signature. Therefore, the impact of alcohol use is still unclear.

These findings do not lessen the importance of smoking and drinking cessation programs to reduce SCCOT incidence. However, they suggest possible new avenues for prevention or treatment of SCCOT. If an unknown tumor initiator is acting in SCCOT, its identification could lead to new prevention strategies. Our analysis of the mutation profile does not indicate an obvious tumor initiator for SCCOT. It may, however, eliminate some candidates (e.g., many environmental mutagens) that would be expected to leave a mutation signature. It is also possible that the same initiator is acting in both YT and OT patients.

If smoking is indeed a tumor promoter in SCCOT, it will be important to determine the mechanism. In lung cancer, smoke has been shown to function as a tumor promoter by increasing inflammation (16), and that might also be the case in SCCOT. Inflammation from other causes promotes many tumor types including; esophageal, pancreatic, and colorectal (19). That connection suggests that inflammation is an important factor in non–smoking-related SCCOT, as previously reported (20, 21). Prevention strategies that include anti-inflammatory agents may prove useful. Although the results of clinical trials of COX inhibitors for prevention of SCCOT have not generally been positive (20, 22), agents that target other inflammatory pathways (such as the NF-kB pathway) still look very promising (20).

Another factor that could be promoting SCCOT is the oral microbiome. The oral cavity is home to a diverse population of microorganisms (23, 24), even in healthy individuals, and some of those microorganisms could also play a role in carcinogenesis (21, 25). They may produce carcinogenic compounds such as acetaldehyde, inducing local inflammation, or they may act through still-undiscovered mechanisms (21, 24, 25). Numerous microorganisms have already been associated with the development of oral cancer (21, 24, 25). Included are species of \textit{Streptococcus}, \textit{Neisseria}, and \textit{Candida}. Conversely, the “normal” microbiome may protect against carcinogenesis and a disrupted microbiome may cause it to no longer be protective (23, 24). Finally, it is unlikely that microorganisms are directly oncogenic in the way that HPV is oncogenic in oropharyngeal cancer. HPV is the only virus detected in SCCOT, and it is found in only 5% of cases. In addition, we found that HPV leaves a distinct mutation signature and expect that other pathogens would also leave distinct mutation signatures. However, we detect no such signature. In sum, the oral microbiome is likely to be involved in SCCOT in some way, but very little is understood about its causal role or the specific organisms involved.

Finally, YT cancer is troubling in that it defies the stereotype of HNSCC, which generally occurs in older men with long histories of heavy smoking and drinking (5). That observation is even more disturbing because the incidence of YT cancer is increasing (1–3). Hence, there is a strong interest in identifying the cause. Clinically, the cancer-specific outcomes for YT patients are similar to those for OT patients (4, 26–29), and we have shown here that the genomic alterations in YT tumors are also similar to those in OT tumors. Therefore, the only evidence suggesting that they are a unique cohort is epidemiologic (1–3). Although statistically significant, the epidemiologic findings are based on a low number of patients, because YT tumors are still quite rare, and the apparent difference
could be a statistical anomaly. Additional epidemiologic studies will be necessary to validate the original reports. Our data indicate that the mechanistic cause of YT cancer is likely to be similar to the mechanistic cause of OT cancer. For example, if smoking causes inflammation to promote OT perhaps another factor that also induces inflammation is promoting YT, resulting in a genomically and clinically similar disease. Similarly a factor that disrupts the oral microbiome in YT could promote carcinogenesis in a way similar to the mechanism in OT. However, the identity of a causative factor in YT cancer, if it exists, is still unknown.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: C.R. Pickering, D.M. Neskey, C.J. Tsai, M.J. Frederick, J.N. Myers
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): C.R. Pickering, M. Zhao, S.A. Jasser, J. Wang, A. Ward, C.J. Tsai, M.V.O. Alves, R. Gibbs, D.A. Wheeler, J. Wang, J.N. Myers

References

Grant Support
This work was supported by the Cancer Prevention Research Institute of Texas grant RP100233 (to J.N. Myers); NIH/National Institute of Dental and Craniofacial Research grant RC2DE020958 (to J.N. Myers); NIH Specialized Program of Research Excellence grants P50CA097917; Cancer Center Support grant P30CA16672; and an Institutional Research grant from the University of Texas M.D. Anderson Cancer Center (to M.J. Frederick) and the PanCan Program.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 6, 2014; revised April 22, 2014; accepted April 23, 2014; published OnlineFirst May 29, 2014.

Acknowledgments
The authors thank patients for their courage and generosity and the members of the Myers laboratory for moral support, technical support, and helpful discussions.

Writing, review, and/or revision of the manuscript: C.R. Pickering, J. Zhang, D.M. Neskey, C.J. Tsai, J.H. Zhou, A.K. El-Naggar, J.N. Weinstein, M.J. Frederick, J.N. Myers
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M. Zhao, S.A. Jasser, A. Ward, M.V.O. Alves, J. Drummond, R. Gibbs, J.N. Myers
Study supervision: R. Gibbs, D.A. Wheeler, J.N. Myers
Running lab experiments: C.J. Tsai
Identify tumors for DNA isolation: J.H. Zhou
Clinical Cancer Research

Squamous Cell Carcinoma of the Oral Tongue in Young Non-Smokers Is Genomically Similar to Tumors in Older Smokers


Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-14-0565

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2014/07/16/1078-0432.CCR-14-0565.DC1

Cited articles
This article cites 27 articles, 8 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/20/14/3842.full.html#ref-list-1

Citing articles
This article has been cited by 4 HighWire-hosted articles. Access the articles at:
/content/20/14/3842.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.