Ancestral-derived effects on the mutational landscape of laryngeal cancer

Meganathan P. Ramakodi a,b,c,d, Rob J. Kulathinal b,c,d, Yujin Chung b,d, Ilya Serebriiskii e,f, Jeffrey C. Liu a,e, Camille C. Ragin a,c,g,h,⁎

a Cancer Prevention and Control Program, Fox Chase Cancer Center-Temple Health, Philadelphia, PA 19111, USA
b Department of Biology, Temple University, Philadelphia, PA 19122, USA
c African-Caribbean Cancer Consortium
d Center for Computational Genetics and Genomics, Temple University, Philadelphia, PA 19122, USA
e Developmental Therapeutics, Fox Chase Cancer Center- Temple Health, Philadelphia, PA 19111, USA
f Kazan Federal University, Kazan, Russia
g Department of Otolaryngology - Head and Neck Surgery, Temple University School of Medicine, Philadelphia, PA 19140, USA
h College of Public Health, Temple University, Philadelphia, PA 19122, USA

ARTICLE INFO

Article history:
Received 17 August 2015
Received in revised form 26 November 2015
Accepted 21 December 2015
Available online 22 December 2015

Keywords:
Laryngeal cancer
Cancer genomics
African-Americans
European-Americans
Mutational landscapes
Context nucleotides

ABSTRACT

Laryngeal cancer disproportionately affects more African-Americans than European-Americans. Here, we analyze the genome-wide somatic point mutations from the tumors of 13 African-Americans and 57 European-Americans from TCGA to differentiate between environmental and ancestrally-inherited factors. The mean number of mutations was different between African-Americans (151.31) and European-Americans (277.63). Other differences in the overall mutational landscape between African-American and European-American were also found. The frequency of C > A, and C > G were significantly different between the two populations (p-value < 0.05). Context nucleotide signatures for some mutation types significantly differ between these two populations. Thus, the context nucleotide signatures along with other factors could be related to the observed mutational landscape differences between two races. Finally, we show that mutated genes associated with these mutational differences differ between the two populations. Thus, at the molecular level, race appears to be a factor in the progression of laryngeal cancer with ancestral genomic signatures best explaining these differences.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Laryngeal cancer afflicts approximately 12,000 new individuals in the United States each year [1,2] with different incidence and survival rates across ethnic groups [1]. This particular cancer type affects more African-American (Afr-Amr) individuals than European-Americans (Eur-Amr) [1] and the five year survival rate for Afr-Amr with laryngeal cancer is consistently lower than that for Eur-Amr [1]. While socioeconomic factors and life styles are associated with the higher incidence and lower survival rates among Afr-Amr [3], we have shown that the contribution of an individual's genetics cannot be ignored [4].

The major risk factors for laryngeal cancer are tobacco smoke and alcohol consumption [5,6]. Pro-carcinogens found in tobacco smoke are absorbed by cells, metabolized to form active carcinogens, and subsequently excreted from the body following detoxification [7]. If the active carcinogens are not excreted from the cell, the carcinogenic compounds may bind to and ultimately damage DNA [7]. The effect of alcohol with tobacco is synergistic; it is hypothesized that alcohol accelerates the absorption and action of tobacco-based carcinogens [8]. Defects in the enzyme activity or metabolic pathway of tobacco metabolism may lead to the accumulation of tobacco carcinogens in the body and increase the risk of tumor progression. Higher levels of nicotine and cotinine (the major nicotine-based metabolite that contributes to cancer development) have been reported in Afr-Amr compared to individuals of European descent, irrespective of smoking levels [9–12]. In addition, reduced metabolic clearance of nicotine to cotinine and decreased excretions of nicotine and cotinine have been observed in Afr-Amr, relative to Caucasians, for similar cigarette consumption [11,12]. Genetic studies have identified gene variants associated with reduced rates of nicotine metabolism in populations with significant African descent [13–15]. African-ancestry related genetic variants associated with susceptibility to cancer chemotherapeutic agents have also been demonstrated [16]. In addition, genetic variants associated with increased risk for head and neck cancers in patients of African descent have also been revealed by meta-analysis [17]. These evidences suggest the possible role of genetic ancestry, together with other non-genetic factors, in increased laryngeal cancer risk and poor survival rate among Afr-Amr. Nevertheless, genome-wide analysis to address the disparity issues in laryngeal cancer has
not been conducted and genome level analyses are warranted to understand the molecular basis of cancer disparity.

The recent advancements in sequencing technologies has enabled researchers to analyze the whole genome/exomes of tumor and matched normal samples of several cancer types including laryngeal cancer [18–21]. Nonetheless, the potential baseline effect of population (racial/ethnic) level genetic variation in laryngeal cancer has not been examined. This study compares the distribution of de novo point mutations that have developed in laryngeal cancers among Afr-Amr and Eur-Amr patients to gain insight into the genetic basis of racial disparities in laryngeal cancers.

2. Materials and methods

2.1. Data source

Level-2 mutation calls based on whole exome sequence data analyses for Head and Neck Squamous Cell Carcinoma were obtained from The Cancer Genome Atlas (TCGA) (http://cancergenome.nih.gov/). These publically available mutation data were manually curated by the TCGA experts.

2.2. Laryngeal cancer data

Dataset-1: We stratified clinical patient data along with their unique IDs for laryngeal cancer from the TCGA portal based on race: Afr-Amr (n = 18) and Eur-Amr (n = 91). Of these, only 13 Afr-Amr and 57 Eur-Amr had data in the publicly available TCGA database and we included all 13 Afr-Amr and 57 Eur-Amr patients in Dataset-1.

Dataset-2: Afr-Amr patients possessed a number of common characteristics: current smoker or current reformed smoker; clinical stage II or stage IV; and all were less than 70 years of age. We matched 36 Eur-Amr with similar characteristics to the Afr-Amr patients: all 13 Afr-Amr and 36 Eur-Amr patients represent Dataset-2.

We retrieved the somatic mutations specific for each individual from TCGA data using custom perl/shell scripts. The clinical data contained patient ID and other metadata. We used patient IDs to retrieve corresponding mutations from the TCGA dataset.

2.3. Statistical analyses

On Dataset-1 and Dataset-2, respectively, we conducted a Mann–Whitney U test to compare the differences in each potential factor such as age and pack years between Afr-Amr and Eur-Amr patients. The effects of race, age, smoking status, pack years, and the number of years smoked on the number of mutations of 13 Afr-Amr and 36 Eur-Amr patients (Dataset-2) were studied using multiple linear regression models. We sequentially removed non-significant covariates at a 5% significance level and the final fitted regression model with significant covariates against mutational load (i.e., the number of mutations) was plotted using R.

The following analyses were carried out for all two datasets. We compared the distribution of individual somatic mutations from each individual’s tumor for each population. Each mutation was classified as transitions (Ti) (substitution of purine to purine or pyrimidine to pyrimidine) and transversions (Tv) (substitution of purine to pyrimidine or vice-versa), their frequencies were estimated for each individual, and distributions were plotted for each group. Ti and Tv were further classified into all six possible mutational changes, C > T, C > A, C > G, T > A, T > G and T > C, and transitional frequencies were estimated for each individual. A Mann–Whitney U test at a 5% significance level was employed to compare the differences in the: (a) number of mutations, and (b) frequency of mutations for each mutation type between Afr-Amr and Eur-Amr patients. We used custom perl/ shell/R scripts for mutation estimates and STATA 10.0 (Stata Corp, College Station, TX) was used for the Mann–Whitney U test.

2.4. Context nucleotide signatures

We studied the context nucleotide signatures for somatic point mutations in Dataset-2 as this dataset contains matched samples for potential risk factors associated with laryngeal cancer. The development of point mutations highly depends on the localized, or contextual, neighborhood sequence that they are located in [22,23]. Context nucleotides for a mutation are adjacent nucleotides of that mutation (i.e., nucleotides that exist 3′ and 5′ adjacent to the point mutation) and studying the context nucleotide signatures may explain the genomics factor associated with observed mutational landscape differences. We used the Bioconductor package, SomaticSignatures [24] to analyze the context nucleotide signatures in Afr-Amr and Eur-Amr patients. In total, we analyzed 96 context nucleotide signatures (as there are 16 possible combinations of the four nucleotides (A, T, G, C) at the 5′ and 3′ end of each of 6 possible mutation types.). Differences in the frequency of each context nucleotide signature between these two ethnic groups were assessed statistically using a Mann–Whitney U test.

2.5. Significantly differently mutated genes

We created a list of genes mutated in one or more Afr-Amr or Eur-Amr patients (Dataset-2) and the differences in frequency of patients with mutations between Afr-Amr and Eur-Amr groups were studied using chi-square test in R. In addition, we obtained a list of 44 cancer driver genes for Head and Neck Squamous Cell Carcinoma (HNSCC) from the Broad GDAC Firehose (gdac.broadinstitute.org). The Broad Institute has identified these cancer driver genes using HNSCC dataset of TCGA. We analyzed the frequency of patients with mutations in these driver genes in Afr-Amr and Eur-Amr groups and the differences between the two populations were examined by chi-square test for homogeneity with continuity correction in R.

3. Results

3.1. Laryngeal cancer samples

Summary statistics of age and pack years for Afr-Amr and Eur-Amr patients for each dataset are given in Table 1. The age and number of pack years were not significantly different between Afr-Amr and Eur-Amr patients in Dataset-1 and Dataset-2. Other clinical characteristics of patients in Dataset-2 are given in Supplementary Table 1.

3.2. Dataset-1: somatic point mutations and mutational landscapes

The distributions of the number of somatic point mutations per sample for Afr-Amr and Eur-Amr were different (Fig. 1A). Specifically, Eur-Amr possessed more point mutations compared to Afr-Amr. The number of mutations ranged from 46 to 1026 with a mean of 277.63 and a median of 186 for Eur-Amr whereas the number of mutations varied from 29 to 313 with a mean of 151.31 and a median of 150 for Afr-Amr patients. At a significance level of 5%, the medians of the number of mutations in Afr-Amr and Eur-Amr were not significantly different (Table 1; P = 0.063).

We classified the somatic point mutations into transitions (Ti) and transversions (Tv) and the medians of Ti and Tv frequencies from Afr-Amr and Eur-Amr patients were found to be significantly different between these two racial groups (Fig. 1C; P = 0.0454). In particular, Afr-Amr patients had a higher proportion of Ti (median = 53.11; Q1 = 47.18; Q3 = 67.01) than Tv (median = 46.89; Q1 = 32.99; Q3 = 52.82). In contrast, Eur-Amr patients had higher Tv proportions (median = 50.6; Q1 = 42.52; Q3 = 57.92) compared to Ti (median = 49.4; Q1 = 42.08; Q3 = 57.48).
The somatic mutations were further categorized into each of the six possible nucleotide changes. The tumors from Afr-Amr and Eur-Amr patients have distinct mutational landscapes (Fig. 1E). Of these six possible mutations, only the frequency of somatic mutation type C → G was found to vary significantly between these two populations (Fig. 1G and Table 1; P = 0.01).

3.3. Dataset-2: somatic point mutations and mutational landscapes

Eur-Amr patients harbored a higher number of point mutations in the tumor (median = 213; Q1 = 129; Q3 = 383) compared to Afr-Amr patients (median = 150; Q1 = 117; Q3 = 177), as we observed in Dataset-1. At a significance level of 5%, the numbers of mutations from Afr-Amr and Eur-Amr patients were not significantly different (Table 1; P = 0.0557). The distribution of the number of mutations per sample for each patient group is shown in Fig. 1B. To study the relative effects of race, age, smoking status, the number of years of smoking and pack years on mutation load, we applied multiple linear regressions. Smoking status, the number of years of smoking and pack years were discarded from the regression model at a 5% significance level. After removing such non-significant covariates, the effects of race and age on the number of mutations remained significant. The final fitted model and the ANOVA table are provided in Supplementary Table 2. As age increases, there is significantly more mutational burden in each population (P = 0.026) and the increasing rates, that is, the effects of age, in Afr-Amr and Eur-Amr are the same (Fig. 2; Supplementary Table 2). The number of mutations in Eur-Amr was larger on average (by 151.91 mutations) than in Afr-Amr when ages were matched (Fig. 2; Supplementary Table 2).

When somatic point mutations were classified as transitions (Ti) and transversions (Tv), we found significantly different distributions of Ti and Tv frequencies between Afr-Amr and Eur-Amr patients (Fig. 1D; P = 0.0425). Afr-Amr patients had a higher proportion of Ti (median = 53.11; Q1 = 47.18; Q3 = 67.01) than Tv (median = 46.89; Q1 = 32.99; Q3 = 52.82). Eur-Amr patients had higher Tv proportions (median = 52.99; Q1 = 42.64; Q3 = 58.88) compared to Ti (median = 47.01; Q1 = 41.12; Q3 = 57.36).

The somatic mutations were further categorized into each of the six possible nucleotide changes. Fig. 1F shows that the tumors from Afr-Amr and Eur-Amr patients have distinct mutational landscapes and

<table>
<thead>
<tr>
<th>Set</th>
<th>Samples matched for</th>
<th>Age</th>
<th>Pack years</th>
<th># Mutations</th>
<th>C → A (%)</th>
<th>C → G (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>None</td>
<td>54.7</td>
<td>34.0</td>
<td>150</td>
<td>16.5</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td></td>
<td>60.9</td>
<td>40.0</td>
<td>60.0</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>57 EA</td>
<td></td>
<td>63.8</td>
<td>60.0</td>
<td>80.0</td>
<td>0.162</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>Smoking status,</td>
<td>54.7</td>
<td>34.0</td>
<td>150</td>
<td>16.5</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>Stage III/IV,</td>
<td>60.9</td>
<td>40.0</td>
<td>60.0</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>36 EA</td>
<td>Age below 70</td>
<td>63.8</td>
<td>60.0</td>
<td>80.0</td>
<td>0.037</td>
</tr>
</tbody>
</table>

Q1: first quartile; Q3: third quartile; M: median; P-val: P value.

* At a significance level of 5%, the proportions were significantly different between AA and EA.
Fig. 1H shows different distributions of frequencies by mutation type. The frequencies of somatic mutation types, C > A and C > G, were found to vary significantly between these two populations (Table 1; P = 0.0372 for C > A; P = 0.0297 for C > G). Tumors from Afr-Amr patients harbor lower C > A (median = 16.49; Q1 = 13.79; Q3 = 26.50) and C > G (median = 11.28; Q1 = 8.25; Q3 = 16.51) somatic mutations compared to Eur-Amr patients (for C > A; median = 23.79; Q1 = 18.21; Q3 = 32.90; for C > G; median = 14.68; Q1 = 12.66; Q3 = 17.72).

3.4. Context nucleotide signatures

The development of somatic mutations depends on the contextual neighborhood of the mutation loci [22,23]. We investigated whether Afr-Amr patients in Dataset-2 possess different context nucleotide signatures at each somatic mutation loci compared to Eur-Amr patients in Dataset-2, thus favoring different mutational landscapes. Fig. 3 shows different context nucleotide patterns for Afr-Amr and Eur-Amr patients with P-values < 0.05. Some context nucleotide frequencies significantly differed between Afr-Amr and Eur-Amr patients. Mutation types C > G and C > A whose frequency distributions are different have four and two, respectively, significantly different context nucleotide patterns between Afr-Amr and Eur-Amr. Among mutation types, C > G showed the most different patterns. Mutation T > A does not have any significantly different patterns and others have 1 or 2 different patterns.

3.5. Significantly differently mutated genes

We obtained a list of genes mutated in one or more Afr-Amr or Eur-Amr patients (Dataset-2). We found 576 genes were mutated in both Afr-Amr and Eur-Amr patients while 37 genes and 1731 genes were mutated only in Afr-Amr and Eur-Amr patients, respectively. Thus, we analyzed a list of 2344 (576 + 37 + 1731) genes mutated in one or more Afr-Amr or Eur-Amr patients. Of these 2344 genes, six genes, RUNX1T1, TTN, NAV3, PIK3CA, KIAA1033, and ZMYM6, were significantly differently mutated between Afr-Amr and Eur-Amr patients (Table 2). We also obtained a list of 44 driver genes identified in HNSCC and compared the frequencies of patients with mutations in these driver genes in our dataset. The results are shown in Fig. 4. Several driver genes were mutated in different frequencies in Afr-Amr as compared to Eur-Amr. The Eur-Amr patients had mutations in 29 of 44 driver genes while Afr-Amr had mutations in 20 of 44 driver genes. Importantly, a driver gene, PIK3CA was significantly differently mutated between Afr-Amr and Eur-Amr (Table 2).

4. Discussion

Disparities have been reported in the incidence rate of laryngeal cancer between Afr-Amr and Eur-Amr [1]. However, our knowledge of the differences of laryngeal cancer at the genetic level between populations is limited. The goal of this study is to understand differences in the mutational landscape of tumors between Afr-Amr and Eur-Amr laryngeal cancer patients. The sample size, its large sample space (i.e., whole exomes), and our risk-matched comparison provided us with sufficient power to identify inherent and significant differences between the cancer genomes of Afr-Amr and Eur-Amr patients.

4.1. Effect of race on mutational landscapes

In this study, we grouped and analyzed two sequentially-filtered datasets. While the first dataset was generated to identify overall mutational landscape differences between Afr-Amr and Eur-Amr patients, Dataset-2 was matched for smoking status, and clinical stage III/IV to test the effect of race after adjusting for smoking status and late stage cancer. If the mutational burdens are independent of race, we expect to see more or less similar mutational landscapes for Afr-Amr and Eur-Amr in our two datasets. However, we observed different mutational landscapes between Afr-Amr and Eur-Amr patients. The results from these two datasets are congruent for most of the observations (Fig. 1 and Table 1) and reveal that ancestral origin is an important factor associated with mutation burdens. In each of the analyses, the number of somatic mutations and the frequencies of Ti and Tv were dissimilar between Afr-Amr and Eur-Amr patients. While the differences of the number of mutations were close to significant or not significant, Ti and Tv frequencies varied significantly (P-value < 0.05) between Afr-Amr and Eur-Amr patients in all datasets. In addition, the somatic mutational landscapes of Afr-Amr patients were different from that of Eur-Amr patients.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Afr-Amr (N = 13)</th>
<th>Eur-Amr (N = 36)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUNX1T1</td>
<td>5 (38.46%)</td>
<td>3 (8.33%)</td>
<td>0.037</td>
</tr>
<tr>
<td>TTN</td>
<td>5 (38.46%)</td>
<td>30 (83.33%)</td>
<td>0.007</td>
</tr>
<tr>
<td>NAV3</td>
<td>0 (0%)</td>
<td>13 (36.11%)</td>
<td>0.031</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>0 (0%)</td>
<td>12 (33.33%)</td>
<td>0.043</td>
</tr>
<tr>
<td>KIAA1033</td>
<td>3 (23.08%)</td>
<td>0 (0%)</td>
<td>0.021</td>
</tr>
<tr>
<td>ZMYM6</td>
<td>3 (23.08%)</td>
<td>0 (0%)</td>
<td>0.021</td>
</tr>
</tbody>
</table>

* Indicates known cancer driver gene from HNSCC.

Table 2 Genes significantly differing in frequency between African-American (Afr-Amr) and European-American (Eur-Amr) laryngeal cancer patients.

Fig. 3. Fraction of context nucleotide signatures for all six mutation types based on TCGA patient data from African-American and European-American populations.
We also studied the effect of race, age, smoking status, number of pack years, and number of years smoked on mutational load on Dataset-2 using multiple linear regression models. We sequentially removed non-significant factors from regression models. Smoking status, the number of years smoked and pack years did not have significant effects on mutation burden. This result also supports that we removed the effect of known risk factors of laryngeal cancers such as smoking levels and studied the effect of race on mutational landscape using Dataset-2. After removing all non-significant factors, race and age showed significant effects on mutation load (P = 0.02 for race, and P = 0.049 for age). Thus analyses suggest that race and age are significant factors related to the number of mutations. As we found from other analyses in different datasets, Eur-Amr patients carry about 151 more mutations on average than Afr-Amr while controlling for the age of patients. Our results support race as a potential factor associated with mutational burdens.

Our Dataset-2 provided stronger power to analyze the mutational landscapes. Dataset-2 showed a significant differences in the frequency of C > A and C > G mutation types between Afr-Amr and Eur-Amr patients. However, Dataset-1 did not show a significant difference in the frequency of C > A mutations between Afr-Amr and Eur-Amr patients. The C > A mutations are related to tobacco smoke [25,26]. In Dataset-1, we did not match the patients for smoking status with Afr-Amr and Eur-Amr patients possessing broad ranges of smoking histories. In fact, the Eur-Amr cohort in Dataset-1 included current smokers (49.1%), current reformed smokers for ≤ 15 years (34.5%), current reformed smokers for > 15 years (12.73%), and life-long non-smokers (3.6%) whereas the Afr-Amr cohort included current smokers (61.5%), current reformed smokers for ≤ 15 years (30.8%), and current reformed smokers for > 15 years (7.8%). These differences in smoking history between Afr-Amr and Eur-Amr patients in Dataset-1 may have obscured the significance of race on C > A mutations. Nevertheless, when the patients were matched for smoking status in Dataset-2, we found a significant difference in the frequency of C > A mutations between the two populations.

While socio-economic, behavior, and lifestyle factors are known to be associated with the higher incidence and lower survival rates of laryngeal cancer in Afr-Amr [3], host factors may also play an important role. We attempted to study the effect of ancestry on mutational burden in laryngeal cancer in our study and findings suggest that genetic ancestry could also be associated with the cancer disparity. Each genome-wide sample contained, on average, hundreds of new
mutations providing ample power for our comparisons. In each of the datasets, we observed a higher number of somatic point mutations in Eur-Amr as compared to Afr-Amr patients, despite the fact that Afr-Amr and Eur-Amr patients have a similar smoking history in Dataset-2. In contrast, earlier studies showed a direct correlation between number of mutations and smoking rate in tobacco-related cancers [27, 28]. If smoking is the only major factor for mutation burden, we expect to see a similar number of mutations in Afr-Amr and Eur-Amr patients in Dataset-2, however, we found a lower mutational load in Afr-Amr patients, adding more evidence for the effect of race on mutations. Since this study only investigates the mutational landscape differences of Afr-Amr vs Eur-Amr patients, we could not study the effect of mutations or the functional impact of these mutations on these populations. These and other molecular mechanisms explaining the lower number of mutations in Afr-Amr patients need to be further explored.

To delve deeper into the reason for the disproportionate mutational landscapes between these two populations, we further analyzed the context nucleotide signatures (+/− 1 bp) of each mutation type in Afr-Amr and Eur-Amr patients (Dataset-2). Our analyses reveal that the Afr-Amr and Eur-Amr populations have different frequencies of context nucleotide signatures [Fig. 3]. Thus, ancestral genomic signatures may play a key role in the observed mutational landscape differences. In addition to these signatures, other genomic factors also might have influenced the mutational landscapes. For instance, population-specific driver mutations and other mutational burdens in the tumor cell could have also contributed to the different mutational patterns observed between Afr-Amr and Eur-Amr patients. However, at present, it is not possible for us to study the influence of driver mutations and mutational burdens on the mutational landscape of these two populations since the raw data are not yet available.

4.2. Differentially mutated genes

This study also shows differences in the distribution of mutated genes between Afr-Amr and Eur-Amr patients (Dataset-2). Several of 44 known HNSCC cancer driver genes were mutated in different frequencies between Afr-Amr and Eur-Amr patients. Eur-Amr patients had mutations in 29 of 44 driver genes, whereas Afr-Amr had mutations in 20 of 44 driver genes. Importantly, Afr-Amr did not have any mutation in a driver gene, PTK3CA whereas many Eur-Amr had mutations in this gene and the difference is significant (P = 0.043). The analyses of 2344 genes reveal Afr-Amr had mutations in significantly limited number of genes compared to Eur-Amr (613 vs 2307; chi-square test P < 2.2e−16). While none of the Afr-Amr patients had mutations in PTK3CA, a higher frequency of Afr-Amr patients had mutations in RUNX1T1, KIAA1033, and ZMYM6 as compared to Eur-Amr (P < 0.05) (Table 2). The gene, RUNX1T1 is known to be a key gene associated with acute myeloid leukemia [29]. In addition, KIAA1033, and ZMYM6 are important for cellular function and cell morphology. Thus, mutations in these genes may be associated with HNSCC progression or survival and needs to be studied further. However, RUNX1T1, KIAA1033, and ZMYM6 were not identified as cancer driver genes in HNSCC. A recent study based on data from 279 TCGA HNSCC patients also showed PTK3CA as a potential driver gene but the genes, RUNX1T1, KIAA1033, and ZMYM6 were not identified as driver genes [21]. The limited inclusion of Afr-Amr patients might have severely biased cancer gene discovery in TCGA cohort and the analyses of more black patients could provide a different list of genes (the TCGA HNSCC cohort currently has ~390 Eur-Amr and ~40 Afr-Amr patients). These observations suggest that Afr-Amr may have a different set of frequently-mutated genes that could be associated with a higher risk for laryngeal cancer. Either more Afr-Amr samples or population-specific analyses are needed to identify additional candidate genes from this population.

5. Conclusion

This study demonstrates the power of a genomics approach to study the etiology of health disparities by analyzing, for the first time, the effect of race on mutational signatures in laryngeal cancer. If laryngeal cancer has a common genetic basis in all populations, we should find a similar mutational burden in Afr-Amr and Eur-Amr patients. Our results reveal different mutation loads between Afr-Amr and Eur-Amr patients. Afr-Amr patients may experience different genetic burdens in addition to other difficulties arising from non-biological factors which make them more prone to HNSCC. Thus, our results support a molecular disparity in laryngeal cancer between Eur-Amr and Afr-Amr as suggested by our mutational landscape analyses. Our study also suggests that genetic background may play an important role in mutational development and cancer progression and highlights the need for more data and specific analyses to identify the driver mutations in Afr-Amr patients. In addition, if we could compare the effects of driver mutations on laryngeal cancer risk and outcome, we may better address why Afr-Amr are at a higher risk for laryngeal cancer than Eur-Amr. More research into the effect size of the genomic level variations found in the laryngeal cancers of Afr-Amr and Eur-Amr as well as genome–environment interactions will further help us to understand the associated baseline risk disparity observed between these two populations.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

Authors thank Dr. Bhawna Dubey and Dr. Sudhir Kumar from, respectively, Fox Chase Cancer Center and Temple University, for their valuable comments on data analyses and the manuscript. We thank Dr. Kara Maxwell, University of Pennsylvania for critical review of the MS. We also thank Dr. J. Robert Beck from Fox Chase Cancer Center for supporting this study. We gratefully acknowledge the efforts of TCGA network for making the data available for research use. This work was supported in part by grants RSG-14-033-01-CPB from the American Cancer Society and CA006927-50 from the National Cancer Institute. We also acknowledge the Russian Government Program of competitive growth of Kazan Federal University for the partial funding support to IS. The results shown here are in whole or part based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ygeno.2015.12.004.

References


