



# The Impact of CA9 Gene Polymorphism (*RS2071676*) on Oral Cancer Susceptibility - An *In-Silico* Study

Suvarna K<sup>1</sup>, Vijayashree Priyadharsini J<sup>2</sup>, Gheena S<sup>3</sup>, Pratibha Ramani<sup>4</sup>, Abilasha R<sup>3</sup>, Hannah R<sup>5</sup>

<sup>1</sup>Department of Oral Pathology, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai-600077, India; <sup>2</sup>Assistant Professor, Department of Microbiology, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai-600077, India; <sup>3</sup>Reader, Department of Oral Pathology, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai-600077, India; <sup>4</sup>Professor and Head, Department of Oral Pathology, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai-600077, India; <sup>5</sup>Senior Lecturer, Department of Oral Pathology, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai-600077, India.

## ABSTRACT

**Background:** Oral cancer accounts for 2 to 4% of all cancer and 90% of oral cancer cases are squamous cell carcinoma. Apart from etiological factors genetic factors also play an important role in oncogenesis. The common type of sequence variations is single nucleotide polymorphisms (SNPs), described as a change in the coding and amino acids sequence in the related proteins which could confer protective or lethal effects on the organism.

**Objective:** To analyze the genetic variation in the CA9 gene (*rs2071676*) and to compare allele frequencies in different populations worldwide. The *rs2071676* variant of the CA9 gene is a missense variant that results in the substitution of the amino acid methionine in place of valine.

**Methods:** Genotype frequencies of the SNP *rs2071676* were collected from the Ensembl database for different populations and the deviations were analyzed. Furthermore, the expression profile of the CA9 gene in HNSCC was assessed using *in-silico* tools. The survival of patients based on the expression of the CA9 gene was also assessed.

**Results:** The present study identified deviations in allele frequencies for *rs2071676* polymorphism between different populations. The minor allele frequency in the ancestral population or the African population was found to be much lower than the other populations such as American, East Asian, and South Asian, where both the alleles showed similar frequencies.

**Conclusion:** The significant deviation between allele frequencies in different groups provides clues about the positive selection of these alleles in certain populations. This led us to further investigate the plausibility of association of the CA9 gene with HNSCC. The intriguing facts observed from the study could aid in revealing the association between the CA9 gene and the progression of oral cancer.

**Key Words:** CA9 gene, Polymorphism, SNP, Oral cancer, Genetic variation, In silico study

## INTRODUCTION

Oral cancer accounts for 2 to 4% of all cancer cases in the world. 90% of oral cancer cases are squamous cell carcinoma (OSCC).<sup>1-3</sup> Globocan in 2012 stated that the worldwide oral cancer incidence is 2.1%.<sup>4</sup> In India, the prevalence of oral cancer is high and it accounts for 45% of all the cancers reported in India.<sup>1,5,6</sup> In Asian countries, the major etiological factors for developing OSCC include tobacco, areca nut

chewing, and smoking.<sup>7,8</sup> Genotoxic substances present in tobacco and areca nuts such as nitrosamines, polycyclic aromatic hydrocarbons, volatile aldehydes, and hydroquinone are converted to carcinogens. Nicotine derived nitrosamine, n-nitrosonornicotine, NNK, areca nut alkaloids, arecoline, arecaine can result in the formation of oral submucous fibrosis, oral precancerous lesions, and oral cancer.<sup>7,9-11</sup> Apart from the etiological factors, genetic factors also play a major

### Corresponding Author:

**Suvarna K**, Department of Oral Pathology, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai-600077, India; Email: [151907003.sdc@saveetha.com](mailto:151907003.sdc@saveetha.com)

ISSN: 2231-2196 (Print)

ISSN: 0975-5241 (Online)

Received: 10.08.2020

Revised: 05.09.2020

Accepted: 08.10.2020

Published: 26.12.2020

role in oncogenesis.<sup>12-15</sup> It is well known that the tumor cell during its course of development undergoes molecular alterations in several cellular molecules including DNA, RNA, and proteins. This could be attributed to the inherent biological properties of the cancer cell.<sup>16,17</sup> The up-regulation and down-regulation of certain genes were found to be involved in extracellular matrix degradation (ECM), epithelial to mesenchymal transition (EMT), and detoxification pathways.<sup>18,19</sup>

The most common type of DNA sequence variation is single nucleotide polymorphisms. This results in errors in the coding sequence, which leads to alterations in the amino acid being coded or incorporated into the protein chain.<sup>12</sup> SNP directly or indirectly affect the gene expression pattern and functions.<sup>12</sup> SNPs at intronic regions of genes may result in three-dimensional changes in the DNA molecules which result in Gibbs free energy change. The possible consequence of it may impact DNA polymerases enzyme activity and transcription factor binding. SNPs can be present as homozygous and heterozygous. An ancestral allele is referred to as the wild type allele and the altered allele is referred to as SNP or variant allele.<sup>7</sup>

*CA9* is a gene located on chromosome 9p13.3.<sup>12</sup> *CA9* gene codes for carbonic anhydrase 9 which is a glycoprotein and a member of a family of metalloenzymes which can catalyze the reversible hydration of carbon dioxide into carbonic acid.<sup>20</sup> *CA9* glycoprotein plays an important role in the stabilization of extracellular pH hydration of carbon dioxide.<sup>20</sup> Solid tumor have a common feature called cellular hyperoxia.<sup>21</sup> Tumor hypoxia causes tumor cells to undergo adaptive changes that can enable them to survive and proliferate.<sup>22</sup> Increased *CA9* synthesis is regulated through the binding of hypoxia-inducible factor (HIF) to a hypoxia response element (HRE) within its basal promoter. This contributes to the neutralization of intracellular pH and acidification of extracellular microenvironment which promotes cell proliferation, adhesion, and invasion.<sup>23</sup> *CA9* expression is highly increased in various cancers and it has been considered as an endogenous marker for tumor hypoxia.<sup>24,25</sup> Several studies have reported various polymorphisms of the *CA9* gene in renal cell carcinoma.<sup>26</sup> The increased expression of the *CA9* gene is also seen in several cases of oral squamous cell carcinoma. Haplotypes of *rs2071676*, *rs3829078*, and *rs1048638* when combined have potential in predicting significance in oral carcinogenesis and the gene-environment interaction of *CA9* polymorphism.<sup>12</sup> This study was used to determine the consequence of a genetic variation (*rs2071676*) in the *CA9* gene and their putative association with OSCC.

## MATERIALS AND METHODS

One of the exon variants of the *CA9* gene was selected for the study based on the literature mining process. Since oral cancer is a disease that is more predominant in patients with

a chronic history of using smoking and smokeless tobacco such as pan, gutka, etc., genetic variants that were more closely related to environmental factors were identified in the gene and selected. A recent study by Chien et al., 2012, demonstrated that patients who are heterozygous for *rs2071676* polymorphism had a higher risk of developing lymph-node metastasis when compared to patients with homozygous AA allele.<sup>12</sup> The Ensembl database was used to acquire the frequency data of the polymorphism variant in different populations ([https://asia.ensembl.org/Homo\\_sapiens/Variation/Population?db=core;r=20:56387216-56388216;v=rs2064863;vdb=variation;vf=52924966](https://asia.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=20:56387216-56388216;v=rs2064863;vdb=variation;vf=52924966)).<sup>27</sup> IMutant suit version 2.0 was used to identify the stability of the protein.<sup>28</sup> PROVEAN and MutPred were the tools used to predict the pathogenicity of the variant.<sup>29</sup> Furthermore, the expression of the *CA9* gene in HNSCC was analyzed using the UALCAN (<http://ualcan.path.uab.edu/cgi-bin/TCGA-survival1.pl?genenam=CA9&ctype=HNSC>) database.<sup>30</sup> Survival curve analysis based on the tumor grade and expression profile was performed to demonstrate the putative role of the *CA9* gene with HNSCC. Furthermore, the expression of the *CA9* gene in HNSc was analyzed using the UALCAN database. Expression in different grades of tumors were assessed viz., grade 1- well-differentiated, grade 2- moderately differentiated, grade 3- poorly differentiated, grade 4- undifferentiated. Transcript per million (TPM) is a normalization method for RNA-seq data. The TPM values used for the generation of box whisker plots were also used to determine the significant difference between the groups. The t-test was performed using a PERL script with the Comprehensive Perl Archive Network (CPAN) module. Combined survival effect analysis of gene expression and other clinical parameters such as race, gender, tumor grade, cancer subtypes were assessed using multivariate Kaplan-Meier survival analysis.

## RESULTS AND DISCUSSION

The *CA9* gene can be considered as a potentially important biomarker for the evaluation of tissues. The *CA9* gene codes for carbonic anhydrase-IX protein. This protein is a tumor-associated cell-surface glycoprotein that is induced during hypoxia. This protein is also involved in adaptation to acidosis and implicates in cancer progression via its catalytic activity and/or non-catalytic functions.<sup>31</sup> *CA9* gene expression in non-cancerous tissues is rare and confined to epithelia of the stomach, gallbladder, pancreas, and intestine.<sup>30</sup> *CA9* expression is significantly up-regulated in several cancers like lung, colon, colorectal, gastric, pancreatic, breast, cervix, bladder, ovaries, brain, head and neck, and oral cancers.<sup>31,32</sup>

Lin et al., identified polymorphic variants in the *CA9* gene which was a predictor of lymph node metastasis of prostate cancer. Their study was the first to demonstrate the correlation between *CA9* gene polymorphism and expres-

sion profile which could enable early detection of low-risk prostate cancer.<sup>34</sup> Brockton et al., identified that aberrant expression of *CA9* gene in OSCC patients correlated with nodal metastasis and poor prognosis.<sup>35</sup> Intracellular alkalosis and extracellular acidosis induced by *CA9* allows tumor cells to survive in hypoxic conditions.<sup>34,35,36</sup> In addition to it, acidification of the extracellular environment and activation of extracellular proteases might contribute to the disruption of cell adhesion structures leading to tissue invasion and metastasis. *CA9* expression was also found to be indicative of aggressive and resistant tumors.<sup>37</sup>

In this study, the global allelic frequency of *rs2071676* polymorphism in the *CA9* gene was found to be 65% for G and 35% for A allele (Figure 1). The global distribution among different populations like African, American, East Asian, European, and South Asian were also included (Figure 2). To derive a more vivid picture about the deviations in allele frequencies for the selected polymorphism, the allele frequency was computed from the data available from the reports of Chien et al, 2012.<sup>12</sup> Since the population described in the report was of Taiwanese population, the East Asian data from the database was selected for comparison and is presented in figures 3 and 4. The results showed a slight deviation in the altered and reference allele frequencies. But when the same data was compared with the ancestral population i.e., the African population [G allele - 92%; A allele - 8%] the allele frequency was found to deviate dramatically. This result showed that the A allele was positively selected in certain populations. Generally, an allele tends to be positively selected when it is beneficial to the individual or a group. Interestingly, we observed that the variant allele frequency to be markedly elevated. The reason for the same could provide clues on the putative association of the *CA9* gene with the progression of HNSCC.

The haplotype of *rs2071676*, *rs3829078*, *rs1048638* combined have a potential predictive value in oral carcinogenesis.<sup>12</sup> Although a single variation in the gene might not always affect the gene expression profile, the cumulative effect of several genes might influence the process of gene expression. The placement of the variant in the gene and the type of alteration will determine the consequences of the variant. The lymph-node metastasis tendencies in oral cancer were highly in heterozygous GA allele combination rather than AA allele combination.<sup>12</sup> The missense mutation of the *CA9* gene can cause altered *CA9* gene expression and protein production. The prognostic value of *CA9* expression has been proven for other cancers such as renal clear cell carcinoma, cervical, colorectal, esophageal, lung, and breast cancer.<sup>12,32,33</sup> These various findings suggest that mutations in the *CA9* gene play an important role in all OSCC progression.

In the present study, the I-Mutant analysis (DDG value: -1.02) revealed that the variation from valine to methionine decreased the stability of the protein (data not shown). Furthermore, the pathogenicity analysis of the resultant protein after substitution with variant amino acid was assessed using Mutpred and PROVEAN tools which returned scores that predicted it to be neutral or non-pathogenic (data not shown). Despite the contrasting results, experimental evidence emphasizing the functional role of the *CA9* gene would aid in deriving the association between the candidate gene and the disease phenotype. Gene expression analysis revealed an increase in the transcript levels with increasing grades of tumor which differed significantly in most of the groups (Figure 5). In addition, the Kaplan-Meier method was used to assess the survival period of head and neck squamous cell carcinoma (HNSCC) patients based on the expression of the selected gene. The results showed a marginal significance on the survival profile of patients with high and medium/low-level expression of *CA9* gene [p value = 0.073]. High expression of *CA9* gene in grade 2 HNSCC patients correlated with poor prognosis compared to grade 1 patient with HNSCC. In other words, there was a sharp decline in the survival of patients upon down-regulation of the *CA9* gene in HNSCC patients (Figure 6a). Furthermore, we observed that the low/medium level expression among African-American and Caucasian population exerted a significant effect on the survival of patients. African-American patients with a low/medium expression of the *CA9* gene exhibited low survival probability when compared to Caucasian patients (Figure 6b).

The cumulative effect of single nucleotide variants might play a vital role in the development and progression of a disease. Since cancer is a multifactorial disease precipitated by the interaction of several candidate genes, a thorough knowledge of the consequences of individual variants and their collective effect would aid in deducing the possible mechanisms underlying the phenotype analyzed. The association of genetic variants of the *CA9* gene with HNSCC is worth studying since it might play a direct role in tumor metastasis and resistance. Identifying significant genetic markers will aid in creating an SNP panel which can be correlated with progressing grades of tumor. Various etiological factors like consumption of tobacco and its products can also induce the aberrant expression of the *CA9* gene via the expression of nuclear hypoxia-inducible factor.<sup>12</sup> Consumption of tobacco is one of the main etiological factors of oral cancers in the South Asian population. Hence a detailed study to analyze the *CA9* gene expression on oral cancer susceptibility in the South Asian population needs to be done to get a more comprehensive picture. Computational analysis has been used to identify potential targets related to the cancer phenotype.<sup>38,39</sup> Hence such studies would aid in solving population-based queries related to polymorphic markers associated with cancer.

## CONCLUSION

In conclusion, the present study is the first of its kind which was designed to assess the frequency of *rs2071676* polymorphism in different populations and reveal the deviation which could be associated with positive selection of an allele in specific populations. The intriguing facts observed from the study could aid in revealing the association between the *CA9* gene and the progression of oral cancer. Further experimental validation is warranted to confirm the selection of minor alleles in the south Asian population. Effects of the various etiological factors of oral cancers like consumption of tobacco products resulting in hypoxia and aberrant expression of *CA9* gene can be better studied in the South Asian population. Genotype analysis of *rs2071676* polymorphism of *CA9* gene in the South Asian population should be performed to identify the positive selection of A allele and its prevalence. However, more experimental evidence should be accumulated to identify the association between *rs2071676* polymorphism of *CA9* gene and oral cancer susceptibility in the South Asian population.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the help and support rendered by the department of Dental research cell and the department of Oral pathology of Saveetha Dental College for their constant assistance with the research.

## Conflict of Interest

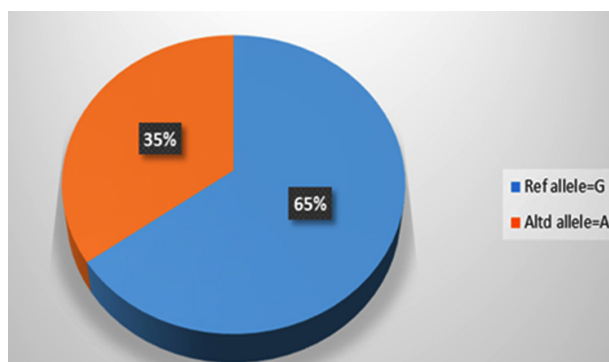
The authors have no potential conflict of interest.

## Source of Funding: Nil

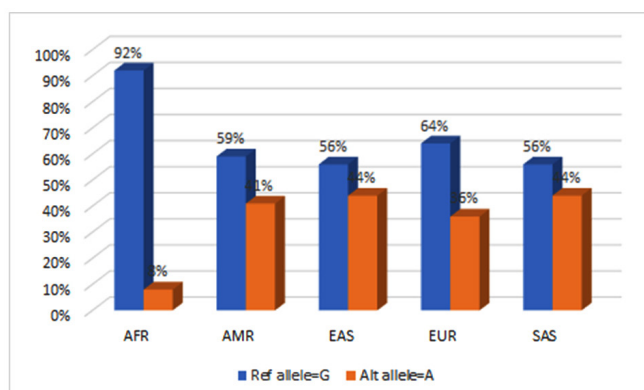
## REFERENCES

- Markopoulos AK. Current Aspects on Oral Squamous Cell Carcinoma. *Open Dentis J* 2012;10:126–30.
- Shree KH, Ramani P, Sherlin H, Sukumaran G. Saliva as a diagnostic tool in oral squamous cell carcinoma—a systematic review with Meta analysis. *Pathol Oncol* 2019;25(2):447–453.
- Viveka TS, Shyamsundar V, Krishnamurthy A, Ramani P, Ramshankar V. p53 Expression Helps Identify High Risk Oral Tongue Premalignant Lesions and Correlates with Patterns of Invasive Tumor Front and Tumor Depth in Oral Tongue Squamous Cell Carcinoma Cases. *Asian Pac J Cancer Prev* 2016;17:189–95.
- Gupta N, Gupta R, Acharya AK, Patthi B, Goud V, Reddy S, et al. Changing Trends in oral cancer – a global scenario. *Nepal J Epidemiol* 2017;6:613–9.
- Sivaramakrishnan MS, Ramani P. Study on the Prevalence of Eruption Status of Third Molars in South Indian Population. *Biol Med (Aligarh)* 2015;7: 245.
- Jayaraj G, Ramani P, Herald J, Sherlin, Premkumar P, Anuja N. Inter-observer agreement in grading oral epithelial dysplasia – A systematic review. *J Oral Maxillof Surg Med Pathol* 2015;27:112–6.
- Multania S, Saranath D. Single Nucleotide Polymorphisms and Risk of Oral Cancer: Indian Case-Control Study. *J Clin Cell Immunol* 2016;7(4): 100448.
- Hannah R, Ramani P, Sherlin HJ, Ranjith G, Ramasubramanian A, Jayaraj G, et al. Awareness about the use, ethics and scope of dental photography among undergraduate dental students dentist behind the lens. *Res J Pharma Tech* 2018;11(3):1012–6.
- Jayaraj G, Sherlin HJ, Ramani P, Premkumar P, Natesan A. Stromal myofibroblasts in oral squamous cell carcinoma and potentially malignant disorders. *Indian J Cancer* 2015; 52(1):87–92.
- Sridharan G, Ramani P, Patankar S, Vijayaraghavan R. Evaluation of salivary metabolomics in oral leukoplakia and oral squamous cell carcinoma. *J Oral Pathol Med* 2019;48(4):299–306.
- Jayaraj G, Sherlin HJ, Ramani P, Premkumar P, Anuja N. Cytomegalovirus and Mucoepidermoid carcinoma: A possible causal relationship? A pilot study. *J Oral Maxillofac Pathol*. 2015;19(3):319–24.
- Chien MH, Yang JS, Chu YH, Lin CH, Wei LH, Yang SF, et al. Impacts of *CA9* Gene Polymorphisms and Environmental Factors on Oral-Cancer Susceptibility and Clinicopathologic Characteristics in Taiwan. *PLoS ONE* 2012;7:e51051.
- Jangid K, Alexander AJ, Jayakumar ND, Varghese S, Ramani P. Ankyloglossia with cleft lip: A rare case report. *J Indian Soc Periodontol* 2015;19(6):690–3.
- Kumar A, Sherlin HJ, Ramani P, Natesan A, Premkumar P. Expression of CD 68, CD 45 and human leukocyte antigen-DR in central and peripheral giant cell granuloma, giant cell tumor of long bones, and tuberculous granuloma: An immunohistochemical study. *Indian J Dent Res* 2015;26(3):295.
- Swathy S, Gheena S, Varsha SL. Prevalence of pulp stones in patients with history of cardiac diseases. *Res J Pharma Tech* 2015;8(12):1625–8.
- Sridharan G, Ramani P, Patankar S. Serum metabolomics in oral leukoplakia and oral squamous cell carcinoma. *J Cancer Res Ther* 2017;13(3):556–61.
- Gupta V, Ramani P. Histologic and immunohistochemical evaluation of mirror image biopsies in oral squamous cell carcinoma. *J Oral Biol Craniofac Res* 2016;6(3):194–7.
- Thangaraj SV, Shyamsundar V, Krishnamurthy A, Ramani P, Ganesan K, Muthuswami M, et al. Molecular Portrait of Oral Tongue Squamous Cell Carcinoma Shown by Integrative Meta-Analysis of Expression Profiles with Validations. *PLoS One*. 2016;11(6):e0156582.
- Gheena S, Ezhilarasan D. Syringic acid triggers reactive oxygen species-mediated cytotoxicity in HepG2 cells. *Hum Exp Toxicol* 2019;38(6):694–702.
- Sedlakova O. Carbonic anhydrase IX, a hypoxia-induced catalytic component of the pH regulating machinery in tumors. *Front Physiol* 014;4:400.
- Runkel S, Vaupel P, Kallinowski F, Menke H, Bäßler KH, Fortmeyer HP, et al. Blood flow, oxygen consumption and substrate utilization of human tumors xenotransplanted into nude rats. *J Cancer Res Clin Oncol* 1986;111:S51–S51.
- Towle HC. Metabolic Regulation of Gene Transcription in Mammals. *J Biol Chem* 1995;270:23235–8.
- Supuran CT. Carbonic Anhydrase Inhibition and the Management of Hypoxic Tumors. *Metabolites* 2017;7(3):48.
- Bussink J, Johannes HA, van der Kogel AJ. Tumor hypoxia at the micro-regional level: clinical relevance and predictive value of exogenous and endogenous hypoxic cell markers. *Radiother Oncol* 2003;67:3–15.
- Ivanov S, Liao SY, Ivanova A, Danilkovitch-Miagkova A, Tarasova N, Weirich G, et al. Expression of Hypoxia-Inducible Cell-Surface Transmembrane Carbonic Anhydrases in Human Cancer. *Am J Pathol* 2001;158:905–19.

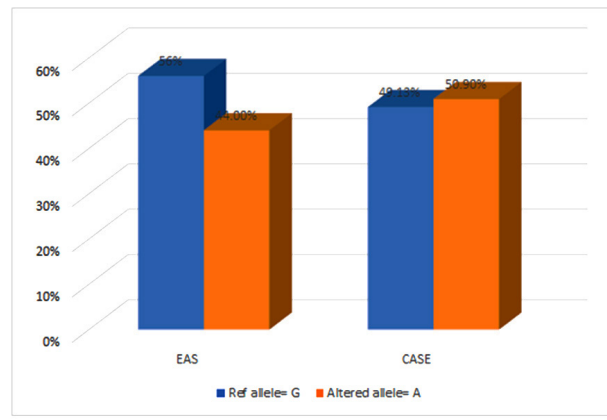
26. de Martino M, Klatte T, Seligson DB, LaRochelle J, Shuch B, et al. CA9 Gene: Single Nucleotide Polymorphism Predicts Metastatic Renal Cell Carcinoma Prognosis. *J Urol* 2009;182:728–34.
27. Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, et al. Ensembl variation resources. Database 2018; 2018:bay119.
28. Capriotti E, Fariselli P, Casadio R. I-Mutant 2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res* 2005;33:W306–10.
29. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 2015; 31(16):2745–7.
30. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* 2017;19(8):649–58.
31. Pastorekova S, Gillies RJ. The role of carbonic anhydrase IX in cancer development: links to hypoxia, acidosis, and beyond. *Cancer Metast Rev* 2019;38(1-2):65–77.
32. Bui MHT, Visapaa H, Seligson D, Kim H, Han K-R, Huang Y, et al. prognostic value of carbonic anhydrase ix and ki67 as predictors of survival for renal clear cell carcinoma. *J Urol* 2004;171:2461–6.
33. Giatromanolaki A. BNIP3 Expression Is Linked with Hypoxia-Regulated Protein Expression and with Poor Prognosis in Non-Small Cell Lung Cancer. *Clin Cancer Res* 2004;10: 5566–71.
34. Lin CY, Wang SS, Yang CK. Genetic polymorphism and carbonic anhydrase 9 expression can predict nodal metastatic prostate cancer risk in patients with prostate-specific antigen levels  $\leq 10$  ng/ml at initial biopsy. *Urol Oncol* 2019;37(11):814.e9-814.e16.
35. Brockton NT, Klimowicz AC, Bose P. High stromal carbonic anhydrase IX expression is associated with nodal metastasis and decreased survival in patients with surgically-treated oral cavity squamous cell carcinoma. *Oral Oncol* 2012;48(7):615–622.
36. Smrutipragnya S, Priyanka D, Santosh KS. Role of Chromium Enriched Tobacco in the Occurrence of Oral Carcinogenesis. *Int J Curr Res Rev* 2020;12(18): 20-24
37. Hoogsteen IJ, Marres HA, Bussink J, van der Kogel AJ, Kaanders JH. Tumor microenvironment in head and neck squamous cell carcinomas: predictive value and clinical relevance of hypoxic markers. A review. *Head Neck* 2007;29: 591–604
38. Vijayashree P, Paramasivam A. Virtual screening of mutations in antioxidant genes and its putative association with HNSCC: An *in silico* approach. *Mutat Res Fund Res Fund Mol M* 2020;821:111710.
39. Anita R, Paramasivam A, Vijayashree Priyadharsini J, et al. The m6A readers YTHDF1 and YTHDF3 aberrations associated with metastasis and predict poor prognosis in breast cancer patients. *Am J Cancer Res* 2020;10: 2546-2554.



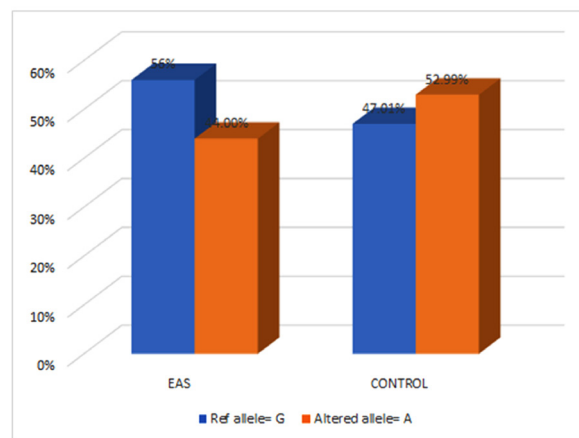
**Figure 1:** Pie graph showing Global distribution of allele frequency for *rs2071676* polymorphism of *CA9* gene [data derived from Ensembl database]. The global reference allele frequency is 65% and represented in blue color and the global altered allele frequency is 35% and represented in orange in color.



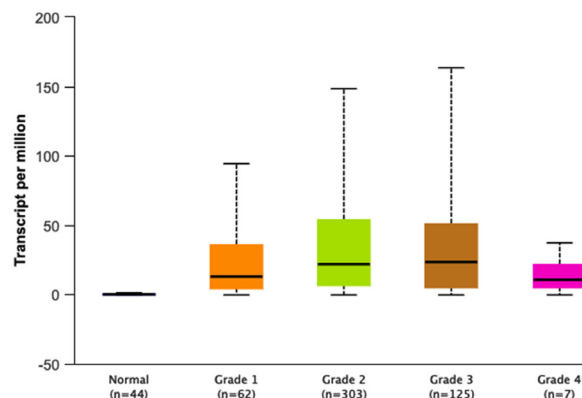
**Figure 2:** The bar graph showing Allele frequency of *rs2071676* polymorphism of *CA9* gene among different populations. X-axis representing the different populations [AFR - African; AMR - American; EAS - East Asian; EUR - European; SAS - South Asian] and Y-axis representing the percentage allele frequency. The reference allele frequency is represented by blue color and altered allele frequency is represented in orange color.



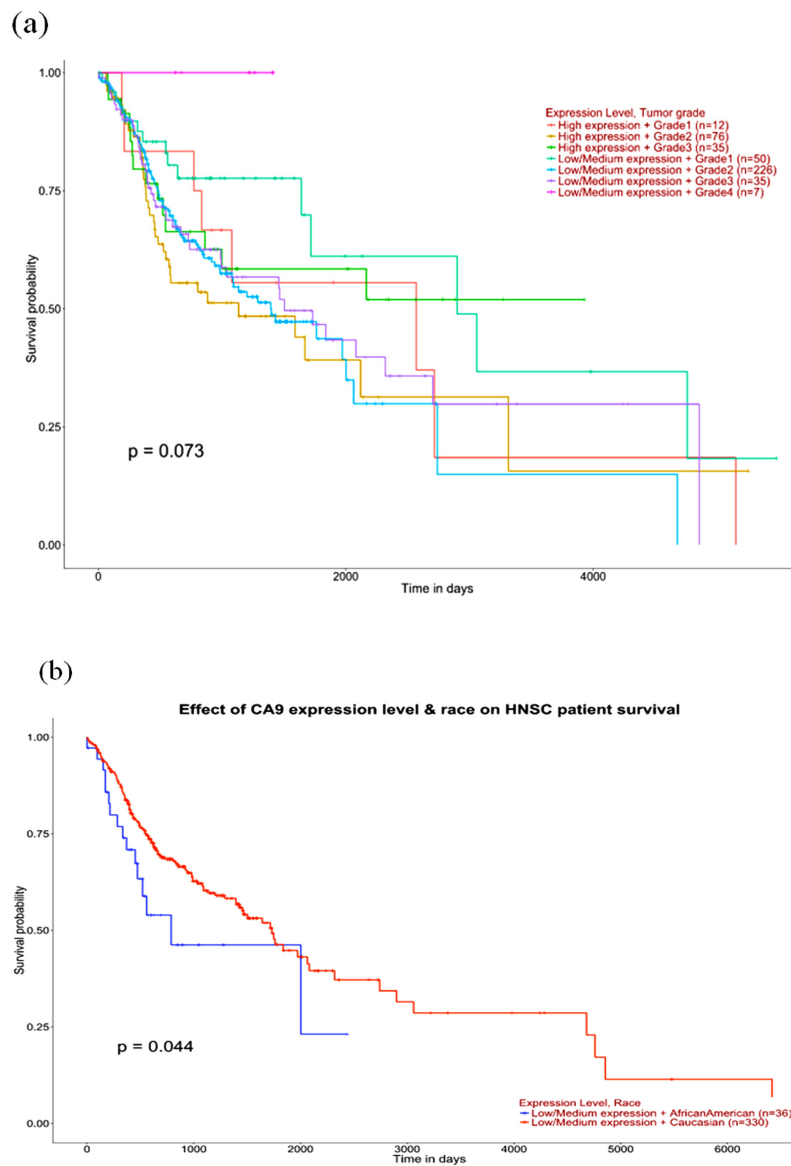
**Figure 3:** The bar graph showing the comparison of allele frequencies between East Asian [Ensembl] and case population [Taiwanese population - East Asian] derived from experimental data [Chien et al, 2012]. X-axis represented the East Asian and case population and Y-axis representing the percentage of allele frequency. The reference allele is represented in blue and altered allele is represented in orange color.



**Figure 4:** The bar graph showing the comparison of allele frequencies between East Asian population [Ensembl] and control [Taiwanese population - East Asian] derived from experimental data [Chien et al, 2012]. X-axis represented the East Asian and case population and Y-axis representing the percentage of allele frequency. The reference allele is represented in blue and an altered allele is represented in orange color.



**Figure 5:** The Boxplot showing expression profile of *CA9* gene in different grades of HNSCC tumor. The X-axis denotes grades of tumor relative to normal expression of *CA9* and Y-axis denotes mRNA counts expressed as TPM (transcript per million). The comparison of gene expression pattern between different grades of HNSC returned significant values between Normal vs Grade 1 ( $p=2.8 \times 10^{-6}$ ), Normal vs Grade 2 ( $p=1.6 \times 10^{-12}$ ), Normal vs Grade 3 ( $p=3.6 \times 10^{-13}$ ), Normal vs Grade 4 ( $p=3.7 \times 10^{-2}$ ), Grade 1 vs Grade 3 ( $p=3.5 \times 10^{-2}$ ), Grade 1 vs Grade 4 ( $p=1.14 \times 10^{-2}$ ), Grade 2 vs Grade 4 ( $p=1.44 \times 10^{-4}$ ), Grade 3 vs Grade 4 ( $p=2.3 \times 10^{-6}$ ).



**Figure 6:** Kaplan-Meier curve (a) showing the effect of differential CA9 gene expression on the survival of HNSCC patients with different grades of the tumor (p-value = 0.073). X-axis represented the time in days and Y-axis representing the survival probability. (b) showing the effect of CA9 gene expression level and race (low/medium expression in African-American vs Caucasian population, p value = 0.044) on HNSCC patient survival. X-axis representing the time in days and Y-axis representing the survival probability.