Genetic association of **SERPINB5** gene polymorphism (rs17071138) with susceptibility to HNSCC - An *In silico* study

Dr Snega T\(^1\), Dr J. Vijayashree Priyadharsini\(^2\), Dr. Abilasha R\(^3\)

Department of Oral and Maxillofacial Pathology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai, India.

Email id: 151907001.sdc@saveetha.com

Assistant Professor, Department of Microbiology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai, India.

Email id: vijayashreej.sdc@saveetha.com

Reader, Department of Oral and Maxillofacial Pathology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai, India.

Email id : abilasha@saveetha.com

*Corresponding authors:

Dr.J.Vijayashree Priyadharsini
Assistant Professor
Department of Microbiology,
Saveetha Dental College and Hospitals,
Saveetha Institute of Medical and Technical Sciences, Saveetha University,
162, PH Road, Chennai 600077,
Tamil Nadu, India.

Email id: vijayashreej.sdc@saveetha.com

**Abstract:**

**Aim:** Human genetic variations are found to influence the disease susceptibility, the way in which an individual responds to drugs and infections. The effect of alleles tends to differ in different ethnic groups or populations. The screening of such rare variants related to the risk of developing disease would aid in preparing an SNP panel.

corresponding to the disease which can be used to assess disease risk or identify vulnerable individuals in a population. The present study aims to perform a population level screening of allele frequencies of (rs17071138) polymorphism of SERPINB5 gene. **Material and methods:** Genotype frequencies of the SNP (rs17071138) were collected from the Ensembl database for different populations and the deviations were analysed. Furthermore, the expression profile of the SERPINB5 gene in HNSCC was assessed using *In-silico* tools. The survival of patients based on the expression of the SERPINB5 gene was also assessed. **Results:** The present study identified deviations in allele frequencies for SERPINB5 polymorphism between different populations. The minor allele frequency in the ancestral population or the African population was found to be much lower than the two populations viz., east Asian and south Asian. **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 upon the plausibility of association of SERPINB5 genes with HNSCC.

**Keywords:** Head and neck squamous cell carcinoma, polymorphism, SERPINB5, variations

**How to cite this article:** Snega T, Priyadharsin JV, Abilasha R (2020): Genetic association of SERPINB5 gene polymorphism (rs17071138) with susceptibility to HNSCC- An in silico study, Ann Trop Med & Public Health; 23(S22): SP232312. DOI: [http://doi.org/10.36295/ASRO.2020.232312](http://doi.org/10.36295/ASRO.2020.232312)

**Introduction:**

The global prevalence of HNSCC accounts for about 5%-10% annually [1]. Head and neck cancer is a heterogeneous disease that comprises various types of cancers involving multiple sub-sites including the oral cavity, oropharynx, larynx, and salivary glands [2]. HNSCC is characterized by a spectrum of chromosomal genetic and molecular attractions [3]. Mainly HNSCC is associated with diverse risk factors with the majority being oral habits such as tobacco and alcohol consumption which induces molecular changes in mucosal epithelial cells [4]. HNSCC can also arise from transformation of benign lesions to malignant state. Malignant transformation is largely seen in OPMDs [5,6]. OPMDs have statistically increased risk of progressing to cancer, but their risk varies according to a range of patients or lesion related factors. Around 300,000 patients are estimated to have an oral cancer world annually, and the incidence is higher in South Asian countries [7]. In 2005, WHO, the term potentially malignant disorder/lesion (OPMD/OPML) was suggested as a replacement for premalignant oral lesions and conditions [7,8],[9]. OPMDs are those lesions of the oral mucosa that are at an increased risk of malignant transformation [10][11] into SCC. The SERPINB5 gene which is a mammary serine protease inhibitor that has a
role in inhibition in cell proliferation, development, invasion, angiogenesis and metastasis [12,13]. Gene polymorphisms of SERPINB5 are rs17071138 T/C rs3744941 C/T that are found in the promoter region. These polymorphisms reduce transcription factor binding affinity that reduces the SERPINB5 gene expression [12].

There are a few similar studies on SERPINB5 gene expression, its polymorphism and environmental factors associated with oral cancer risk. Hsiu-tsai et al., has identified rs17071138 T/C rs3744941 C/T and rs8089104 T/C gene polymorphism of SERPINB5 that may be a factor to increase susceptibility to oral cancer. Single nucleotide polymorphisms (SNPs) are important factors in HNSCC susceptibility. It was reported previously that rs17071138 TC SERPINB5 messenger RNA expression in whole blood was significantly down regulated compared to rs17071138 wild type (WT) homozygous genotype [14]. No other studies have investigated the impact of the SERPINB5 gene to HNSCC in Indian population. The present in silico study was conducted to determine whether the deviations observed for the rs17071138 polymorphism of SERPINB5 gene has any effect relating to the risk or associated susceptibility of HNSCC.

Material and methods:

This study is an in silico study performed with rs17071138 polymorphism of SERPINB5 gene based on literature mining process. Since HNSCC is more prevalent in patients with chronic history of oral habits (forms of tobacco), the genetic variant which is very closely related to environmental factors was identified in the gene and selected. The data retrieved from the study done by Tsai et al., in 2017 with the case group of 741 patients and control group of 601 patients were used as reference data. 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) [15]. Furthermore, the expression of the gene in HNSCC was analyzed using the UALCAN (http://ualcan.path.uab.edu/cgi-bin/TCGA-survival1.pl?genenam=SERPINB5&ctype=HNSC) database. Survival curve analysis based on the tumor grade and expression profile was performed to demonstrate the putative role of SERPINB5 gene with HNSC [16]. Gene expression data is expressed as transcripts per million (TPM) which 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 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 [16].

**Results and discussion:**

The global phenotype data obtained from the Ensembl database for *rs17071138* polymorphism of the *SERPINB5* gene showed the frequency of T allele as 90% and C allele as 10% (Figure 1). The frequency distribution of the alleles of the selected polymorphism in different populations are shown in figure 2. A comparative analysis was performed with the documented allele frequencies of cases and control from research article Tsai, et al., with a global population of *rs17071138* polymorphism of *SERPINB5* gene. Since the population studied by Tsai et al., was Taiwanese population, a comparison between east Asian population and data from the research analysis was performed (Figure 3 and 4). Although the allele frequency documented in study and database were more or less similar in east Asians, the comparison between the ancestral population which is the African population [T-96% and C-4%] and south Asian population showed a marked deviation [T-74% and C-26%] (Figure 2). Oral potentially malignant disorders (OPMDs) have statistically increased risk of progressing to cancer, but the risk varies according to a range of patient or lesion related factors [17–19]. The WHO 2005 classification recognizes five histopathological stages in epithelial precursor lesions – squamous hyperplasia, mild dysplasia, moderate dysplasia, severe dysplasia, carcinoma *in situ* [20,21]. The genetic factor plays an important role in oral cancer and oral potentially malignant disorders [22,23] by influencing its susceptibility and disorder of potential oral cancer, related SNPs was recommended for the early detection of possible candidates for malignancy development/oral cancer development [24,25]. There are two crucial aspects in pathogenesis of tumor development which is the proliferation and metastasis of tumor cells [26]. Tumor suppressor genes play an important role in anticancer programs [27].

*SERPINB5* is unique and is an unusual member of Serpin superfamily, compared to other members in secondary structure; there is no protease inhibited by *SERPINB5* via the classical serpin pathway [28] and also it is a 42 kDa protein which includes inhibitory members that target proteases, as well as non-inhibitory members that possess a diverse array of functions. The expression of the gene was assessed using *in silico* tools which showed significant increase in the transcript number as the grade increases (figure 5). The expression of *SERPINB5* showed a marked difference between the races with a special emphasis on the African American population and the Asian...
population (figure 6). Although, the effect of the variant on gene expression may not be concluded with the data available unless experimentally validated. The present investigation could be the first step which has embarked on the identification of the relationship between genetic variants and the gene expression profile. The Kaplan-Meier survival analysis returned significant value (p=0.017) in comparison of the differential SERPINB5 gene expression between Caucasian (high) and African Americans (low) (figure 7). The analysis is indicative of the fact that HNSCC patients with low or medium expression in the African American population present with a poor survival rate when compared to high expression Caucasian groups (Figure 6). It is a universal fact that the body naturally tries to eliminate something that is abnormal or defective [29]. A hormone negative regulatory element is present in the promoter region of SERPINB5 gene [30]. The critical role of the respective gene is that it inhibits both tumor growth and metastasis and it also has a pro-apoptotic, anti-metastatic and anti-angiogenic properties exhibiting an inhibitory effect on cancer motility, invasiveness and metastatic ability [28,31]. It is the only biomarker that has significant association with carcinoma histology grade [27]. The SERPINB5 gene has been characterized as a tumor suppressor in several cancers. Epidermal growth factor (EGF) induced epithelial mesenchymal transition (EMT) of esophageal carcinoma cells can be inhibited by the SERPINB5 gene that suppresses proliferation, migration and invasion. Even induces transition of tumor cells into benign cells [32]. It also inhibits vasculogenic mimicry in non-small cell lung cancer (NSCLC) cells, and its expression level is positively correlated with prognostic implication in NSCLC patients [33]. In breast cancer and prostate cancer it inhibits apoptosis [34,35]. SERPINB5 suppresses prostate tumor growth, invasion and metastasis by inhibiting histone deacetylation, and activates neutrophils and B cell-dependent antitumor immune response [36]. Also its role in gastric and hepatocellular cancer has been significant with its critical function in suppression of tumor [14,24]. Therefore, SERPINB5 could be an effective marker for diagnosis and therapy of cancers. There is no other similar literature to support the SERPINB5 gene role in HNSCC.

Conclusion:

The study carried out by Hsui-Ting demonstrated the association of rs17071138 polymorphism of SERPINB5 in Taiwan population - East Asia. Hence this study was done as comparative analysis to identify the difference in the allele frequencies as mentioned in the study with frequency available in the database. Since, there was a marked deviation in the data analyzed, we expect a similar kind of difference in the South Indian population.
also. Interestingly, the south Asian population showed a dramatic increase in the C allele frequency, which was significantly higher than any other population. Genotype analysis in the south Indian population is warranted to derive clues about the positive selection of “C” allele of the rs17071138 polymorphism of SERPINB5 in east and south Asian groups.

Acknowledgement:

The authors would like to acknowledge the help and support rendered by the department of Oral Pathology & Microbiology of Saveetha Dental College and Hospitals and the management for their constant assistance with the research.

Authors’ contributions:

All the authors contributed to the design and implementation of the research, to the analysis of the results and to writing the paper and approved the final manuscript.

Conflict of interest:

The authors declare that they have no competing interests.

References:


Snega et al (2020): SERPINB5 gene polymorphism (rs17071138) with susceptibility to HNSCC


Figure legends

Figure 1: Pie chart depicting the global frequency data for the rs17071138 polymorphism of SERPINB5 gene as derived from the Ensembl database. The orange colour denotes T allele frequency = 90% and blue colour denotes C allele frequency = 10%.

Figure 2: Data obtained from Ensembl database for the rs17071138 polymorphism of SERPINB5 showing allelic frequency distribution among different populations [AFR - African; AMR - American; EAS - east Asian; EUR - European; SAS - south Asian].

Figure 3: Comparison of allele frequencies between East Asian allele frequency [Ensembl] and case allele frequency derived from experimental data [Tsai et al, 2017].

Figure 4: Comparison of allele frequencies between East Asian allele frequency [Ensembl] and control allele frequency derived from experimental data [Tsai et al, 2017].

Figure 5: The expression profile of SERPINB5 gene in different grades of OPMD tumor. The comparison of gene expression pattern between different grades of OPMD returned significant values between Normal vs Grade 1 (p=5.9 X 10^{-4}), Normal vs Grade 2 (p=1.9 X 10^{-2}), Grade 1 vs Grade 2 (p=9.7 X 10^{-3}), Grade 1 vs Grade 3 (p=3.03 X 10^{-8}), Grade 1 vs Grade 4 (p=1.16 X 10^{-4}), Grade 2 vs Grade 3 (p=9.8 X 10^{-10}), Grade 2 vs Grade 4 (p=3.5 X 10^{-3}).

Figure 6: Expression of SERPINB5 in OPMD based on the patient’s race. The comparison of gene expression between different races returned significant values between the groups Normal vs Asian (p=1.3 X 10^{-02}), Caucasian vs African American (p=2.6 X 10^{-02}), Caucasian vs Asian (p=4.6 X 10^{-02}), African-American vs Asian (p=4.4 X 10^{-02}).

Figure 7: The effect of differential gene expression level and race (high expression in Caucasian vs Low/medium expression in African-American population, p value = 0.017) on HNSC patient survival
Figure 1: Pie chart depicting the global frequency data for the rs17071138 polymorphism of SERPINB5 gene as derived from the Ensembl database. The orange colour denotes T allele frequency = 90% and blue colour denotes C allele frequency = 10%.
Figure 2: Bar chart depicting the frequency distribution of alleles as obtained from the Ensembl database for the rs17071138 polymorphism of SERPINB5 gene. The frequency of “T” and “C” allele is denoted by orange and blue bars respectively. The X axis denotes the different populations viz., African, American, east Asian, European, south Asian and global frequency data. Y axis denotes the percentage of allele frequencies. There exist a marked difference in the C allele frequency between the south and east Asian populations when compared to the ancestral African population.

Figure 3: Bar chart depicting the comparison of allele frequencies between East Asian allele frequency [Ensembl] and case allele frequency derived from experimental data [Tsai et al, 2017]. The blue and orange colour represents
the “T” and “C” allele frequency respectively. The X axis denotes the East Asian population and case group [Tsai et al, 2017] and the Y axis denotes the percentage of allele frequencies.

Figure 4: Bar chart depicting the comparison of allele frequencies between East Asian allele frequency [Ensembl] and control allele frequency derived from experimental data [Tsai et al, 2017]. The blue and orange colour represents the “T” and “C” allele frequency respectively. The X axis denotes the East Asian population and control group [Tsai et al, 2017]. The Y axis denotes the percentage of allele frequencies.
Figure 5: Box-Whisker plot showing relative expression profile of SERPINB5 gene in different grades of HNSCC. The X axis denotes the TCGA samples and Y axis denotes the transcripts per million values. The comparison of gene expression pattern between different grades of HNSC returned significant values between normal vs grade 1 (p=5.9 X 10^{-4}), normal vs grade 2 (p=1.9 X 10^{-2}), grade 1 vs grade 2 (p=9.7 X 10^{-3}), grade 1 vs grade 3 (p=3.03 X 10^{-8}), grade 1 vs grade 4 (p=1.16 X 10^{-4}), grade 2 vs grade 3 (p=9.8 X 10^{-10}), grade 2 vs grade 4 (p=3.5 X 10^{-3}). A p value less than 0.05 is considered to be significant.
Snega et al (2020): SERPINB5 gene polymorphism (rs17071138) with susceptibility to HNSCC

Figure 6: Box-whisker plot showing expression of SERPINB5 based on the HNSC patient’s race. The x axis denotes the expression of SERPINB5 among different races and y axis denotes mRNA counts expressed as TPM (transcripts per million). The comparison of gene expression pattern between different races returned significant values between the groups Normal vs Asian (p=1.3 X 10^{-02}), Caucasian vs African American (p=2.6 X 10^{-02}), Caucasian vs Asian (p=4.6 X 10^{-02}), African-American vs Asian (p=4.4 X 10^{-02}). A p value less than 0.05 is considered to be significant.
Figure 7: Kaplan–Meier plots showing the association of SERPINB5 expression in combination with the race with HNSC patient's survival. The x-axis represents time in days and y-axis shows the survival probability. The red line denotes low/medium level expression of SERPINB5 in African American population, while the blue line denotes high level expression of SERPINB5 in Caucasian population. The effect of differential gene expression level and race (high expression in Caucasian vs Low/medium expression in African-American population) returned a significant p value of 0.017. The results indicated a poor survival rate was observed in patients of African American population exhibiting low/medium level expression of the SERPINB5 gene. A p value less than 0.05 is considered to be significant.