An Investigation on the Genetic Alterations of the ALG3 Gene and Their Possible Correlation with the Development of HNSCC and LUSC

Vipra Sharma,1 Anitha Pandi,2 Smiline Girija Aseervatham Selvi,1 Paramasivam Arumugam,2 Vijayashree Priyadharsini Jayaseelan2

1Department of Microbiology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India
2Centre for Cellular and Molecular Research, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India

Abstract

Objectives: The ALG3 gene has been associated with various hallmarks of cancer including proliferation and metastasis. However, the relationship between the genetic alterations in ALG3 with the pathophysiology of HNSCC and LUSC is poorly understood. The present study aimed to demonstrate the genetic alterations in the ALG3 gene and its consequences at the molecular level in head and neck squamous cell carcinoma (HNSCC) and lung squamous cell carcinoma (LUSC).

Methods: The cBioportal database was employed to identify the genetic alterations. UALCAN was used to demonstrate the gene expression and survival analysis. The microRNAs targeting ALG3 were demonstrated using the miRDB database.

Results: The ALG3 gene exhibited 20% and 51% alteration in HNSCC and LUSC patients respectively. There was a statistically significant increase in the expression of the ALG3 gene in both datasets. The HNSCC patients presenting with high expression of ALG3 were found to exhibit a low survival probability when compared to the low/medium expression group.

Conclusion: The gene amplification of ALG3 correlated well with the gene expression status and survival of patients in the HNSCC dataset. This provides evidence of the possible involvement of the ALG3 gene with HNSCC, however, experimental evidence is warranted to prove this association.

Keywords: Genetic variation, mutation, gene expression, prognosis, survival, carcinoma


Post-translational modification of crucial proteins is associated with several physiological and pathological outcomes.1 The glycans play a pivotal role in the development of cancer. Among different types of glycosylation processes, N-glycosylation is a major modification, which is of 3 types, high-mannose, complex and hybrid. There are six mannosyl transferases viz., beta 1,4-mannosyltransferase encoded by ALG1 gene, alpha 1,3 mannosyl transferases encoded by ALG2 and ALG3, alpha 1,2 mannosyl transferases encoded by ALG9 and ALG11 and alpha 1,6 mannosyl transferases encoded by ALG12. They function as key modulators of pathways such as cellular signalling and communication, dissociation and invasion, cellular matrix interactions, angiogenesis and immune evasion. Glycosylation not only causes tumour metastasis but also triggers transformation into a malignant phenotype.
The genetic alterations such as translocations, amplification, deletions and mutations in candidate genes are key modifiers of the gene and protein expression. The translocation results in the production of fusion proteins, whilst amplification and deletions can result in an increase or decrease in copy numbers of the mRNA transcripts. The consequence of single nucleotide substitutions often known as mutation/variation can have a profound effect on the functional or structural aspects of a protein. The ALG3 gene has been reported to exhibit differential expression in several cancer types including oesophageal squamous cell carcinoma[4], cervical cancer[5], and hepatocellular carcinoma[6]. Interestingly, the gene and protein expression demonstrated several-fold increases when compared to paired normal tissues.

Furthermore, ALG3 is also considered to be a potential prognostic marker wherein the increased expression resulted in poor overall survival of lung adenocarcinoma patients. Computational approaches have largely contributed to the selection of candidate genes from a large collection of genes. Several studies related to gene families[8, 9], gene-gene or protein-protein interaction networks[10] have been conducted to identify the possible association of the candidate genes with cancer phenotypes. Despite the availability of data related to gene or protein expression in different cancer types, the influence of genetic modifications on ALG3 gene expression in HNSCC (Head and neck squamous cell carcinoma) and LUSC (Lung squamous cell carcinoma) has not yet been reported. These two datasets were chosen for the present analysis because of the kind of similar risk factors associated with them. The results accumulated through several experimental reports, aided in designing the present study which was employed to identify the association between the genetic alteration observed in the ALG3 gene, its influence over gene expression and the survival status of HNSCC and LUSC patients. Also, the epigenetic control of ALG3 in terms of small non-coding RNA such as microRNA was analyzed.

Methods

Sample Dataset
The samples used in the present study were from two datasets viz., head and neck squamous cell carcinoma (HNSCC) and Lung Squamous cell carcinoma (LUSC) (TCGA, Firehose Legacy). The demographic and clinical details of the patients of both datasets were included in Table 1.

Gene Alteration Analysis
The CIBIOPortal database (http://cibioportal.org) is an essential platform that hosts clinical and molecular information from numerous cancer types. The platform allows for the analysis of specific genes in terms of mutations, gross abnormalities, and expression of genes in relation to genetic alterations.[11, 12] The ALG3 gene was included in the query and each of the datasets was selected. The query returned oncoprint data that was further assessed for the type of variation and their frequencies were documented.

Gene Expression Analysis
The gene expression data of ALG3 of both datasets were demonstrated using the UALCAN (http://ualcan.path.uab.edu) database. This portal is user-friendly and was employed to perform an exhaustive analysis of a queried gene using TCGA gene expression data. The survival analysis was performed to generate the Kaplan-Meier survival plot. The survival plots were based on different methods of stratification of the data based on gender, tumour grade and ethnicity. The overall gene expression profiles were also organised into two types, low/medium expression and high expression for assessing the survival probability. The “survival” and “survminer” packages were used to generate survival plots. The comparison of gene expression profiles was performed using a log-rank test.[13]

miRNA Expression Analysis
The gene expression can be markedly influenced by epigenetic factors such as methylation of DNA, modifications in histone proteins and non-coding RNAs. Among these factors, microRNAs are considered potential drivers of cancer phenotypes and are associated with several cancer types. The analysis of microRNAs targeting the ALG3 gene could provide a clue about the differential expression pattern demonstrated by the gene. The miRDB is the database to identify microRNAs targeting specific gene transcripts. The upregulation of microRNA results in a decrease in mRNA copies of the target gene which eventually leads to a decrease in protein expression. Thus, the prediction of miRNAs targeting the differentially expressed gene is vital in understanding the role of epigenetic factors in the process of carcinogenesis. The miRDB database (http://mirdb.org) was employed to identify the microRNAs targeting ALG3 gene transcripts.[14, 15]

Protein-Protein Interaction Analysis
Since, cancer is a complex polygenic disorder, understanding the protein network interactions of the ALG3 gene becomes vital. The STRING (Version 10.5) database was used to derive the network of the ALG3 gene. The association events were categorised into co-expressions, gene neighbourhoods, gene fusions, co-occurrence, experiments, data mining and text-mining. The protein-protein interac-
tion data thus obtained from different sources were collected, analyzed and collated to derive an interpretation.[16]

Results

Genetic Alterations
The present study follows an in silico approach to demonstrate the prognostic significance of the ALG3 gene in HNSCC and LUSC datasets. The HNSCC and LUSC share similar risk factors and hence these two datasets were combined to dissect the similarities and differences in gene expression profiles among the patients. The HNSCC dataset had 528 samples which included 386 males and 142 females of age between 19-90 years. This dataset included 515 smokers and 352 patients with a history of alcohol intake. The majority of HNSCC patients presented at the grade 2 neoplasm (59%). On the other hand, there were 511 patients in the LUSC (TCGA, Firehose Legacy) dataset, of which 373 were male, 131 were female and seven patients whose gender remained unknown (Table 1a and 1b). The LUSC patients belong to the age group between 50 - 85 years. The majority of the patient population were smokers and data was unavailable for 24 other patients. The oncoprint analysis demonstrated 20% (HNSCC) and 51% (LUSC) alterations in the ALG3 gene (Figs. 1a and 1b). The predominant alteration in both the groups was gene amplification, however a missense mutation and splice site mutation were observed in HNSCC and LUSC datasets respectively.

Gene Expression and Survival
Owing to gene amplification, an increase in gene expression was anticipated. Interestingly, the gene expression of the ALG3 gene was found to be significantly upregulated in both datasets. The gene expression in normal vs. primary tissue derived from HNSCC patients revealed a p-value of 1.624 x 10-12 (Fig. 2a). A similar presentation was observed in the LUSC group (Fig. 2b). The Kaplan Meier plot demonstrated the survival probability of patients presenting with high and low levels of ALG3. A statistically significant asso-

<table>
<thead>
<tr>
<th>Table 1a. Demographic details of the HNSCC dataset (TCGA, Firehose Legacy) used for analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Mutation count</td>
</tr>
<tr>
<td>Diagnosis age</td>
</tr>
<tr>
<td>Smoking status</td>
</tr>
<tr>
<td>Alcohol history</td>
</tr>
<tr>
<td>Neoplasm Histologic grade</td>
</tr>
<tr>
<td>Race category</td>
</tr>
<tr>
<td>Race category</td>
</tr>
<tr>
<td>Race category</td>
</tr>
<tr>
<td>Race category</td>
</tr>
<tr>
<td>Race category</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1b. Demographic details of the LUSC dataset (TCGA, Firehose Legacy) used for analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Race</td>
</tr>
<tr>
<td>Race</td>
</tr>
<tr>
<td>Race</td>
</tr>
<tr>
<td>Race</td>
</tr>
<tr>
<td>Mutation count</td>
</tr>
<tr>
<td>Cancer type</td>
</tr>
<tr>
<td>Smoking Status</td>
</tr>
<tr>
<td>Smoking Status</td>
</tr>
</tbody>
</table>

Figure 1. Oncoprint data demonstrating gene alterations in the ALG3 gene in (a) HNSCC and (b) LUSC datasets.
A significant association was observed between the survival probability of patients and ALG3 expression (p=0.017) in HNSCC groups (Fig. 3a). The survival status of HNSCC patients with high expression of the ALG3 gene was found to be markedly reduced when compared to the low or medium-expression group. Therefore, it is evident that the overexpression of the ALG3 gene was associated with poor prognosis. On the other hand, no significant association was found between expression levels and survival state in the LUSC dataset (p=0.17) (Fig. 3b). As with the LUSC dataset, the gene expression profile was found to correlate with the gene amplification state. However, this presentation did not influence the survival of the patients. Such an observation can be attributed to the epigenetic factors that should be assessed further to reveal the role of ALG3 in LUSC.

**Protein-Protein Interaction Network**

Cancer is a complex polygenic disorder, where multiple genes are dysregulated and the cumulative effect of all the genetic aberrations leads to a malignant phenotype.

---

**References**

Sharma et al., Genomic Aberrations in ALG3 Gene in LUSC and HNSCC / doi: 10.14744/ejmo.2024.23872
Hence, a protein interaction network analysis was initiated to identify the proteins that are reacting with ALG3. The process provided preliminary data with a set of 10 genes of the ALG family viz., ALG9-2, ALG6, ALG12, ALG11, ALG5, ALG10B and protein belonging to Dolichol phosphate-mannose biosynthesis regulatory protein family which included DPM1, 2 and 3 (Fig. 4).

miRNA Expression

The over-expression of the ALG3 could be attributed to the epigenetic factors that modulate the expression process. The hub of microRNAs targeting the ALG3 gene was acquired from the miRDB database. Seven microRNAs hsa-miR-544b, hsa-miR-9900, hsa-miR-374c-3p, hsa-miR-3202, hsa-miR-10393-5p, hsa-miR-10393-5p, hsa-miR-571 and hsa-miR-767-3p were found targeting the ALG3 gene (Table 2). The increased expression of a target gene can be attributed to the downregulation of the specific miRNAs. The present study did not identify such a microRNA; more intense investigation on the role of other epigenetic modifiers in the expression of the ALG3 gene is warranted.

Discussion

Head and neck squamous cell carcinoma and lung squamous cell carcinoma are considered to be the leading cause of death due to cancer in the South Asian regions.[17] These cancer types have similar risk factors associated with them such as smoking, tobacco chewing, exposure to carcinogenic chemicals or pollutants etc.[18] The differentially expressed biomarkers are being explored for their application as prognostic, diagnostic or therapeutic leads. These markers provide clues about the interaction networks involved in the disease pathogenesis. Given this, the ALG3 gene was investigated for genetic alterations, gene expression and the possible role of microRNA targets in regulating the expression of the ALG3 gene. In the present study, we observed that the expression of the ALG3 gene was high in both datasets when compared to the normal tissues. The expression profile positively correlated with the survival status of patients wherein the overexpression of ALG3 led to poor survival rate in HNSCC. As with the LUSC dataset, the genetic alteration and expression had a positive correlation, but the increased expression was not found to alter the survival of LUSC patients significantly. Several genetic and epigenetic factors might play a role in establishing this observation. The pan-cancer analysis reiterated the fact that ALG3 is overexpressed in several cancer types such as breast, glioblastoma, hepatocellular carcinoma and many more (data not shown).

A study conducted by Shao et al. investigated the potential mechanism of action of ALG3 in OSCC. They employed techniques such as qPCR and Western blot to measure the gene expression and protein expression of ALG3. They observed that an increase in the expression of ALG3 resulted in aggressive behaviour of OSCC cells including lymph node metastasis. The depletion of ALG3 resulted in a concomitant reduction in the expression of MCM7/CCNB2/CDK1/PCNA, all of which are involved in the cell cycle process. The over-expression of ALG3 transcripts also correlated with the poor survival of OSCC patients.[19] The qRT-PCR has investigated the expression of all the members of asparagine-linked glycosylation in breast cancer cases. Among all the genes assessed, ALG3 was found to be overexpressed in breast cancer cells which were radioresistant, a process mediated by the glycosylation of TGF-beta receptor II. The upregulation of ALG3 correlated well with elevated radioresistance, and stemness which contributed to poor prognosis among breast cancer patients. Consolidation of the results reported that ALG3 can serve as a marker indicative of radioresistance.[20] The results were in close agreement with the observations made in the present study.

Wu and colleagues employed computational tools to as-
sessed the gene expression profile of glycosyltransferase-related genes (GTRGs) in HNSCC. Several pathways leading to tumour immunity were assessed for their gene expression pattern. It was suggested that the signatures obtained could aid in discriminating between the individuals who benefit and those who do not benefit from immunotherapy. Also, they identified that the tumour mutational burden was much higher in the candidate genes such as TP53, CDKN2A, MUC17 and CSDM1. A high throughput sequencing of non-coding RNAs in patients diagnosed with oral lichen planus and OSCC revealed potential gene targets of long non-coding RNAs. The ALG3 gene was the differentially expressed gene that topped the list of upregulated genes in the OSCC group compared to normal tissues, thus providing substantial evidence of the involvement of this gene in the pathophysiology of OSCC.

An interesting finding on the role of the ALG3 gene in non-small cell lung cancer (NSCLC) was demonstrated by Ke and the team. They revealed that the expression of ALG3 in NSCLC tissues was markedly higher and was associated with tumour stage, poor differentiation of tissues and lymph node metastasis. This presentation also correlated well with the survival of patients which was compromised to a larger extent. The knockdown of ALG3 expression resulted in diminished expression of N-cadherin and Vimentin, which are known markers observed in invasive tumour cells. There was a concomitant inhibition of proliferation and migration of NSCLC cells. Additionally, they also identified miR-98-5p that specifically binds to 3′-UTR of the ALG3 gene and reduces the oncogenic functions.

The evidences discussed here provides a vivid picture of the role of the ALG3 gene in HNSCC. This study is the first of its kind to analyze gene expression and survival of patients with HNSCC and LUSC following the gene alterations observed. The study reiterated the fact that genetic alterations induced via carcinogens or other exposure factors are the key drivers of malignant phenotypes. Although genetics play a very vital role in tumour initiation, the progression of tumours is largely influenced by epigenetic factors. Hence, a proper exploration of the genetic and epigenetic components in specific cancer types would provide more information on the molecular pathogenesis of cancer. Owing to the computational approach that was used for the investigation of datasets, there were certain limitations to be addressed viz., (a) the study design is an in silico approach which requires further experimental validation, (b) the patient group were not categorised based on the habits or exposures, (c) several other epigenetic markers could influence the gene expression, hence a more exhaustive study design has to be planned to derive concrete evidence on the role of ALG3 with HNSCC.

Conclusion

The study demonstrated the consequences of ALG3 gene alteration on the expression profile and survival of HNSCC patients. The differential expression pattern exhibited by this gene can make it a potential candidate for diagnosis of HNSCC. Also, targeting the gene products with small-molecule inhibitors or drugs could lead to the design of therapeutic agents. Nevertheless, exploring the pathways and networks connecting the ALG3 gene can broaden the theragnostic role of this gene in HNSCC.

Disclosures

Acknowledgement: The authors are grateful to all the Cohorts and groups involved in the compilation of data from patients for public use. Our sincere thanks also go to all the patients who have indirectly contributed to the scientific community by providing consent for sharing their data for research use.

Ethics Committee Approval: The study was approved by the Local Ethics Committee.

Peer-review: Externally peer-reviewed.

Conflict of Interest: None declared.

Authorship Contributions:

Concept – V.P.J.; Design – V.P.J., PA; Supervision – V.P.J., PA; Materials – VS, AP; Data collection &/or processing – VS, AP; Analysis and interpretation – VS, AP, PA; Literature search – VS, AP, SGAS; Writing – VS, SGAS, VPJ; Critical review – VPJ.

References