Multi-omics analysis of *NFE2L2* mutated TCGA-Cervical Squamous Cell Carcinoma patients

Sri Vidya Ramisetti1, Akhileshwar Namani2, 3 *

1 Department of Biotechnology, GITAM School of Sciences, GITAM (Deemed to be University), Visakhapatnam, 530045, India; email : ramisetti.srividya678@gmail.com

2 Department of Biotechnology, GITAM School of Sciences, GITAM (Deemed to be University), Visakhapatnam, 530045, India; email: anamani@gitam.edu

3 Current address: Sri Shankara Cancer Hospital and Research Centre, Bangalore, India; email: akileshwarnamani@ssnccpr.org

***Correspondence: akileshwarnamani@ssnccpr.org**

Abstract

Genetic alterations in *NFE2L2* gene have been identified across various cancers and the dysregulation of the NRF2 pathway due to these alterations leads to drug and radioresistance in several cancers. Identification of biomarkers associated with these alterations allows the researchers and clinicians to identify the personalized medicine and quicker diagnosis. In this current study, we carried out an integrated, multi-omics, multi-database analysis of exome, transcriptomics data's of *NFE2L2* altered TCGA-Cervical squamous cell carcinoma (CSCC) patients against wild type counterparts. Finally, we discovered the genes associated with *NFE2L2* alterations, identified the prognostic genes which could be used as potential biomarkers in the *NFE2L2* mutated CSCC patients. Our finding might be useful to identify the early diagnosis of *NFE2L2* mutated CSCC patients.

Keywords

NFE2L2, Cervical Cancer, biomarkers, therapeutic strategies, multi-omics

Introduction

Cervical cancer is the fourth most common cancer amongst in women, accounting for approximately 6.5% of all female cancer cases worldwide [1]. The Cancer Genome Atlas (TCGA) is a publicly funded project that aims to catalog and discover major cancer-causing genome alterations to create a comprehensive “atlas” of cancer genome pro-files [2]. *NFE2L2* is a gene that encodes the transcription factor NRF2 (nuclear factor erythroid 2-related factor), which is the key regulator of oxidative stress in normal cells [3]. Genetic alterations such as mutations and amplification in the *NFE2L2* gene can affect the stability, localization, and activity of the NRF2 protein. These alterations have been identified in many cancers, including cervical squamous cell carcinoma (CSCC), and dysregulation of NRF2 signaling due to these alterations leads to tumorigenesis, drug and radiation resistance. Identifying biomarkers associated with these alterations allows the researchers and clinicians to develop personalized medicine and faster diagnosis [4].

Methodology

1.Identification of genetic alterations of NRF2 in TCGA-CSCC

The cBioportal for cancer genomics website was used to identify the *NFE2L2* mutational landscape and amplifications in CSCC patients from the TCGA pan-cancer study (n =251).

2.Analysis of differentially expressed genes (DEG’s) in NRF2 altered TCGA-CSCC

Based on the NRF2 genetic alterations of TCGA-CSCC, we stratified the total number of patients into two groups and designated them as *NFE2L2*-altered (n=20) and wild-type (n=231) (without NFE2L2 alterations), respectively. The mRNA expression pro-files (RNA Seq-RSEM batch normalized from Illumina HiSeq_RNASeqV2) were checked to identify the DEG’s in these two groups. The *NFE2L2* alterations result in upregulation of its downstream genes [6]. From the list of upregulated genes, we can conclude that they are the driving genes behind tumorigenesis and cancer progression.

3.Functional annotation and survival analysis

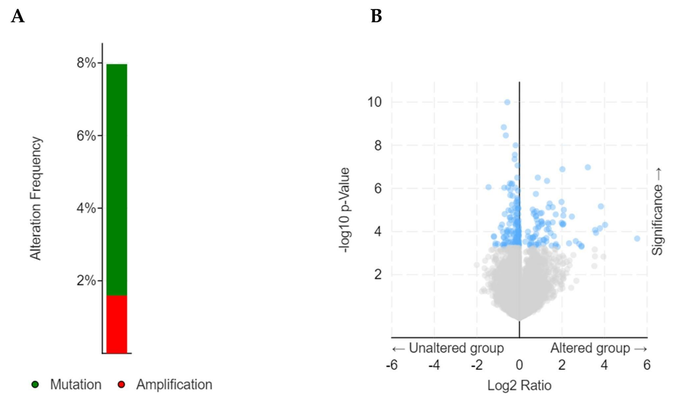
The Functional annotation of the upregulated genes from *NFE2L2* altered patients was performed by a web tool named DAVID [7]. This analysis provides the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway information for genes.

The Kaplan–Meier Plotter [8] tool was used to evaluate the prognostic value of the 29 upregulated genes identified in the *NFE2L2* altered patients in TCGA-CSCC cohort. Briefly, for TCGA-CSCC cohort, the patient samples are divided into two risk groups such as low-risk and high-risk groups based on the prognostic index (PI).

4.Identification of NRF2-binding sites by in silico analysis

LASAGNA-Search 2.0 [9] is an integrated web tool for searching and visualizing transcription factor binding sites (TFBS). In this study, LASAGNA-Search 2.0 with cutoff p-values <0.001 was used to identify the NRF2 TFBS within the promoter regions of upregulated genes from *NFE2L2*-altered patients. The search was restricted to the -2 kb up-stream human promoter region relative to the transcription start site.

Results



C

Gene Ontology Biological Processes		
GO BP term	Gene Count	p value
GO:0009636~response to toxic substance	4	2.20E-04
GO:0042572~retinol metabolic process	3	0.002384012
GO:0016488~farnesol catabolic process	2	0.002687962
GO:0042377~vitamin K catabolic process	2	0.005368977
GO:0042361~menaquinone catabolic process	2	0.005368977
GO:0034599~cellular response to oxidative stress	3	0.006617267
GO:0036101~leukotriene B4 catabolic process	2	0.006706884
GO:0042376~phylloquinone catabolic process	2	0.006706884
GO:0006805~xenobiotic metabolic process	3	0.008248696
KEGG Pathways		
hsa00040:Pentose and glucuronate interconversions	3	0.002043251
hsa01100:Metabolic pathways	9	0.004649962
hsa00140:Steroid hormone biosynthesis	3	0.006297662
hsa00983:Drug metabolism - other enzymes	3	0.010311172
hsa01240:Biosynthesis of cofactors	3	0.034935112
hsa00790:Folate biosynthesis	2	0.049557718

Figure 1. (A) Percentage of *NFE2L2* genetic alterations in TCGA-CSCC patients. (B) Volcano plot showing the DEG’s between *NFE2L2*-altered vs wild-type patients. (C) GO and KEGG pathway analysis of upregulated genes in *NFE2L2*-altered patients

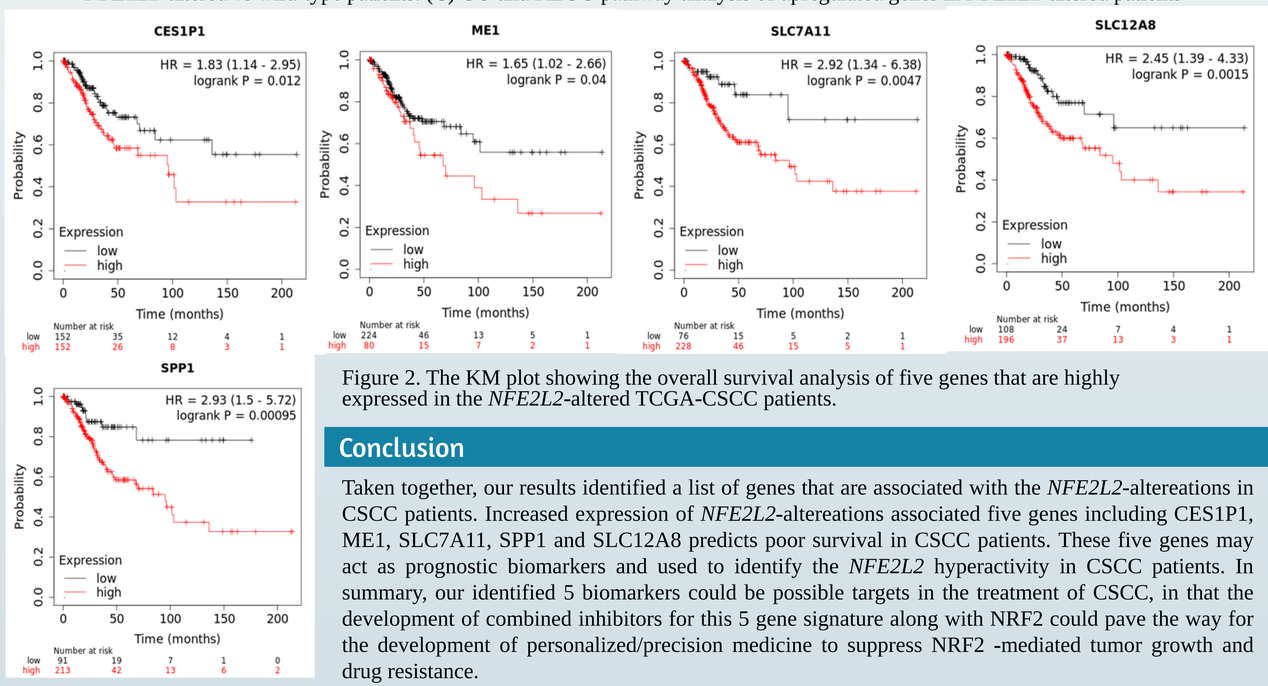


Figure 2. The KM plot showing the overall survival analysis of five genes that are highly expressed in the *NFE2L2*-altered TCGA-CSCC patients.

Conclusion

Taken together, our results identified a list of genes that are associated with the *NFE2L2*-altereations in CSCC patients. Increased expression of *NFE2L2*-altereations associated five genes including CES1P1, ME1, SLC7A11, SPP1 and SLC12A8 predicts poor survival in CSCC patients. These five genes may act as prognostic biomarkers and used to identify the *NFE2L2* hyperactivity in CSCC patients. In summary, our identified 5 biomarkers could be possible targets in the treatment of CSCC, in that the development of combined inhibitors for this 5 gene signature along with NRF2 could pave the way for the development of personalized/precision medicine to suppress NRF2 -mediated tumor growth and drug resistance.