The expression and clinical value of LRP11, FUBP1 and TET1 in cervical cancer

Xiaojie Zhao1*, Suhua Li1*, Jiexin Wen**, Yawei Li**

1 Inspection Center, the First Hospital of Hebei Medical University, Shijiazhuang, Hebei050031, China
2 Department of Blood Transfusion, the First Hospital of Hebei Medical University, Shijiazhuang, Hebei050031, China
3 Department of Gynaecology, the First Hospital of Hebei Medical University, Shijiazhuang, Hebei050031, China

ABSTRACT

Cervical cancer is the second leading cause of cancer death among women worldwide. Identification of effective genes along with biological markers as targeting agents is very necessary for the diagnosis and treatment of this disease. Bioinformatics techniques along with genetic and molecular investigations have provided the possibility of studying different levels of information such as the genome, transcriptome, proteome, and metabolome with high depth and accuracy. The collection of these data provides comprehensive and valuable information about the investigated phenotypes, including complex diseases such as cancer. In this study, we examined three genes LRP11, FUBP1, and TET1 related to cervical cancer. The results of this study showed that the level of expression of these genes is high in lymph nodes and the thyroid and is less in the pancreas and liver. Also, the expression level of the FUBP1 gene is higher than that of LRP11, and the expression level of the LRP11 gene is higher than that of TET1. Regarding the structure and proteomics of the studied genes, it can be seen that due to the presence of more domains in the LRP11 and FUBP1 genes, these genes probably independently participate in various functions and have a wider range of activity than the TET1 gene. Also, the analysis of the stability of the examined genes showed that the stability of the FUBP1 gene is relatively higher than that of the TET1 gene, and this gene is also more stable than the LRP11 gene. Considering that these genes are effective key genes for the early detection of cervical cancer, it is hoped that they will be used as markers in the diagnosis and treatment of cervical cancer.

Introduction

In 2018, about 18 million new cancers were registered, and according to the statistics and information mentioned in the same year, cervical cancer ranked fourth among women's cancers in terms of disease incidence and mortality rate (1). It is one of the most dangerous and deadly diseases of the 20th century and it is expected to be one of the great challenges to human health in the 21st century. While infection with human papillomaviruses plays an important role in the development of cervical cancer, the presence of the virus alone is not sufficient to cause cancer. Genetics and environmental factors such as smoking and unhealthy eating habits also play a role in this. Today, mankind is constantly exposed to various carcinogenic factors such as mutations, radioactive substances, and chemical and environmental pollution (2). Every year, approximately 470,000 new cases of cervical cancer are diagnosed, and approximately 230,000 women around the world are diagnosed due to this. Most of the deaths (approximately 80%) occur in developing countries (3). The prevalence of cervical cancer is different all over the world the highest rate is in East Africa and the lowest in West Asia (4).

In addition to tumor suppressors and carcinogenic oncogenes, various receptors such as EPHA4, EPHA5, and EPHB2, endothelin receptors EDNRB and EDNRA, as well as nuclear receptors NR2C1 and NCOA3 are involved in cervical cancer disease (5). The use of new approaches, such as microarray gene expression profiles and RNA-seq, has provided an unprecedented look at the function of the genome. The accurate extraction of such data and their analysis makes it possible to accurately identify the genes and biological pathways involved in various diseases, including cancer. Considering the vast amount of information available in these data and their huge dimensions, the analysis of these data requires the use of information and bioinformatics networks that both facilitate the understanding of information for biologists and provide many tools for the analysis of existing networks.

Although infection with the human virus is necessary for the development of cervical cancer, it alone is not sufficient to explain cervical cancer; therefore, other risk factors such as genetic factors, specific host and their importance in the induction of cervical cancer should be known. Considering that gene expression profile studies in the last decade have provided significant molecular findings about oral cancer, sufficient screening and effective treatment strategies have not yet been achieved (5). In the investigations, genes such as LRP11, FUBP1, and TET1 have been identified in cervical cancer. The present study was conducted to investigate the effect of these genes on cervical cancer. In general, the bioinformatics analysis and expression analysis of these influential genes can be further evaluated in order to develop new therapeutic and
LRP11 gene

The LRP11 gene with accession number NM_032832 is located on chromosome 6 (6q25.1). This gene activates the binding of phosphoproteins and it is predicted that its location is in the plasma membrane. The number of amino acids, molecular weight, and isoelectric point of the number of exons were determined in Table 1 for the LRP11 gene. LRP11 has an effective role in proliferation, migration, and invasion and can be used as a useful prognostic marker for cancer diagnosis. Also, the LRP11 gene is known as one of the therapeutic targets for patients with HSIL and cervical cancer (6-7). Figure 1(A) shows the location of this gene on the chromosome.

FUBP1 gene

The FUBP1 gene with accession number NM_032832 is located on chromosome 1 (1p31.1). The number of amino acids, molecular weight, and isoelectric point of the number of exons were determined in Table 1 for the FUBP1 gene. The protein made by this gene is a single-stranded protein that binds to a series of upstream elements such as FUSE and c-myc. Binding to FUSE leads to the up-regulation of the c-myc factor, which occurs mostly in undifferentiated cells. This protein has 3 specialized domains, the N-terminal region and the C-terminal region of the central DNA-binding domain. Evidence suggests that the N-terminal domain may suppress the activity of the C-terminal domain. This gene plays a role in the nervous systems of lung diseases, hepatitis, as well as oral diseases, and cervical cancer, and its expression has been seen in malignant tissues (8-12). Figure 1 (B) shows the location of this gene on the chromosome.

TET1 gene

The TET1 gene with accession number NM_030625 is located on chromosome 10 (10q21.3). The number of amino acids, molecular weight, and isoelectric point of the number of exons were determined in Table 1 for the TET1 gene. The TET1 gene is also known as LCX and KIAA1676. And they play a role in epigenetic processes such as DNA methylation and gene expression. This gene belongs to the Tet methylcytosine dioxygenase family and plays an effective role in the activation of other genes. Diseases associated with TET1 include Glioblastoma and Rett Syndrome and cervical cancer. The important paralog of this gene is TET2 (13-17). Figure 1(C) shows the location of this gene on the chromosome.

Materials and Methods

First, the sequences of LRP11 (NM_032832), FUBP1 (NM_003902), and TET1 (NM_030625) genes were obtained from the NCBI database. The length of these proteins was 500, 644, and 2136 amino acids, respectively. Then the exact location of these genes was determined using the UCSC database. The 3D structure of proteins and Ramachandran plot were determined using the MBC database and the molecular weight and isoelectric point of proteins were determined using the ProtScale database (18-21). Then the cellular comparison of LRP11, FUBP1, and TET1 genes and the expression of these genes were analyzed by the Human Protein Atlas OMIM database.

Results

Ontology of LRP11, FUBP1 and TET1 genes

One of the most important molecular functions of the LRP11 gene is the binding of phosphoproteins. The effective biological processes of the LRP11 gene include response to heat, cold, hunger, mechanical stimuli and stress. In general, the LRP11 gene is an integral part of the cell membrane (22-23). The molecular functions of the FUBP1 gene include binding to nucleic acid, DNA, RNA, and pro-

Figure 1. A) Chromosome 6, the red area is where the LRP11 gene is located (6q25.1). B) Chromosome 1, the red area is where the FUBP1 gene is located (1p31.1). C) Chromosome 10, the red area is where the TET1 gene is located (10q21.3).

<table>
<thead>
<tr>
<th>Name</th>
<th>LRP11</th>
<th>FUBP1</th>
<th>TET1</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORGANISM</td>
<td>Homo sapiens (Human)</td>
<td>Homo sapiens (Human)</td>
<td>Homo sapiens (Human)</td>
</tr>
<tr>
<td>Accession number nucleotide</td>
<td>NM_032832</td>
<td>NM_003902</td>
<td>NM_030625</td>
</tr>
<tr>
<td>Accession number protein</td>
<td>NP_116221.3</td>
<td>NP_003893.2</td>
<td>NP_085128.2</td>
</tr>
<tr>
<td>Gene ID</td>
<td>84918</td>
<td>8880</td>
<td>80312</td>
</tr>
<tr>
<td>Chromosome</td>
<td>6</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Cytogenetic location</td>
<td>6q25.1</td>
<td>1p31.1</td>
<td>10q21.3</td>
</tr>
<tr>
<td>Chromosome location bp</td>
<td>149818757-149864359</td>
<td>77944055-77979072</td>
<td>68560337-68694487</td>
</tr>
<tr>
<td>nucleotide length</td>
<td>3634bp</td>
<td>6714bp</td>
<td>9612 bp</td>
</tr>
<tr>
<td>protein length</td>
<td>500aa</td>
<td>644aa</td>
<td>2136aa</td>
</tr>
<tr>
<td>Molecular weight (Da)</td>
<td>53311.35</td>
<td>67560.39</td>
<td>235308.68</td>
</tr>
<tr>
<td>Isoelectric point</td>
<td>6.04</td>
<td>7.18</td>
<td>8.53</td>
</tr>
<tr>
<td>Total Exon</td>
<td>7</td>
<td>20</td>
<td>12</td>
</tr>
</tbody>
</table>
tein and acting as a transcription factor. Biological processes related to the FUBP1 gene also include transcriptional regulation, transcription from RNA polymerase II promoter and positive regulation of gene expression. Also, the FUBP1 gene is a cellular component of the nucleus and nucleoplasm (24). The TET1 gene is also a cellular component of the nucleus. Its molecular functions include the activation of RNA polymerase II transcription factor, DNA binding to specific sequences, binding to metal ions such as iron and zinc, oxidoreductase activity, dioxygenase and methylcytosine dioxygenase. The effective biological process in the gene also includes chromatin organization and negative regulation and silencing of chromatin dependent on methylation, DNA demethylation, cell differentiation, transcription regulation, and glycosylation, positive regulation of cell proliferation and methylation of histones and maintenance of stem cell population (25-27).

Three-dimensional structure
Molecular homology modeling using the SWISS-MODEL database in Expasy led to the prediction of the three-dimensional structure of LRP11, FUBP1, and TET1 proteins based on sample 1at02 with the highest percentage of similarity (Figure 2). Then the proteins related to LRP11, FUBP1, and TET1 were determined (Figure 3). QMEAN-DisCo Global and GMQE numbers showed the overall quality of the model, which usually varies between zero and one, and higher numbers indicate high model quality. The value of QMEAN-DisCo Global and GMQE for the LRP11 gene is equal to 0.45 ± 0.05 and 0.27, for the FUBP1 gene it is equal to 0.71 ± 0.07 and 0.21, and for the TET1 gene, it is equal to 0.64 ± 0.05 and 0.09. The estimation of protein quality was determined based on the QMEAN z-scores scale. According to the QMEAN Z-scores, it was found that there is a good match between the model structure and the experimental structures of the same size.

Figures 4, 5, and 6 show the different QMEAN z-scores, including QMEAN, Cβ interactions, all-atom interactions, solvation, and torsion for the LRP11, FUBP1, and TET1 genes. Then the Ramachandran diagram related to LRP11, FUBP1, and TET1 proteins was determined to determine the energy level and stability in terms of two angles φ and ψ in the proteins (Figure 7). Considering that in LRP11 protein, the percentage of Ramachandran's favorite amino acids was 77.10%, in FUBP1 protein, the percentage of Ramachandran's favorite amino acids was 93.45%, and in TET1 protein, the percentage of Ramachandran's favorite amino acids was 84.06%. Therefore, the proposed model is suitable for the three-dimensional structure of proteins.

Investigation of molecular domains in LRP11, FUBP1 and TET1 genes
The LRP11 gene has 3 protein domains. The first domain is MANEC, which is between 84 and 184 amino acids and has an E-value of 6.74e-44. The second domain is PKD, located between 215 and 299 amino acids with an E-value=1.72e-03. The third domain is LDLa, which is between 309 and 346 amino acids with an E-value=1.09e-05.

FUBP1 gene contains 4 KH domains, located in the distances of 99-169, 184-256, 274-344, and 375-448, respectively, E-value1=1.09e-17, E-value2=2.33e-17, E-value3=1.86e-16 and E-value4=1.019e-14. The second protein and acting as a transcription factor. Biological processes related to the FUBP1 gene also include transcriptional regulation, transcription from RNA polymerase II promoter and positive regulation of gene expression. Also, the FUBP1 gene is a cellular component of the nucleus and nucleoplasm (24). The TET1 gene is also a cellular component of the nucleus. Its molecular functions include the activation of RNA polymerase II transcription factor, DNA binding to specific sequences, binding to metal ions such as iron and zinc, oxidoreductase activity, dioxygenase and methylcytosine dioxygenase. The effective biological process in the gene also includes chromatin organization and negative regulation and silencing of chromatin dependent on methylation, DNA demethylation, cell differentiation, transcription regulation, and glycosylation, positive regulation of cell proliferation and methylation of histones and maintenance of stem cell population (25-27).

Figure 2. A) Three-dimensional structure of LRP11 protein. B) Three-dimensional structure of FUBP1 protein. C) Three-dimensional structure of TET1 protein.

Figure 3. A) Interacting proteins for LRP11 Gene. B) Interacting proteins for FUBP1 Gene. C) Interacting proteins for TET1 Gene.
domain is called DUF1897, which is located between 572-599 and 602-626, E-value DUF1897 (1)= 2.80e-08, E-value DUF1897 (2)= 2.70e-10. 

**TET1** gene has a zf-CXXC domain located between 583 and 624 and E-value=3.70e-11 (Figure 8).

The results of examining the expression of LRP11, FUBP1, and TET1 genes in different organs

Figures 9, 10, and 11 show the expression of LRP11, FUBP1, and TET1 genes in different organs. According to the graphs, it can be seen that the expression of the LRP11 gene is the highest in the organs of the brain, prostate, and thyroid gland, and the lowest in the bone marrow. In the FUBP1 gene, the highest level of expression was observed in the lymph node and the lowest level of expression was observed in the pancreas and liver. In the case of the TET1 gene, the highest level of expression was observed in the placenta and the lowest level of similar expression of the FUBP1 gene was observed in the pancreas and liver.

**Discussion**

Cervical cancer, as one of the challenges facing women's health and treatment, is of particular importance. The molecular and bioinformatics technique has provided the possibility of studying different levels of information such as the genome, transcriptome, proteome, and metabolome with high depth and accuracy, which provides comprehensive and valuable information about the investigated phenotypes, including complex diseases such as cancer. Comprehensive databases such as GEO and NCBI contain bioinformatics data for different phenotypes and organisms. Extracting valuable and comprehensive information also requires the use of appropriate and accurate methods. Since the genes examined in this research are closely related to cervical cancer, it is very necessary to examine the expression of these genes in different organs and lead to the prediction and improvement of the field of medicine and treatment, especially in cancer (28). In gene-expression analysis of *LRP11*, *FUBP1*, and *TET1* genes showed that the expression level of these genes is high in lymph nodes and thyroid, and lower in the pancreas and liver. Also, the expression level of the *FUBP1* gene is higher than that of *LRP11*, and the expression level of the *LRP11* gene is higher than that of *TET1*. Regarding the structure and proteomics of the studied genes, it can be seen that due to the presence of more domains in the *LRP11* and *FUBP1* genes, these genes probably independently participate in various functions and have a wider range of activity than the *TET1* gene. Also, the matching of QMEANDisCo Global and GMQE figures showed that the stability of the FUBP1 gene is relatively higher than the *TET1* gene, and the *TET1* gene is more stable than the *LRP11* gene. These results were completely consistent with the data obtained from the Ramachandran chart. Considering that these genes are effective key genes for the early detection of cervical cancer and there is much evidence that they are effective in suppressing tumors related to cervical cancer, they can be used as markers in the diagnosis and treatment of cervical cancer.

**References**


