

In silico analysis of fragile histidine triad involved in regression of carcinoma

Muhammad Asif Rasheed, Fatima Tariq, Sara Afzal, Shazia Mannan

Abstract

Hepatocellular carcinoma (HCCa) is a primary malignancy of the liver. Many different proteins are involved in HCCa including insulin growth factor (IGF) II, signal transducers and activators of transcription (STAT) 3, STAT4, mothers against decapentaplegic homolog 4 (SMAD 4), fragile histidine triad (FHIT) and selective internal radiation therapy (SIRT) etc. The present study is based on the bioinformatics analysis of FHIT protein in order to understand the proteomics aspect and improvement of the diagnosis of the disease based on the protein. Different information related to protein were gathered from different databases, including National Centre for Biotechnology Information (NCBI) Gene, Protein and Online Mendelian Inheritance in Man (OMIM) databases, Uniprot database, String database and Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Moreover, the structure of the protein and evaluation of the quality of the structure were included from Easy modeler programme. Hence, this analysis not only helped to gather information related to the protein at one place, but also analysed the structure and quality of the protein to conclude that the protein has a role in carcinoma.

Keywords: Hepatocellular carcinoma, Fragile histidine triad, FHIT, Proteomics, Bioinformatics.

Introduction

In our previous study,¹ we presented the analysis of mothers against decapentaplegic homolog 4 (SMAD4) protein, while the current study is based on the analysis of fragile histidine triad (FHIT) protein involved in hepatocellular carcinoma. The proteomics studies are helpful in understanding the cell processes, and the effect of proteins on the cell processes. Moreover, it can also be helpful in studying the effect of the environment and the cell processes on the proteins. Hence, through applying the proteomics tools and identifying the proteins in a disease and the healthy individuals' samples, we may identify the biomarkers to differentiate both the classes.

.....
Department of Biosciences, COMSATS Institute of Information Technology, Sahiwal, Pakistan.

Correspondence: Muhammad Asif Rasheed.

Email: asif.rasheed@ciitsahiwal.edu.pk

Such discoveries may also lead to protein-based diagnostic tools and better understanding of a disease state. For example, protein expression pattern or profile during cancer is different compared to a normal healthy situation. Such unique proteins which are present in the diseased condition and not in the healthy cells can be used as markers for the disease study as well as target during the disease treatment. The analysis of the proteins using a proteomics approach is very useful and has different applications. For example, proteins isolated from body tissues gives a complete idea of the tissue situation as well as provided a good basis to study other processes such as post-translational modifications, protein functionality or protein complexes.²

Bioinformatics introduce new algorithms in the field of proteomics to handle large and heterogeneous data sets. With the integrated use of "-omics" disciplines, the identification and the characterisation of candidate genes, proteins and molecules involved in a given disease will probably represent one of the milestones of future healthcare.³ Proteomics deals with the identification of the proteins produced by cells in normal and diseased conditions, while metabolomics monitors the role of small molecules (lipids, sugars and amino acids) involved in daily cellular function.⁴ There are approximately 30,000 genes in the human genome and the number of proteins is likely at least three times higher, as resulting from alternative splicing and post-translational modifications.⁵

Hepatocellular Carcinoma

Hepatocellular carcinoma (HCCa) is a primary malignancy of the liver. Most cases of HCCa are secondary to either a viral hepatite infection (hepatitis B or C) or cirrhosis (alcoholism) which is the most common cause of hepatic cirrhosis.⁶ In countries where hepatitis is not endemic, most malignant cancers in the liver are not primary HCCa but metastasis of cancer from elsewhere in the body, e.g., the colon. HCCa is the fifth most common cancer and the third most common cause of cancer death worldwide. There were few effective therapeutic options available for those suffering from advanced disease⁷ but progress has been seen in surgical treatment of hepatocellular carcinoma.⁸ Moreover, many immune-based approaches have shown efficacy in achieving disease regression and representing the most promising new treatment

approach.⁹ Administration of levocarnitine and/or branched chain amino acids during invasive treatments reduced blood ammonia (NH₃) concentration and suppressed the albumen.¹⁰ HCCa is reported to be the second most common cause of non-islet cell tumour hypoglycaemia (NICTH) which was controlled by using systemic chemotherapy.¹¹ This type of cancer occurs more often in men than women and is usually seen in people with age of 50 or older.¹² However, age varies in different parts of the world. Moreover, different risk factors of HCCa include hepatitis B virus (HBV) and alcohol intake, passive smoking, indoor air pollution and pesticide exposure. Moreover, fruit and tea intake may significantly lower the risk of HCCs.¹³ Furthermore, different factors influencing the prognosis of patients with HCCa were explored.¹⁴ Hence in order to prevent HCCa, clinicians should provide health education to overcome such risk factors and encourage HBV and hepatitis C virus (HCV) carriers to undergo annual physical examination and receive adequate treatment.¹⁵ Specifically, hepatitis B should be treated in order to prevent HCCa.¹⁶

Different up-regulated proteins during the disease include insulin growth factor (IGF) II, a disintegrin and metalloproteases (ADAM) 9, signal transducers and activators of transcription (STAT) 3, suppressors of cytokine signalling (SOCS) 3, and cyclin D1 while the down-regulated proteins during the disease include collagen I, SMAD 4, fragile histidine triad (FHIT), and SOCS1.¹⁷ Other proteins include, POU class 5 homeobox 1 (OCT4), baculoviral IAP repeat containing 5 (BIRC5), cyclin D1 (CCND1), ATP binding cassette subfamily G member 2 (BCRP), SRY-box 2 (Sox2), Glutathione S-transferases (GST), NCK adaptor protein 1 (NCK1), human leukocyte antigen DQ (HLA-DQ), miR-106b, c-Myc, Ki67 and selective internal radiation therapy (SIRT).

Analysis of FHIT by Different Databases and Software Tools

The bioinformatics analysis of FHIT protein was performed by using different tools, software and databases related to bioinformatics. Different details related to the proteins were gathered from different databases and software including the National Centre of Biotechnology Information (NCBI) gene, protein and Online Mendelian Inheritance in Man (OMIM) databases and UniProt database. Basic Local Alignment Search Tool (BLAST) was used to get the closest proteins related to FHIT in protein databank (PDB) database. Moreover, Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway database was used to show that the protein is involved in tumour suppression. Furthermore, String database was used to show the interacting proteins with FHIT while

easy modeller programme was used to model the protein structure and to draw Ramachandran plot.

FASTA Sequence of the Protein

The FASTA sequence of FHIT protein consists of 147 amino acids taken from NCBI protein database, which is: >gi|21595364|gb|AAH32336.1| FHIT protein [Homo sapiens]

```
MSFRFGQHLIKPSVFLKTELSFALVNRKPVVPGHVLVCLRPV
ERFHDLRPDEVADLFQTTQRVGTVEKHFHGTSLTFMQDGP
EAGQTVKHVHVHVLPRKAGDFHRNDSIYEELQKHKEDFPAS
WRSEEMAAEAAAALRVYFQ18
```

Chromosome Location

The protein is located at chromosome 3 and the exact location of the protein is 3p14.2. Moreover, the molecular weight of the protein is 16858 Da. Different domains of FHIT protein include Histidine triad (HIT) domain, HIT-like domain and Histidine triad conserved site domain.¹⁹

Table: Detail of the top eight closest proteins compared to the query FASTA sequence of fragile histidine triad (FHIT) protein in protein databank (PDB) database using Basic Local Alignment Search Tool (BLAST) tool.

Serial number	Accession	Query coverage	E value	Identity
1	1FHIA	100%	3e-106	100%
2	2FHIA	100%	2e-105	99%
3	3FITA	100%	5e-104	99%
4	1EMSA	96%	3e-33	49%
5	4Q61A	59%	2e-04	34%
6	3ANOA	63%	6e-11	34%
7	2C82A	40%	0.86	33%
8	3ZIOA	40%	1.1	33%

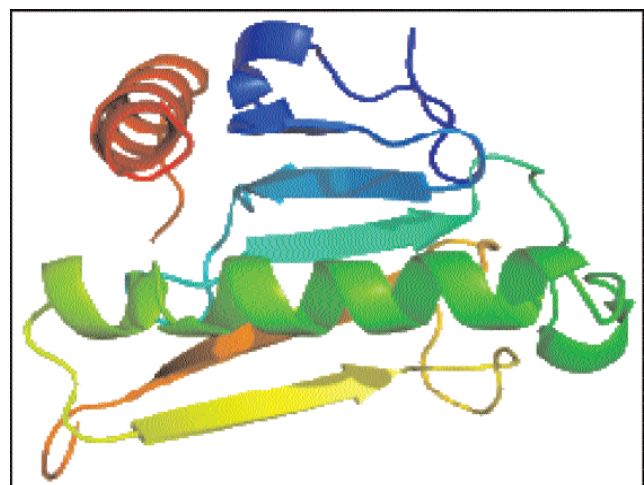


Figure-1: Three-dimensional protein structure of fragile histidine triad (FHIT) showing alpha helices and beta sheets in the protein. The protein structure was generated using 1FHIA as template by easy modeller software.^{34,35}

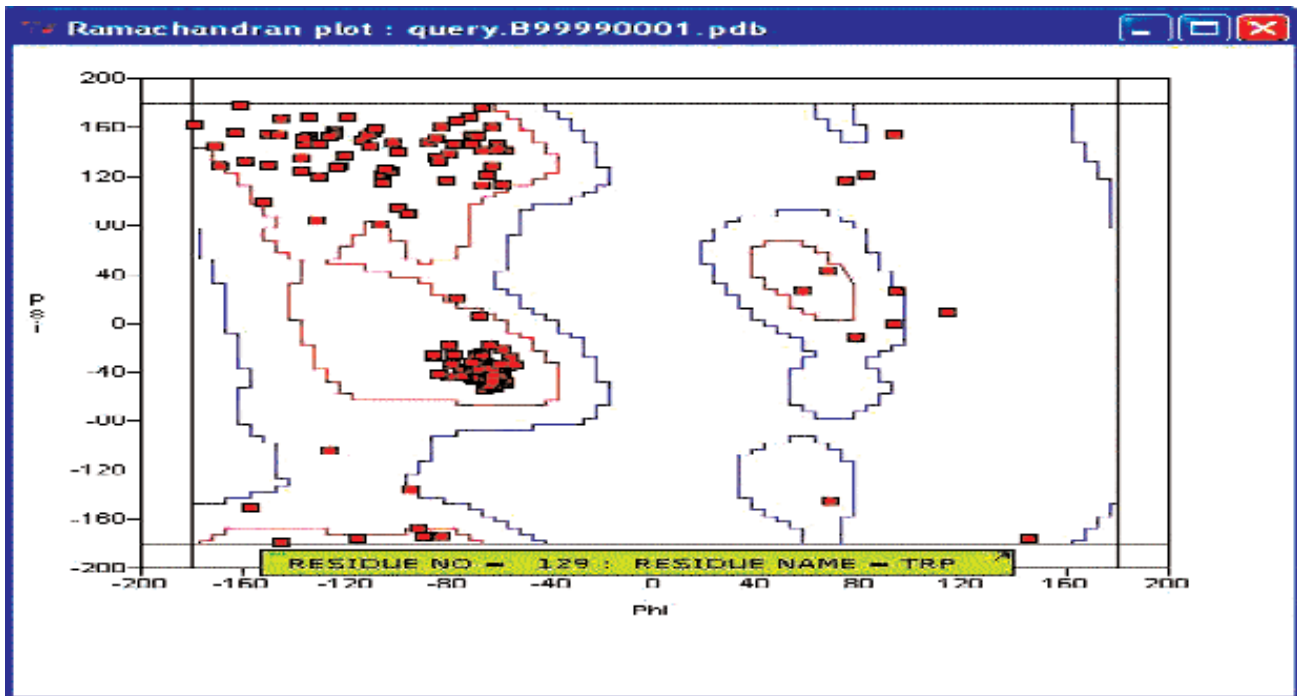


Figure-2: Ramachandran plot to check the protein's quality for fragile histidine triad (FHIT). The red dots are showing protein residues. Most of the residues are present in most favourable regions suggesting a good quality protein structure. The plot was generated by easy modeller software.³⁵

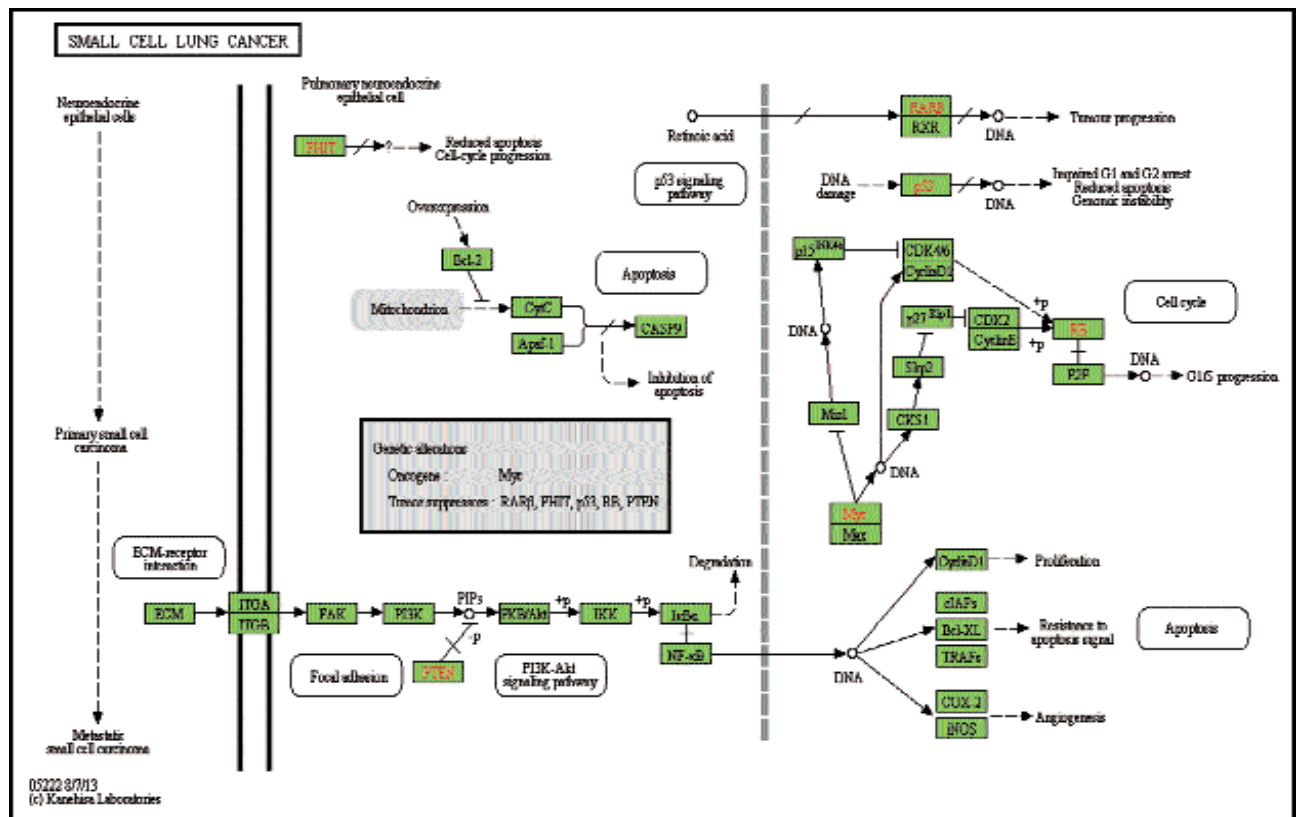


Figure-3: Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway for the protein fragile histidine triad (FHIT). It is showing FHIT as tumour suppressor protein and unavailability of the protein leads to reduced apoptosis and cell-cycle progression.³⁶

BLAST was used to collect eight closest sequences compared to the query FASTA sequence of FHIT protein in PDB database. The PDB codes of selected proteins with other details were noted (Table). Hence, the template chosen to predict the structure of the protein was 1FHIA. This template was predicted by Pace H.C. et al in 1998 by X-ray diffraction method with the resolution of 3.10 Å.²⁰ The structure of FHIT was predicted by using easy modeller software (Figure-1), while the quality of the predicted structure was checked by Ramachandran plot (Figure-2). There were three most favourable regions in the plot and most of the residues were present in most favourable regions, suggesting a

good quality protein structure.

Role of FHIT in other Carcinoma

Among 50% of oesophageal, stomach and colon carcinomas, inconsistent FHIT transcripts were identified.²¹ Moreover, loss of FHIT promoted carcinogens in human bronchial epithelial cells.²² Furthermore, in a series of small cell lung cancers (SCLCs) and non-small cell (NSCLC) types, FHIT gene structure and transcription were also analysed.²³ The study was performed by reverse transcription polymerase chain reaction (RT-PCR) where in 11 of 14 SCLC tumours, abnormal-sized transcripts were found. In 9 of these cases, both normal and abnormal

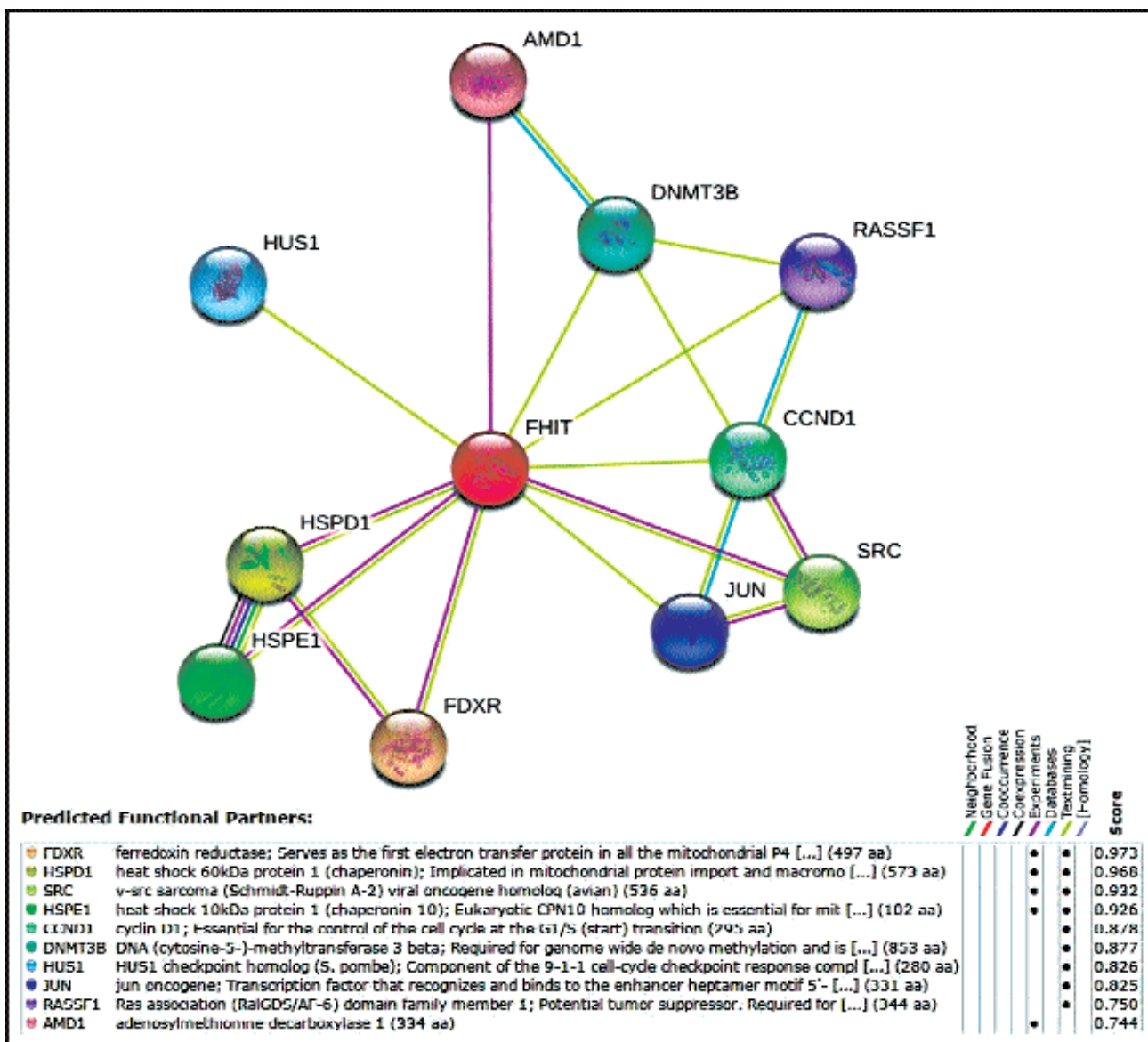


Figure-4: Fragile histidine triad (FHIT) interaction with other proteins. The highest interaction with score 0.973 is found with ferredoxin reductase (FDXR) protein. The image and the related details were retrieved from STRING database.³⁷

sized transcripts were present. Moreover, abnormal transcripts were found in 18 of 25 NSCLC tumours. One or 2 abnormal sized bands which were accompanied by a normal sized transcript suggested the presence of normal cells within the tumours. These researchers reported loss of heterozygosity for microsatellite markers internal to and flanking the FHIT locus. Furthermore, 11 of 12 tumours that exhibited abnormal FHIT transcripts showed allelic loss at 1 or more of the loci. Hence, the researchers suggested the inactivation of the FHIT gene by a mechanism of loss of 1 allele and altered expression of the remaining alleles in these tumours. They further concluded that accumulation of high levels of intracellular diadenosine tetraphosphate and the stimulation of deoxyribonucleic acid (DNA) synthesis and proliferation resulted in the loss of function of the FHIT (Figure-3) which may occur as a consequence of physical, chemical, and biological agents. FHIT was associated with breast cancer,²⁴ thyroid tumours,²⁵ lung cancer,^{26,27} and acute lymphoblastic leukaemia.²⁸ Hence, the tumour suppressor gene proteins (FHIT and others) expression can be one of the factors influencing prognosis and valuable for clinical treatment.²⁹

String is a known and predicted protein-protein interaction database which shows the highest interaction of FHIT with ferredoxin reductase (FDXR) (Figure-4) and it has also been detected by in vitro assay.³⁰

Since the discovery of FHIT gene in 1996, more than 350 studies have been published.³¹ Hence, FHIT is altered in many human tumours, particularly in those caused by environmental carcinogens, such as those present in tobacco smoke.³² In many of these tumours, particularly in those induced by tobacco or other environmental carcinogens, alterations of FHIT occur very early during the multistep process of carcinogenesis. Infection with FHIT recombinant viruses may cause regression, and prevention of tumours in experimental animals showed that FHIT-negative cancer cells are very sensitive to the expression of FHIT.³³ Thus, it is logical to predict the development of a gene therapy approach for the treatment and prevention of FHIT-negative human cancers. Moreover, the analysis of the protein using a proteomics approach is very useful and may have different applications. Bioinformatics introducing new algorithms in the field of proteomics to handle large and heterogeneous data sets and bioinformatics analysis of FHIT protein involved in different carcinomas may lead to better diagnosis as well as the treatment of the disease.

Conclusion

This review paper presented a summary of findings by a

number of studies regarding significant role of FHIT in carcinogenesis. It is clear that the protein has a role in tumour regression. Moreover, bioinformatics has really made the proteomics research easy and efficient by introducing new algorithms in the field of proteomics to handle large and heterogeneous data sets. Proteins and molecules involved in a given disease will probably represent one of the milestones of future healthcare.

Disclaimer: None.

Conflicts of Interest: None.

Source of Funding: None.

References

1. Rasheed MA, Afzal S, Tariq F, Mannan S. In silico analysis of SMAD4 involved in hepatocellular carcinoma. *Asian J Agri Biol* 2014; 2: 28-33.
2. Görg A, Weiss W, Dunn MJ. Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 2004; 4: 3665-85.
3. Tanke HJ. Genomics and proteomics: The potential role of oral diagnostics. *Ann N Y Acad Sci* 2007; 1098, 330-4.
4. Garcia I, Tabak LA. Beyond the "omics": Translating science into improved health. *J Am Dent Assoc* 2008; 139, 392-5.
5. Wright JT, Hart TC. The genome projects: Implications for dental practice and education. *J Dent Educ* 2002; 66, 659-71.
6. Kumar V, Fausto N, Abbas A, (editors). Robbins & Cotran Pathologic Basis of Disease. 7th ed. Philadelphia: Saunders, 2003; 914-7.
7. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005; 55: 74-108.
8. Li KY, Liu LX, Yin DL. Progress in surgical treatment of hepatocellular carcinoma. *Zhonghua Wai Ke Za Zhi* 2016; 54: 148-52.
9. Liu D, Staveley-O K. Immune-based Therapy Clinical Trials in Hepatocellular Carcinoma. *J Clin Cell Immunol* 2015; 6: 376.
10. Iwasa M, Sugimoto R, Ishihara T, Sekoguchi-Fujikawa N, Yoshikawa K, Mifuji-Moroka R, et al. Usefulness of Levocarnitine and/or Branched-Chain Amino Acids during Invasive Treatment for Hepatocellular Carcinoma. *J Nutr Sci Vitaminol* 2015; 61: 433-40.
11. Huang J, Chang P. Refractory hypoglycemia controlled by systemic chemotherapy with advanced hepatocellular carcinoma: A case report. *Oncol Lett* 2016; 11: 898-900.
12. Parkin DM, Ohshima H, Srivatanukul P, Vatanasapt V. Cholangiocarcinoma: epidemiology, mechanisms of carcinogenesis and prevention. *Cancer Epidemiol Biomarkers Prev* 1993; 2: 537-44.
13. Niu J, Lin Y, Guo Z, Niu M, Su C. The Epidemiological Investigation on the Risk Factors of Hepatocellular Carcinoma. *Medicine (Baltimore)* 2016; 95: e2758.
14. Rong WQ, Yu WW, Wu JX, Wu F, Wang LM, Tian F, et al. Analysis of prognostic factors in patients with hepatocellular carcinoma (≥5 cm) underwent hepatectomy. *Zhonghua Wai Ke Za Zhi* 2016; 54: 89-93.
15. Jane S, Lin M, Chiu W, Lai L, Chen P, Chen M. Early detection of unhealthy behaviors, the prevalence and receipt of antiviral treatment for disabled adult hepatitis B and C carriers. *BMC Public Health* 2016; 16: 146.
16. Li YW, Yang FC, Lu HQ, Zhang JS. Hepatocellular carcinoma and hepatitis B surface protein. *World J Gastroenterol* 2016; 22: 1943-52.
17. Tannapfel A, Anhalt K, Hausermann P, Sommerer F, Benicke M, Uhlmann D, et al. Identification of novel proteins associated with hepatocellular carcinomas using protein microarrays. *J Pathol* 2003; 201: 238-49.
18. <https://www.ncbi.nlm.nih.gov/protein/21595364> First accessed in

- March-April, 2013 [title missing]
19. <http://www.uniprot.org/uniprot/P49789> First accessed in March-April, 2013. [title missing]
 20. Pace HC, Garrison PN, Robinson AK, Barnes LD, Draganescu A, Rösler A, et al. Genetic, biochemical, and crystallographic characterization of Fhit-substrate complexes as the active signaling form of Fhit. *Proc Natl Acad Sci USA* 1998; 95: 5484-9.
 21. Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, et al. The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell* 1996; 84: 587-97.
 22. Boylston J, Brenner C. A knockdown with smoke model reveals FHIT as a repressor of Heme oxygenase 1. *Cell Cycle* 2014; 13: 2913-30.
 23. Sozzi G, Veronese ML, Negrini M, Baffa R, Cotticelli MG, Inoue H, et al. The FHIT gene 3p14.2 is abnormal in lung cancer. *Cell* 1996; 85: 17-26.
 24. Zaki S, Abdel-Azeez H, El Nagar M, Metwally K, Ahmed M. Analysis of FHIT gene methylation in egyptian breast cancer women: association with clinicopathological features. *Asian Pac J Cancer Prevent* 2015; 16: 1235-9.
 25. Koc M, Aktimur R, Kagan Gokakin A, Atabey M, Koyuncu A, Elagoz S, et al. Expression of FHIT, p16, p53 and EGFR as prognostic markers in thyroid tumors of uncertain malignant potential. *J BUON* 2015; 20: 567-72.
 26. Yu Y, Liu X, Yang Y, Zhao X, Xue J, Zhang W, et al. Effect of FHIT loss and p53 mutation on HPV-infected lung carcinoma development. *Oncol Lett* 2015; 10: 392-8.
 27. Wu D, Hsu N, Wang Y, Lee M, Cheng Y, Chen C, et al. c-Myc suppresses microRNA-29b to promote tumor aggressiveness and poor outcomes in non-small cell lung cancer by targeting FHIT. *Oncogene* 2014; 34: 2072-82.
 28. Malak CA, Elghanam DM, Elbossaty WF. FHIT Gene Expression in Acute Lymphoblastic Leukemia and its Clinical Significance. *Asian Pac J Cancer Prev* 2015; 16: 8197-201.
 29. Chen Y, Wang X, Li F, Zhang L, Ma L, Liu Y. Relationship between expression of P27, Fragile Histidine Triad (FHIT), phosphatase and tensin homolog deleted on chromosome ten (PTEN), P73, and prognosis in esophageal squamous cell carcinoma. *Ann Diagn Pathol* 2015; 19: 33-6.
 30. Trapasso F, Pichiorri F, Gaspari M, Palumbo T, Aqeilan RI, Gaudio E, et al. Fhit interaction with ferredoxin reductase triggers generation of reactive oxygen species and apoptosis of cancer cells. *J Biol Chem* 2008; 283: 13736-44.
 31. Huebner K, Croce CM. Cancer and the FRA3B/FHIT fragile locus: it's a HIT. *Br J Cancer* 2003; 88: 1501-6.
 32. Izzotti A, Pulliero A. Molecular damage and lung tumors in cigarette smoke-exposed mice. *Ann NY Acad Sci* 2015; 1340: 75-83
 33. Huebner K, Croce CM. FRA3B and other common fragile sites: the weakest links. *Nat Rev Cancer* 2001; 1: 214-21.
 34. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H et al. The Protein Data Bank. *Nucleic Acids Res.* 2000; 28: 235-42. (Cited 2013 March-April). Available from URL: <http://www.rcsb.org/pdb/explore/explore.do?structureId=1FHI>.
 35. Easy modeller software. (Cited 2013 March-April). Available from URL: <http://modellergui.blogspot.sg/>.
 36. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 2012; 40(Database issue): D109-14. (Cited 2013 March-April). Available from URL: http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa05222&keyword=FHIT.
 37. STRING database. (Cited 2013 March-April). Available from URL: <http://www.string-db.org/>.
-