Original article

Gene co-expression analysis and Network biology studies in Indian population reveals functional similarities between Gastric cancer and other metabolic disorders

Blessantoli Mohandas[1], J. Jannet Vennila[2], Nikhil Ruban[3]

Abstract:

Objective: Gastric cancer (GC) is a multifactorial disease and known to have been associated with metabolic disorders. Gene expression profiling among various GC populations will help to strategize diagnosis and treatment. The current study employed microarray data analysis (MDA) and network biology methods to understand the significant genes in a GC Indian population and its association with other metabolic disorders. ***Materials and Methods:*** The microarray datasets of GC Indian population (Bangalore) was retrieved from Gene Expression Omnibus (GEO), normalized and analyzed using GeneSpring. With the fold change of ± 1.5, the differentially expressed genes (DEG) were identified. An interactome was built to study interactions and generate gene clusters. Statistical (centrality) parameters were applied to identify highly connected clusters followed by functional enrichment to identify significant pathways associated with the GC genes. ***Results and Discussion***: MDA identified 7181 DEGs (3984 up regulated genes and 3197 down regulated) and the interactome yielded 16552 interactions and two sub clusters. Cluster 1 was found to be statistically fit. The functional characteristics of the significant genes in this cluster revealed their association with adrenal cortex hormone insufficiency, thyroid disorders and deficiencies in kidney water resorption. ***Conclusion***: It is inferred from our study that, deficiency in Thyroid, Adrenal hormones and Antidiuretic hormone (ADH) functions has fair share in the prognosis and pathogenesis of GC Indian population. Henceforth, GC should not be viewed as separate entity in the series of cancers and gene expression profiling will help in improvising personalized medicine.

Keywords: Gastric Cancer; Network biology; Microarray Data Analysis; Metabolic disorders and Gastric Cancer; Gene expression profiling

Bangladesh Journal of Medical Science Vol. 21 No. 03 July'22 Page : 688-693
DOI: https://doi.org/10.3329/bjms.v21i3.59586

Introduction

Gastric cancer (GC) or Stomach cancer is the fifth most predominant cancer in the world, accounting with 7,23,000 cancer death rate per year and the third most prevalent cause of cancer-related mortality[1]. Approximately 50% of GC people live in East Asia, Southern Asia and China, primarily may be due to poor dietary patterns and *Helicobacter pylor*i infection[2]. In the early phase, GC is asymptomatic and therefore 80 to 90% of GC patients are diagnosed only in advanced stage. Defining new molecular genetic markers through gene expression profiling would help in better understanding of the GC prognosis. The GC gene expression pattern varies among populations depending on genetic makeup,

1. Blessantoli Mohandas, Department of Biochemistry, School of Agricultural and Biosciences, Karunya Institute of Technology & Sciences, Coimbatore-614114.
2. J. Jannet Vennila, MCC MRF Innovation Park, Madras Christian College, Chennai-600059.
3. Nikhil Ruban, Department of Humanities and Social sciences, Birla Institute of Technology and Sciences, Pilani Campus, Rajasthan-333 031.

Correspondence: Dr. J. Jannet Vennila, Director, Shift II and Coordinator, MCC MRF Innovation Park, Madras Christian College, Chennai-600059. Email- jannetvennila@gmail.com

dietary pattern and environmental influence. The genes involved in other metabolic pathways can also be deemed to be significant in GC pathogenesis and progression. Although records have shown 57,394 increasing new cases in India[1], there are only very few data available on gene expression profiling in GC Indian population. This study focused on systematic evaluation of expression and network analysis of significant genes in GC and maps their relevance to important molecular/ metabolic mechanism. This will help to strategize diagnosis, development and treatment process in GC Indian population.

Method

***Data selection*:** Microarray dataset of GC, Bangalore, India was retrieved (GSE 22804) from Gene Expression Omnibus (GEO). The dataset comprised of surgically resected GC samples and their paired disease-free peripheral nonmalignant tissue. Figure 1 gives the flowchart of the methodology adopted out in this study.

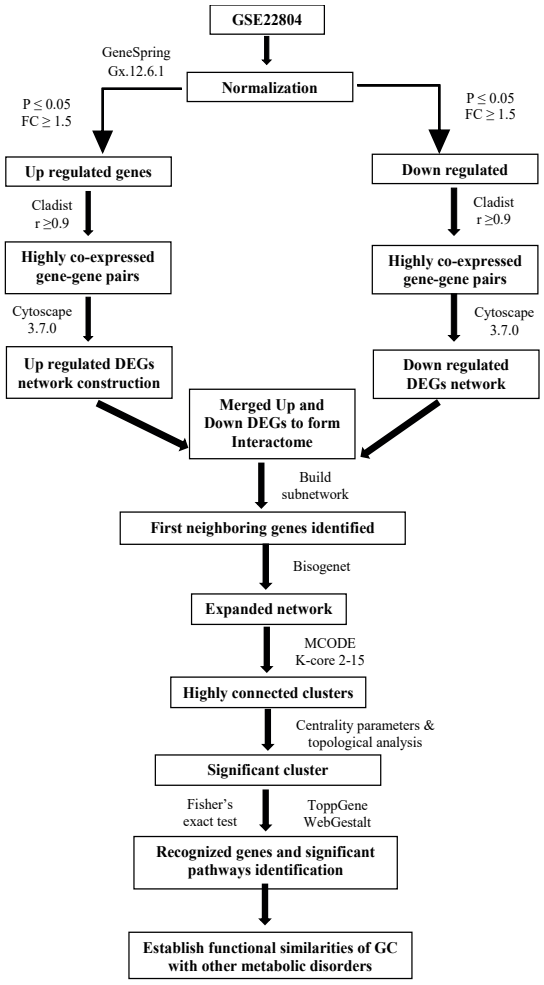

Figure 1: Schema charted in the study

Microarray data analysis (MDA): The Microarray Data Analysis (MDA) was carried out using GeneSpring GX 12.6.1, an Agilent microarray data analysis software[9]. Normalization of the dataset was performed using Percentile shift algorithm for Agilent platform array[10]. The differentially expressed genes (DEGs) were identified (Fold change of ± 1.5) and were categorized into 'up' and 'down' regulated based on fold change.

Network studies: The gene-gene correlation of all the DEGs was studied using Cladist[12]. The Interactome of the DEGs was built using Cytoscape 3.2.0 (http://cytoscape.org)[13]. With the help of GC marker genes reported in earlier studies (TP53[26], IL1B[27], IL1RN[28], VEGFA[29], TNF[30], CDH1[31], CLDN1[32], CLRN3[33], INHBA[34], and SULF1[35]), their first neighboring genes were selected in the network and a subnetwork was built. The interactions were further enriched using Bisogenet (Cytoscape plugin). Molecular Complex Detection (MCODE) was implemented to identify densely connected regions in GC networks (sub-clusters). K-core algorithm was used to identify the highly interconnected nodes (genes) in the network using MCODE. Centrality measures were used to predict the hub genes (highly connected genes) and to identify the node's functional gene. Using the Cytoscape plugin Network Analyzer, the topological and statistical understandings of each sub-clusters were calculated. Centrality parameters - betweeness centrality (BC), closeness centrality (CC), Cluster Coefficient (CLC), degree and topological centrality (TC) were considered to validate the prominence of biological network[16] to predict highly connected genes which are denoted as "significant genes". WebGestalt and ToppGene were used to examine the functional characteristics of the significant genes[23,24]. The pathways which represent many of the significant genes were studied to establish their correlation with other metabolic functions.

Ethical clearance: Not applicable

Results

***Microarray Data Analysis*:** GC dataset of Indian Population (GSE22804) contained 28 samples (14 GC patients and 14 normal of the same patient). Grouping of the dataset was done as 'normal' and 'diseased'. Based on the fold change cut off (±1.5) and p-value (0.05) the differentially expressed genes (DEGs) were selected. 3984 up regulated genes and 3197 down regulated genes were identified.

***Network construction and analysis*:** The above

DEGs were analyzed for gene-gene correlation which revealed additional 15,586,704 co-expressed genes for up regulated and 10,220,809 for down regulated genes respectively. Among these genes, 5611 upregulated and 1,92,156 downregulated genes were observed to be statistically correlated (r value ≥ 0.9). These genes were merged together to build the interactome. The network was extended with enrichment of additional interactions using Bisogenet. And the network enriched with 16552 edges (interactions).

***Identification of highly connected sub-clusters in interactome*:** K-core algorithm (MCODE) was used to identify the highly interconnected nodes (genes) in the network and two highly interconnected clusters were extracted from the interactome.

Topological and Centrality measures: Based on the topological and centrality measures, the cluster 1 was found to be statistically fit with R^2 value ≥ 0.9 and predicted to possess more number of functional and significantly relevant genes.

***Functional enrichment for the significant cluster*:** The functional enrichment of the cluster 1 genes was performed further for identifying the significant pathways associated with the GC genes (Table 1).

Table 1: Significant pathways associated with GC genes

Pathways	Total genes	Genes common to GC	P value
Thyroid hormone synthesis[b]	72	5	0.0016
Autoimmune thyroid disease[b]	47	3	0.0018
Proximal tubule bicarbonate reclamation[c]	23	3	0.0025
Toll-like receptor signaling pathway	99	4	0.0031
Adrenal cortex hormone insufficiency pathway[a]	147	5	0.0031
Fatty acid biosynthesis	12	2	0.0087
Bile secretion	68	4	0.0088
Vasopressin-regulated water reabsorption[c]	42	3	0.0137
Endocrine and other factor-regulated calcium reabsorption	46	3	0.0175
Cytosolic DNA-sensing pathway	57	3	0.0307

[a]Adrenal cortex hormone insufficiency disorders; [b]thyroid disorders; [c]deficiency in kidney water reabsorption. Statistical significance of P≤0.05

It can be inferred from Table 1, that genes associated with GC are also significant in pathways associated with Adrenal cortex hormone insufficiency (5 genes), thyroid disorders (8 genes) and deficiencies in kidney water reabsorption (6 genes). The possible role of these GC genes in interconnecting with the above metabolic disorders is diagrammatically explained in Figure 2.

APC-Adenomatous polyposis coli, SOD -Superoxide dismutase, ADH-Antidiuretic hormone, ATP1B4-ATPase Na+/K+ Transporting Family Member Beta 4, TSHR - Thyroid Stimulating Hormone Receptor, IFNA8- Interferon Alpha 8, SLC26A4- Solute Carrier Family 26 Member 4, GNAS- Guanine Nucleotide Binding Protein (G Protein) Alpha Stimulating Activity Polypeptide, PDE11A- Phosphodiesterase 11A, KMT2A- Lysine Methyltransferase 2A, WNT8B- Wnt Family Member 8B, AIPL1- Aryl Hydrocarbon Receptor Interacting Protein Like 1, DYNC2H1-Dynein Cytoplasmic 2 Heavy Chain 1. Red arrows represent up-regulation and Green arrows represent down-regulation

Discussion

Gastric cancer and metabolic disorders: Functional Similarities

Gastric cancer and thyroid disorders: The risk factors in GC are associated with thyroid disease and vice versa[7]. A study by Carvalho and Fighera[37] has shown that the prevalence of gastrointestinal diseases increased in patients with autoimmune thyroid dysfunction. Similarly, in autoimmune thyroid disorder celiac disease the atrophic gastritis are highly prevalent[38]. It is assumed that the development of certain gastric cancers involves dietary iodine which is also related to thyroid dysfunction[39]. The reason could be morphological and functional similarities of stomach and thyroid tissue which use the membrane active transport movement mechanism to concentrate iodides. And thyroid hormone is also a potent cofactor for tumor-suppressing genes[40].

***Gastric cancer and deficiency in Kidney water reabsorption*:** Inadequate antidiuretic hormone (ADH) leads to hyponatremia and excessive water reabsorption. Prolonged nausea and vomiting in hyponatremia may disturb the digestive system and induce GC[42].

***Gastric cancer and Adrenal cortex hormone insufficiency*:** Chabre and his research team have observed that the abnormal expression of gastric inhibitory polypeptide (GIP) receptor enables

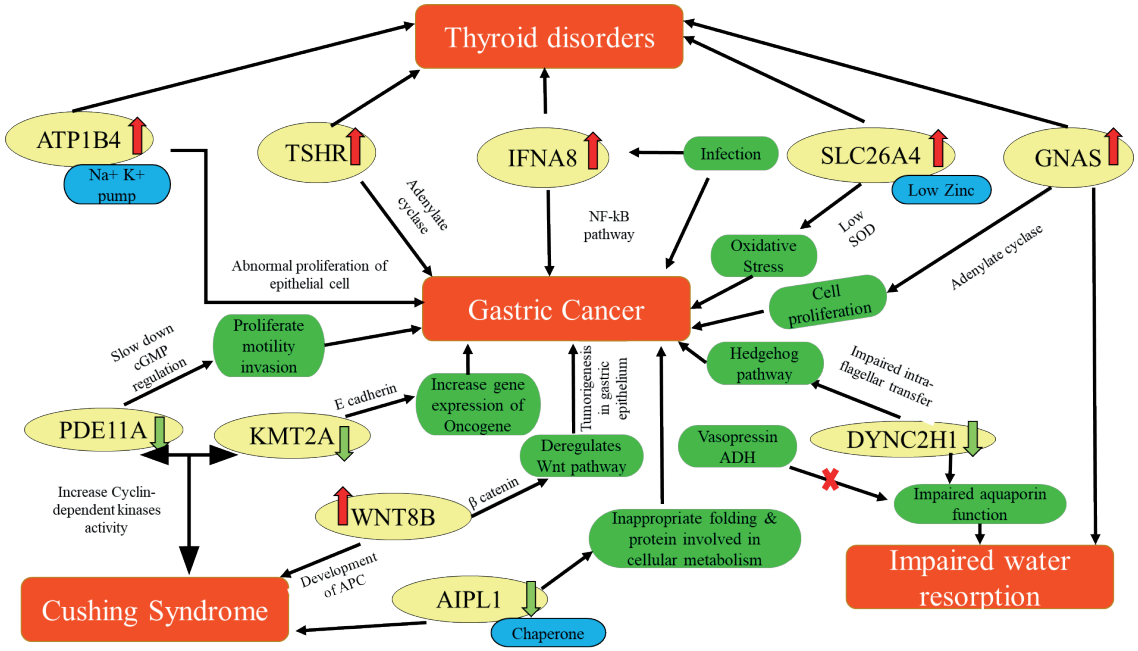

Figure 2: Possible interconnectivity of GC genes with other metabolic disorders

adrenocortical cells to react to food intake, with a cAMP increase in which both cortisol secretion and tumor proliferation can be stimulated[43]. GIP and Glucagon can regulate the gastric acid secretion in humans. GIP is also a potent releaser of gastric somatostatin, a secretion acid inhibitor *in vivo* and *in vitro*[44]. GIP receptor may also play a role in the development of tumor cells. The GIP effects were similar to the effects of ACTH on tumor cells.

Conclusions

Gastric cancer (GC) is a heterogeneous disease known to associate with environmental and genetic predisposing factors[45]. And from our network biology study in Indian population and reports on previous literatures, it was evident that, GC has close association with other metabolic disorders. Deficiency in Thyroid, Adrenal hormones and Antidiuretic hormone (ADH) functions has been found to have distinguished share in the prognosis and pathogenesis of GC. It is suggested that GC should not be viewed as separate entity in the series of cancers. GC may arise due to defective functions of the above said metabolic disorders. Therefore, a, multilevel screening of metabolic functions and gene expression profiling is suggested before the treatment of GC, could pave a better solution in treatment strategy. In addition, as genetic predisposition is an important feature in GC, the markers may not be universal for all the patients worldwide. Populations wise gene expression profiling can be more appropriate. It is recommended that GC gene expression profiling of different populations within India could still validate the above findings.

Conflict of interest statement: The authors declare that they have no conflict of interests.

Funding: Nil

Authors' contributions

Blessantoli Mohandas acquired data and analysed the data. J.Jannet Vennila designed the study and drafted the manuscript for important intellectual content. Nikhil Ruban assisted in manuscript preparation, editing and review. The manuscript was read and approved by all the authors.

Acknowledgments

The authors wish to thank the technical support received from the Department of Biotechnology, Karunya Institute of Technology & Sciences, Coimbatore, India

References:

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A and Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *International Journal of Cancer.* 2019; 144 (8): 1941-1953.
2. Rahman R, Asombang AW and Ibdah J.A. Characteristics of Gastric Cancer in Asia. *World Journal of Gastroenterology.* 2014; 20: 4483-4490.
3. Pirini F, Noazin S, Jahuira-Arias MH, Rodriguez-Torres S, Friess L, Michailidi C, Cok J, Combe J, Vargas, G and Prado W. Early detection of gastric cancer using global, genome-wide an IRF4, ELMO1, CLIP4 and MSC DNA methylation in endoscopic biopsies. *Oncotarget*. 2017; 8 : 38501-38516.
4. Kim TH, Choi SJ, Lee YH, Song GG and Ji JD. Gene expression profile predicting the response to anti-TNF treatment in patients with rheumatoid arthritis; analysis of GEO datasets. *Joint Bone Spine*. 2014; 81: 325-330.
5. Pawson T and Linding R. Network medicine. *FEBS letters.* 2008; 582:1266-1270.
6. Chand Y and Alam MA. Network biology approach for identifying key regulatory genes by expression based study of breast cancer. *Bioinformation.* 2012; 8:1132-1138.
7. Tornow S and Mewes, H Functional modules by relating protein interaction networks and gene expression. *Nucleic Acids Research.* 2003;31:6283-6289.
8. George G, Singh S, Lokappa SB and Varkey J. Gene co-expression network analysis for identifying genetic markers in Parkinson's disease-a three-way comparative approach. *Genomics*. 2019; 114 (4):819-830.
9. Chu L, Scharf E and Kondo T. GeneSpringTM: tools for analyzing microarray expression data. *Genome Informatics*. 2001;12:227-229.
10. Allison DB, Cui X, Pag, GP and Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nature reviews genetics.* 2006; 7: 55-65.
11. Zhao M, Chen, L, Liu Y and Qu H. GCGene: a gene resource for gastric cancer with literature. *Oncotarget*. 2016; 7:33983-33993.
12. Ng A, Bursteinas B, Gao Q, Mollison E and Zvelebil M. pSTIING: a 'systems' approach towards integrating signalling pathways, interaction and transcriptional regulatory networks in inflammation and cancer. *Nucleic acids research.* 2006; 34:527-534.
13. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003; 13: 2498-2504.
14. Bader G.D and Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics*. 2003; 4: 2-10.
15. Raman M.P, Singh S, Devi PR and Velmurugan D. Uncovering potential drug targets for tuberculosis using protein networks. *Bioinformation.* 2012; 8:403-406.
16. Assenov Y, Ramírez F, Schelhorn SE, Lengauer T and Albrecht M. Computing topological parameters of biological networks. *Bioinformatics.* 2007; 24: 282-284.
17. Yu H, Kim PM, Sprecher E, Trifonov V and Gerstein M. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS computational biology*, 2007; 3: e59.
18. Snijesh V and Singh S. Molecular modeling and network based approach in explaining the medicinal properties of *Nyctanthes arbortristis*, *Lippia nodiflora* for rheumatoid arthritis. *Journal of Bioinformatics and Intelligent Control*. 2014; 3:31-38.
19. Diestel R. The cycle space of an infinite graph. *Combinatorics, Probability and Computing*, 2005; 14: 59-79.
20. Azuaje FJ, Zhang L, Devaux Y and Wagner DR. Drug-target network in myocardial infarction reveals multiple side effects of unrelated drugs. *Scientific reports.* 2011; 1: 52-59.
21. Jeong H, Tombor B, Albert R, Oltvai ZN and Barabási AL. The large-scale organization of metabolic networks. *Nature*, 2000; 407: 651-654.
22. Lima-Mendez, G and van Helden J. The powerful law of the power law and other myths in network biology, *Molecular BioSystems.* 2009; 5:1482-1493.
23. Zhang B, Kirov S and Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic acids research.* 2005; 33: W741-W748.
24. Chen J, Bardes E.E, Aronow BJ and Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic acids research.* 2009; 37: W305-W311.
25. Frigge M, Hoaglin DC and Iglewicz. Some implementations of the boxplot. *The American Statistician* 1989; 43:50-54.
26. Graziano F, Cascinu S, Staccioli MP, Catalano V, Rossi MC, Baldelli AM, Giordani P, Muretto P. Catalano, Potential role and chronology of abnormal expression of the Deleted in Colon Cancer (DCC) and the p53 proteins in the development of gastric cancer. *BMC cancer*, 2001; 1: 9-13.
27. Kumar S, Kumar A and Dixit VK. Evidences showing

association of interleukin-1B polymorphisms with increased risk of gastric cancer in an Indian population. *Biochemical and biophysical research communications.* 2009; 387: 456-460.

28. Glas JR, Török HP, Schneider A, Brünnler GN, Kopp R, Albert ED, Stolte M and Folwaczny C. Allele 2 of the interleukin-1 receptor antagonist gene is associated with early gastric cancer. *Journal of clinical oncology.* 2004; 22: 4746-4752.
29. Feng CW, Wang LD, Jiao LH, Liu B, Zheng S. and Xie XJ. Expression of p53, inducible nitric oxide synthase and vascular endothelial growth factor in gastric precancerous and cancerous lesions: correlation with clinical features. *BMC cancer*. 2002; 2: 8-14.
30. El-Omar EM, Rabkin CS, Gammon MD, Vaughan TL, Risch HA, Schoenberg, JB, Stanford JL, Mayne ST, Goedert J and Blot WJ. Increased risk of noncardia gastric cancer associated with proinflammatory cytokine gene polymorphisms. *Gastroenterology.* 2003; 24: 1193-1201.
31. Kuraoka K, Oue N, Yokozaki H, Kitadai Y, Ito R, Nakayama H and Yasui W. Correlation of a single nucleotide polymorphism in the E-cadherin gene promoter with tumorigenesis and progression of gastric carcinoma in Japan. *International journal of oncology.* 2003; 23: 421-427.
32. Eftang LL, Esbensen Y, Tannæs TM, Blom GP, Bukholm IR and Bukholm G. Up-regulation of CLDN1 in gastric cancer is correlated with reduced survival. *BMC cancer.* 2013; 13: 586-592.
33. Marimuthu JH, Jakharia A, Subbannayya A, Keerthikumar Y, Kashyap S, Goel MK, Balakrishnan R, Dwivedi L, Pathare S, Dikshit S, Maharudraiah JB, Singh J, Sameer S, Kumar, Vijayakuma, GS, Veerendra Kumar M., Premalatha KV, Tata CS, Hariharan P, Rao, R, Prasad JC, Chaerkady TS, Kumar R. and Pandey A. Gene Expression Profiling of Gastric Cancer. *Journal of Proteomics & Bioinformatics.* 2011; 4:74-82.
34. Oshima T, Yoshihara K, Aoyama T, Hasegawa S, Sato T, Yamamoto N, Akito N Shiozawa, M, Yoshikawa T and Numata K. Relation of INHBA gene expression to outcomes in gastric cancer after curative surgery. *Anticancer research*. 2014; 34:2303-2309.
35. Li J, Mo ML, Chen Z, Yang J, Li QS, Wang DJ, Zhang H, Ye YJ, Xu JP and Li HL. HSulf-1 inhibits cell proliferation and invasion in human gastric cancer. *Cancer science*, 2011; 102: 1815-1821.
36. Kandemir E, Yonem A and Narin Y. Gastric carcinoma and thyroid status. *Journal of international medical research.* 2005; 33: 222-227.
37. Carvalho GAD and Fighera TM. Effect of gastrointestinal disorders in autoimmune thyroid diseases. *Translational Gastrointestinal Cancer.* 2014; 4: 76-82.
38. Rossi M, Negri E, Foschi R, Franceschi S and La Vecchia C. Relation between goiter and autoimmune thyroid disease, and gastric cancer. *International journal of cancer.* 2007; 120: 951-952.
39. Venturi S and Venturi M. Iodide, thyroid and stomach carcinogenesis: evolutionary story of a primitive antioxidant. *European journal of endocrinology*. 1999; 140: 371-372.
40. Guernsey DL and Leuthauser SW. Correlation of Thyroid Hormone Dose-dependent Regulation of K-ras Protoncogene Expression with Oncogene Activation by 3-Methylcholanthrene: Loss of Thyroidal Regulation in the Transformed Mouse Cell. *Cancer research*, 1987; 47: 3052-3056.
41. Wang CS, Lin KH and Hsu YC. Alterations of thyroid hormone receptor α gene: frequency and association with Nm23 protein expression and metastasis in gastric cancer. *Cancer letters.* 2002; 175: 121-127.
42. Hwang K, Jeon DH, Jang HN, Bae EJ, Lee JS, Cho HS, Chang SH. and Park DJ. Inappropriate antidiuretic hormone syndrome presenting as ectopic antidiuretic hormone-secreting gastric adenocarcinoma: a case report. *Journal of medical case reports.* 2014; 8: 185-191.
43. Chabre O, Liakos P, Vivier J, Chaffanjon P, Labat-Moleur FO, Martinie M, Bottari SP, Bachelot I, Chambaz EM and Defaye GV. Cushing's syndrome due to a gastric inhibitory polypeptide-dependent adrenal adenoma: insights into hormonal control of adrenocortical tumorigenesis. *The Journal of Clinical Endocrinology & Metabolism.* 1998; 83: 3134-3143.
44. Gespach C, Emami S and Rosselin G. Gastric inhibitory peptide (GIP), pancreatic glucagon and vasoactive intestinal peptide (VIP) are cAMP-inducing hormones in the human gastric cancer cell line HGT-1. Homologous desensitization of VIP receptor activity. *Biochemical and biophysical research communications*, 1984; 120: 641-649.
45. Yılmaz HT, Ergen S, Baydar, İdris, Çarlıoğlu A and Kurt A. Relatıonshıp between C-erb-B2 expressıon and other prognostıc factors in gastrıc cancer. *Bangladesh Journal of Medical Science*, 2018; 18(1): 124-129.