Research Journal of Recent Sciences _____ ISSN 2277-2502
Vol. 3**(ISC-2013),** 295-301 **(2014)** *Res. J. Recent. Sci.*

Gene Expression in Gastric Cancer for Singapore and UK Population: An *Insilico* Comparative Approach

Neethu G. Nair[1], Mohandhas Blessantoli[2] and J. Jannet Vennila[3]
[1,2]Department of Bioinformatics, Karunya University, Coimbatore , Tamilnadu, INDIA
[3]Department Biotechnology, Karunya University, Coimbatore, Tamilnadu, INDIA

Available online at: www.isca.in, www.isca.me
Received 20th November 2013, revised 10th January 2014, accepted 17th February 2014

Abstract

Gastric adenocarcinoma is a common cancer of the digestive tract in worldwide. It occurs in men above the age of 40. This cancer is common in Asia, parts of South America, and eastern and central Europe. And it accounts for 800,000 deaths worldwide per year. Prognosis is poor due to the late detection at its advanced stage. Gene expression analysis was done for Singapore and UK Gastric cancer datasets using GeneSpring. Differentially expressed Gastric cancer Stage 1 genes were identified by performing a t-test with a p-value cutoff of 0.005 and fold change ≥ 2.Only the differentially expressed genes that satisfied the criteria was enriched to find the functional insight of the stage 1 genes in Gastric cancer. The enriched genes were analyzed for their gene ontology using a Cytoscape plug-in Bingo. Molecular interactions among the stage 1 genes were done using a Cytoscape plug-in MiMI. hnRNPC gene showed a good interaction with its neighbouring genes. SNP regions in hnRNPC gene were identified and two of them were found to be located in the promoter region. hnRNPC gene interacts with Grb 2 gene in the progression of Gastric cancer. Majority of SNPs are located in the promoter regions for Gastric cancers from Singapore and USA, it can be suggested as one of the prime reasons for its up regulation in Gastric cancer. Thus hnRNPC be gene can be used as a potential marker in the early diagnosis of Gastric cancer in Homo sapiens.

Key words: Gastric Cancer, Singapore, UK, SNP, *hnRNPC.*.

Introduction

Gastric cancer or stomach cancer refers to a type of cancer that arises in any part of the stomach. Gastric Cancer is an asymptomatic type of cancer whose symptoms are nonspecific in its early stages. By the time symptoms occur, the cancer has often reached an advanced stage and may have also metastasized to different parts of the body, which is one of the main reasons for its relatively poor prognosis[1]. Gastric cancer causes about 8,00,000 deaths worldwide per year. It occurs most often in men over age 40. This form of Gastric cancer is very common in Japan, Chile and India. In 2010, the highest incidence rates for both sexes were in Eastern and Southern Asia and the lowest were in Northern and Southern Africa[2]. Compared to other countries Gastric cancer incidence rate is low in India. But in Chennai it is in higher rate[3]. A prospective case control study from Trivandrum evaluated dietary risk factors for Gastric Cancer and found that high consumption of rice, spicy food, chilly and high-temperature food increased the risk of developing Gastric cancer[4]. Gastric cancer is an important public health issue in worldwide. But due to asymptomatic and nonspecific symptoms in its early stage detection becomes a difficult task. Expression of Gastric cancer related genes varies with lifestyle and geographical distribution among population, so the actual mechanism behind the progression is unknown. Correlating Gastric Cancer with its nature of gene expression and occurrence helps in identifying suitable targets at gene level.

Material and Methods

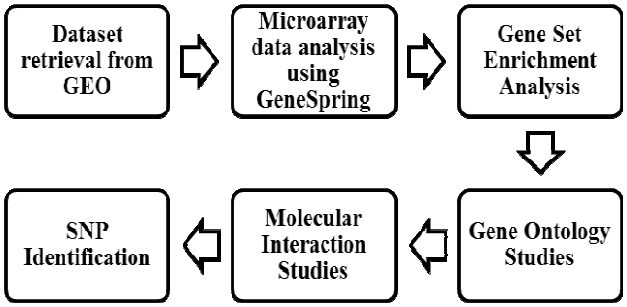

Figure-1
Methodology

Dataset retrieval from GEO: Gastric cancer microarray datasets of Singapore (GSE37023) and UK (GSE15460) were downloaded from the GEO database for comparative studies between populations. These microarray datasets contain normal and gastric cancer tissue samples. These samples were analyzed for their platforms to arrive at uniformity and samples arising

from GPL96 (Affymetrix Gene Chip Human Genome U133A Array) was only considered for the study.

Microarray data analysis using GeneSpring: GeneSpring was used to analyze the expression profiles of genes in datasets. Datasets were divided into two expression data matrices; one matrix includes all normal samples and the other all tumour samples. The data were subjected to statistical analysis for normalization using Robust MultiArray (RMA) algorithm. T-test was used to study the overall significance of differentially expressed genes. Here, Benjamin Hochberg False Discovery Rate (FDR) analysis was used to correct false discovery rates. Genes were filtered based on p-value (0.005) and false discovery rate of less than 1%[5].The obtained list of genes was further filtered by subjecting it to fold change (≥2.0). Further, an unsupervised hierarchical clustering of differentially expressed genes across all Gastric cancer samples was done.

Genes that were "up regulated" were clustered into one group and the genes that were "down regulated" clustered into another group in both the populations. Genes that were differentially expressed in stage 1 was only focussed in the present study. stage 1 is the preliminary stage in Gastric cancer while it progresses to final metastasis in stage 4. Since the present study mainly focuses on finding early markers, the gene expressed in stage 1 was taken into consideration.

Gene Set Enrichment Analysis: Gene set enrichment analysis (GSEA) was performed to determine differentially expressed (up regulated and down regulated) genes that were enriched corresponding to specific functional pathways. Gene set enrichment analysis and pathway analysis using differentially expressed genes in stage 1 of Gastric cancer was carried out. A list of up regulated and down regulated genes was searched in Gastric cancer expression data for Singapore and UK population against a curated gene sets for canonical pathways from the molecular signature database (MsigDB).Using a P value cutoff ≤ 0.005 and FDR ≤ 0.25 genes were filtered and analyzed the functional insight of UP and DOWN regulated genes.

Gene Ontology Studies: After gene enrichment the pathway analysis in Stage 1 Gastric cancer differentially expressed genes was carried out using Cytoscape Bingo plug-in. The biological networks gene ontology tool (Bingo) is an open-source Java tool to determine which Gene Ontology (GO) terms are significantly overrepresented in a set of genes. Bingo can be used either on a list of genes, pasted as text, or interactively on sub graphs of biological networks visualized in Cytoscape[6]. Differentially expressed genes obtained after filtering based on fold change cut off (FC >=2.0) in stage 1were taken as the input list for pathway analysis.

Molecular Interaction Studies: The Cytoscape MiMI Plug-in retrieves molecular interactions and displays the interaction networks along with its attributes. MiMI integrates data from multiple well known protein interaction databases using an intelligent deep-merging approach. It facilitates access to the molecular interaction data assembled in MiMI from the Cytoscape user interface for the biological users to readily explore and analyze interaction data in MiMI[7]. Only genes that show good interactions with its neighbours in both the populations were taken for further analysis.

SNP Identification: Single nucleotide Polymorphism (SNP) was to correlate their expression and the polymorphic changes with Gastric cancer. Previously reported Gastric cancer associated genes and their SNPs were compiled for comparison. SNPs can also be used to study genetic abnormalities in cancer. SNPs can be used to study loss of heterozygosity (LOH). LOH occurs when one allele of a gene is mutated in a deleterious way and the normally functioning allele is lost. LOH occurs commonly in oncogenesis[8].

Results and Discussion

Dataset retrieval from GEO: Microarray Gastric cancer datasets for Singapore and UK populations were retrieved from Gene expression database (GEO) for gene expression studies. Only datasets pertaining to Gastric cancer staging and array type Affymetrix human chip (HG-U133A) were taken for the study. 83 samples from Singapore and 30 samples from UK were taken for the analysis. Samples were categorized based on their stages.

Microarray data analysis: After the selection of microarray gastric cancer datasets for Singapore and UK population, gene expression studies were done using GeneSpring. After grouping of samples into normal and tumour normalization of the data's were done by RMA algorithm to remove the degraded probes**.** After grouping samples in two groups namely control and tumor data's were subjected to normalization to remove all the mismatch probes and summarize all the perfect probes. After performing the normalization the data's converges from the extremes to a common point. Now statistical analysis was done based on a fold-change cut off of 2 and the statistical analysis using t test, 126 genes was observed to be up regulated (UK – 59, Singapore – 68), while 99 genes were down regulated (UK – 82, Singapore – 17) in Gastric cancer tumour tissues when compared to normal tissues. Among these, 19 genes were seen to be up regulated and 6 down regulated in UK and 28 in up regulated and 11 in down regulated for Gastric cancer tissues.

The stage wise gene expression in Gastric cancer was studied to predict novel genes pertaining to particular stage especially stage 1 which is the preliminary and early stage of Gastric cancer. Few genes were observed to be uniquely expressed in stage 1 (Singapore population: 28 up regulated and 11 down regulated, UK gastric cancer population: 19 up regulated and 6 down regulated). These genes can be targeted for novel markers to diagnose Gastric cancer at early stage in these populations.

Table-1
Unique Genes in Stage 1 of Singapore and UK in Gastric Cancer Population

Up regulated Genes in Singapore	Up regulated Genes in UK	Down regulated Genes in Singpaore	Down regulated Genes in UK
Akr1c2	*Abca8*	*Arpc1b*	*Actr2*
Akr7a3	*Azgp1*	*Cdc42bpa*	*Bbx*
Aldh3a1	*Chga*	*Myo6*	*Ddx24*
Anxa10	*Ckb*	*Ptma*	*Dkc1*
C11orf9	*Cxcl12*	*Tpr*	*Hist1h4c*
Ca9	*Defb1*	*Zc3h15*	*Kiaa1033*
Cfd	*Duox1*		*Mfhas1*
Cyb5a	*Fmo5*		*Nktr*
Dhrs7	*Ghr*		*Nop56*
Duox2	*Gsta1*		*Srsf2ip*
Galnt6	*Hnrnpc*		
Gpx3	*Kcnj15*		
Hba1	*Kdm5d*		
Hpgd	*Mt1e*		
Igkc	*P2ry14*		
Igkv1-5	*Pou2af1*		
Igll5	*Scgb2a1*		
Klf4	*Tcn1*		
Klk11	*Zbtb16*		
Lepr			
Muc5ac			
Pgc			
Ppap2b			
Ptger3			
Rab11fip			
Serpina5			
Hnrnpc			
Tff2			

Table-2
Enrichment of Stage 1 Genes in Singapore and UK Gastric Cancer Population

Population	Description	p-value	FDR
Singapore	RNA binding	4.6E-4	1.69E-1
	Nucleic acid binding	8.03E-4	1.48E-1
UK	ATPase activity	1.16E-4	4.97E-4
	Hydrogen ion trans-membrane transporter activity	1.41E-4	2.49E-4

Note: FDR is false discovery rate; FDR and p-value is computed to find the significant genes responsible for Gastric cancer.

The list of UP and DOWN regulated genes uniquely expressed in stage 1 of both the Gastric cancer populations is also mentioned in table-1.

Gene set enrichment analysis: To gain the functional insights from stage1 genes that were differentially expressed in Gastric cancer, gene set enrichment analysis and pathway analysis was carried out.

Table 2 shows that genes in Singapore population are involved in RNA binding and nucleic acid binding. Genes in UK population is involved ATPase activity and Hydrogen transmembrane transporter activity. P-values and FDR q-values for each function are also given. In statistical significance testing the p-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true[9]. False discovery rate (FDR) control is a statistical method used in multiple hypotheses testing to correct for multiple comparisons. In a list of statistically significant findings, FDR procedures are designed to control the expected proportion of incorrectly rejected null hypotheses [10].

Gene ontology studies: Cytoscape plug in Bingo is a tool to determine which Gene Ontology (GO) categories are statistically overrepresented in a set of genes or a subgraph of a biological network. Bingo maps the predominant functional themes of a given gene set on the GO hierarchy and outputs this mapping as a Cytoscape graph. Comparative studies reveals digestion pathway as a common pathway in both the populations.

Table 3 shows the ontological functions for the Stage 1genes in Singapore and UK population. Stage 1 genes in Singapore are involved in digestion and cartilage development process. Stage 1 genes in UK are involved in only Digestion process. After performing enrichment studies in Singapore and UK population, enriched genes were taken to perform gene ontology in both populations to find the pathway enrichment for the genes involved to get an idea about the functions.

Molecular interaction studies: After the gene ontology, the enriched genes molecular interactions were studied by MiMI to find the highly interacted genes. The Cytoscape Plugin MiMI retrieves data from the MiMI repository of molecular interactions developed by National Center for Integrative Biomedical Informatics (NCIBI) and displays the interaction networks using Cytoscape. Protein interaction data exists in many repositories, each with its own data format, molecule identifier and supplementary information.

Table-3
Gene Ontology for Stage 1 Genes Expressed in Singapore and UK Gastric Cancer Population

Population	Genes		Ontology
Singapore	*CHIA*	*PGA3*	Digestion
	CCKBR	*PGA4*	
	CAPN9	*GA5*	
	GRB2	*GKN1*	
	PRSS3	*HNRNPC*	
	COL2A1		Cartilage development involved in endochonral bone morphogenesis
	COL1A1		
	GHR		
UK	*TFF2*	*AKR1C2*	Digestion
	PGA3	*CHIA*	
	GRB2	*CCKBR*	
	PGA4	*PGC*	
	MUC5AC	*CAPN9*	
	PGA5	*HNRNPC*	

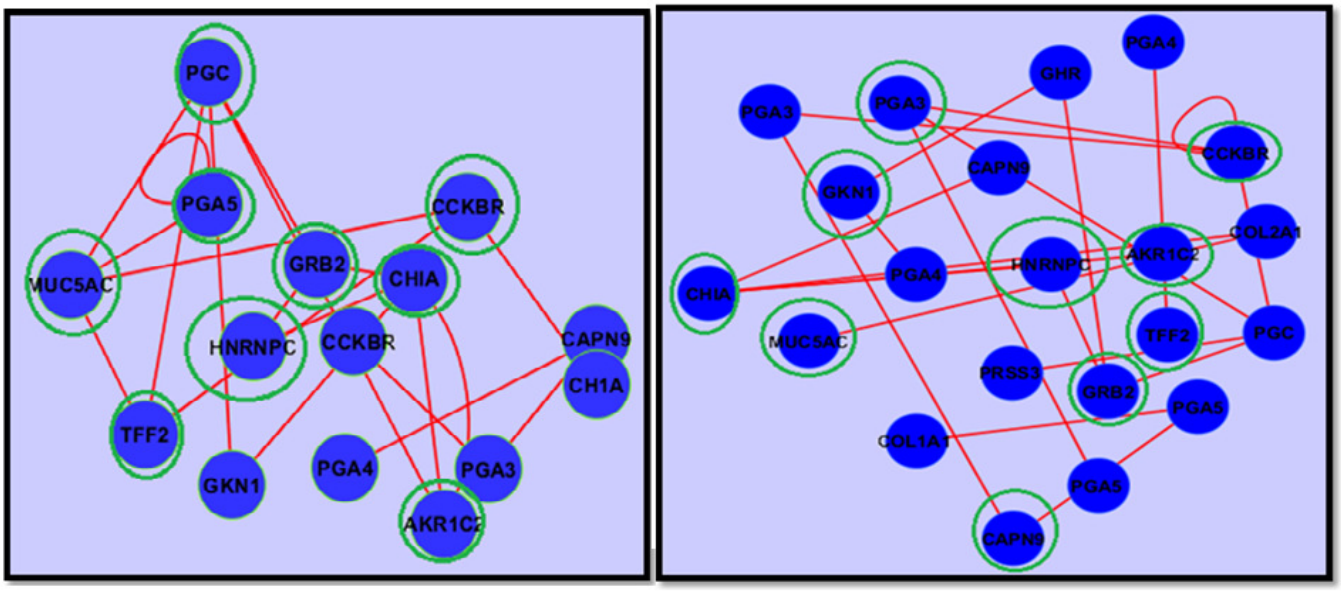

Figure-2
Stage 1 Interacting Genes in Singapore and UK Gastric Cancer Population

Figure 2 shows the interaction of genes in digestion process for Singapore and UK population in Cytoscape. Interactions among the genes were formed using MiMI plug-in of cytoscape. The Stage 1 genes which are present in both the populations are highlighted in green colour. MiMI integrates data from multiple well-known protein interaction databases using an intelligent deep-merging approach. Cytoscape plugin that facilitate access to the data assembled in MiMI. Among all the interacting genes, hnRNP C gene was found to be highly interacting in both the population.

SNP identification: The genes that show good interactions with its neighbours in both the populations were analyzed for presence of single nucleotide Polymorphism (SNP). This was to correlate their expression and the polymorphic changes with Gastric cancer. Previously reported Gastric cancer associated genes and their SNPs were complied for comparison (table 4).

Table-4
SNPs Reported in Genes Associated with Gastric Cancer

Sl.No	SNPs
1	GSTT1, HNRNPC [11]
2	NAT2, XRCC1[12]
3	CYP1A1, GSTP1[13]
4	IB(IL1B)[14]
5	COX2[12]
6	MMP3[15]
7	IL1[16]
8	TGF beta[17]
9	MLH1[18]
10	BRCA2, MRC1[19]
11	GSTM1[12]
12	IL10, XDP[20]
13	MSH2[21]

As shown in table 4, among the 17 gastric cancer associated genes reported to have SNP, only hnRNPC was reported in the present study in Stage 1 of both Singapore and UK population. It also showed highest interaction with other genes in both the population.

Therefore attempt was made to study the SNP in hnRNPC gene. The SNP id was collected for hnRNPC gene from NCBI. Then the sequence was collected. Start codon and the promoter regions were identified manually. Through this study it was identified that two of the SNPs were located in the promoter region (GC BOX and TATA BOX). Hence hnRNPC can be a marker to detect the early progression of gastric cancer.

Heterogeneous Nuclear Ribonucleoprotein C (hnRNP C) was reported as a one of the gene causing Gastric cancer. hnRNP C gene which plays a major role in progression of Gastric cancer. hnRNP C gene is found to interact with Grb2 gene in the progression of Gastric cancer. hnRNP C gene is found to be up regulated in Gastric cancer with a p-value 0.012. hnRNP C gene plays a major role in early steps of spliceosome assembly and pre-mRNA splicing. Growth factor receptor-bound protein 2 also known as Grb2 is an adaptor protein involved in signal transduction and cell communication[11].

Figure 3 shows the sequence obtained from NCBI database. Initially the start codon was identified and promoter regions ie, TATA BOX which are located 25bp upstream, CAAT BOX which are located 75 bp upstream and GC BOX which located 90bp upstream were found. Then the single nucleotide polymorphisms were identified and found two of the SNPs were located in the GC BOX and TATA BOX. Hence it may one of the reasons for the up-regulation of Gastric cancer. So hnRNPC can be used as a marker for the early diagnosis of Gastric cancer.

Microarray gastric cancer datasets were retrieved from GEO and statistical analysis was done to find the differentially expressed genes in stage 1. Genes were filtered based on p-value cut off ≥ 0.005 and false discovery rate of <1%. Based on a fold-change cut off ≤ 2, 68 genes in Singapore population and 59 genes in UK population were found to be up regulated while 17 genes Singapore population and 82 genes in UK population were down regulated in tumour tissues as compared to normal tissues in stage 1. Genes that were up regulated were clustered into one group and the genes that were down regulated clustered into another group. To gain functional insights from the stage 1 genes that were differentially expressed in Gastric cancer enrichment analysis and network analysis were performed. GSEA was performed to determine differentially expressed genes that were enriched corresponding to specific functional pathways. Biological network analysis was carried out using Cytoscape, to generate an interaction database for stage 1 genes of Gastric cancer. Differentially expressed genes were given as input, which resulted in the generation of a complex network based on the connectivity between the genes of stage 1 in both Singapore and UK population. The generated network had various nodes which form highly interconnected sub networks. Most of the stage 1 genes in Singapore population were involved in RNA binding and nucleic acid binding. As well as the genes in UK population were involved in protein binding. Ontology studies showed digestion pathway as the pathway in both the populations. In this pathway hnRNPC played a major role. So the study was focused on SNPs of hnRNPC gene. The SNPs of hnRNPC genes was found to be located in the GC BOX and TATA BOX. It may be one of the reasons for the up regulation of Gastric cancer. The hnRNPs are RNA binding proteins and they complex with heterogeneous nuclear RNA (hnRNA). These proteins are associated with pre-mRNAs in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism and transport. Since it is interacting with Grb2, which is an important splicing molecule. This leads to the synthesis of Grb2 protein which is reported to have a strong association in Gastric cancer. So hnRNPC can be used as a marker for early detection of Gastric cancer.

Table-5
SNP Reported in hnRNPC Gene

SNP Ids	Sequence
rs1130990	TCGGTCCCCACTATGGGCCTGGAGCT(C/G)TACCTGGACCTGCTGTCCCAGCCCT
rs56106137	GGGCTGGGACAGCAGGTCCAGGTACA(A/G)CTCCAGGCCCATAGTGGGGACCGAC
rs149896285	GCAGGTCCAGGTACAGCTCCAGGCCC(A/G)TAGTATCGGGGGACCGACCTGACCG

Table 5 shows the reported SNPs of hnRNPC gene and with their SNP Ids.

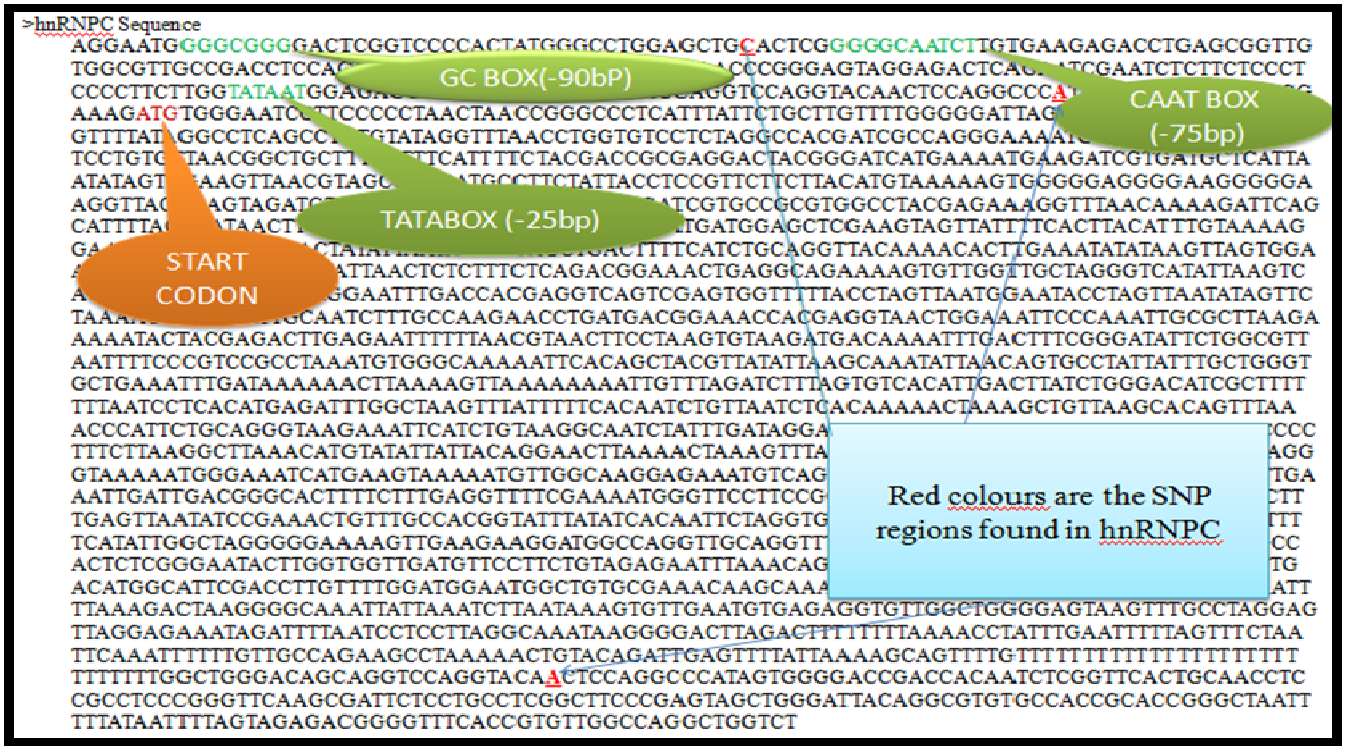

Figure-3
SNPs in hnRNP C for Gastric Cancer

Conclusion

Microarray gastric cancer datasets were retrieved from GEO and statistical analysis was done to find the differentially expressed genes in stage 1. Genes were filtered based on p-value cut off ≥0.005 and false discovery rate of <1%. Based on a fold-change cut off ≤ 2, 68 genes in Singapore population and 59 genes in UK population were found to be up regulated while 17 genes Singapore population and 82 genes in UK population were down regulated in tumour tissues as compared to normal tissues in stage 1. Genes that were up regulated were clustered into one group and the genes that were down regulated clustered into another group. To gain functional insights from the stage 1 genes that were differentially expressed in Gastric cancer enrichment analysis and network analysis were performed. GSEA was performed to determine differentially expressed genes that were enriched corresponding to specific functional pathways. Biological network analysis was carried out using Cytoscape, to generate an interaction database for stage 1 genes of Gastric cancer. Differentially expressed genes were given as input, which resulted in the generation of a complex network based on the connectivity between the genes of stage 1 in both Singapore and UK population. The generated network had various nodes which form highly interconnected sub networks. Most of the stage 1 genes in Singapore population were involved in RNA binding and nucleic acid binding. As well as the genes in UK population were involved in protein binding. Ontology studies showed digestion pathway as the pathway in both the populations. In this pathway hnRNPC played a major role. So the study was focused on SNPs of hnRNPC gene. The SNPs of hnRNPC genes was found to be located in the GC BOX and TATA BOX. It may be one of the reasons for the up regulation of Gastric cancer. The hnRNPs are RNA binding proteins and they complex with heterogeneous nuclear RNA (hnRNA). These proteins are associated with pre-mRNAs in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism and transport. Since it is interacting with Grb2, which is an important splicing molecule. This leads to the synthesis of Grb2 protein which is reported to

have a strong association in Gastric cancer. So hnRNPC can be used as a marker for early detection of Gastric cancer.

References

1. Gao Y., Hu N., Han X.Y., Ding T. and Giffen C., Risk factors for esophagel and gastric cancers in Shanxi Province, China: A case-control study, *Cancer Epidemiology,* **35**, 91–99 **(2011)**

2. Ferlay J., Shin H.R., Bray F., Forman D., Mathers C. and Parkin D.M., Cancer Incidence and Mortality Worldwide, *Globocan,* **12,**184-187 (**2010**)

3. Gajalakshmi V., Swaminathan R. and Shanta V., An Independent Survey to Assess Completeness of Registration: Population Based Cancer Registry, Chennai, India, *Asian Pacific Journal Cancer Prevention,* **2,** 179-183 **(2001)**

4. Mathew A., Gangadhara P., Varghese C. and Nair M.K., Diet and Gastric Cancer: a case-control study in South India, *European Journal of Cancer Prevention,* **9**, 89-97 **(2000)**

5. Gomes L.I., Waleska K., Adriane P., Fernando A., Gustavo H. and Reis L.F., Expression profile of malignant and non-malignant lesions of oesophagus and stomach: differential activity of functional modules related to inflammation and lipid metabolism, *Cancer research,* **65**, 1087-1231**(2003)**

6. Maere S., Heymans K. and Kuiper M., BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks, *Bioinformatics,* **21,** 3448-3449 (**2005**)

7. Jing G., Alex S., Ade V., Glenn T., Terry E.W., Barbara R., Mirel H.V., Jagadish I. and David J., States Integrating and Annotating the Interactome using the MiMI plugin for Cytoscape, *Bioinformatics,* **25,** 137-138 (**2009**)

8. Drabovich A.P. and Krylov S.N., Identification of base pairs in single-nucleotide polymorphisms by MutS protein-mediated capillary electrophoresis, *Analytical Chemistry,* **78**, 2035-2038 (**2006**)

9. Goodman S.N., Toward Evidence-Based Medical Statistics. 1: The P Value Fallacy, *Annals of Internal Medicine,* **130**, 995–1004 (**1999**)

10. Benjamini Y. and Hochberg Y.F., Controlling the false discovery rate: a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society,* **57,** 289–300 **(1995)**

11. Yadav A., Jain D.K. and Kumar C., Preparation and characterization of microparticulate system of propranolol hydrochloride, *Journal of Drug Delivery and Therapeutics,* **2**,15-20 **(2011)**

12. Malik M.A., Upadhyay R., Mittal R.D., Zargar S.A., Modi D.R. and Mittal B., Role of xenobiotic-metabolizing enzyme gene polymorphisms and interactions with environmental factors in susceptibility to gastric cancer in Kashmir Valley, *Gastrointestinal Cancer,* **40**, 26–32 **(2009)**

13. Gonzalez C.A., Sala N. and Capella G., Genetic susceptibility and gastric cancer risk, *International Journal of cancer*, **100**, 249-260 **(2002)**

14. Kumar C., Jain DK. and Yadav A., Preparation and characterization of microparticulate system of propranolol hydrochloride, *Journal of Drug Delivery and Therapeutics,* **2,** 15-20 **(2011)**

15. Dey S., Stalin S., Gupta A., Saha D. and Kesh K., Matrix metalloproteinase3 gene promoter polymorphisms and their haplotypes are associated with gastric cancer risk in eastern Indian population, *Molecular Carcinogenesis,* **51**, 42–53 **(2011)**

16. Moorchung N., Srivastava AN., Gupta NK., Achyut BR. and Mittal B., The Histopathology of Chronic gastritis, *Indian Journal of Pathology and Microbiology,* **50,** 18-24 **(2007)**

17. Moutri L. and Ali W., COX1 and COX2 Polymorphisms and Gastric Cancer Risk in a Polish Population, *Anticancer Research,* **27**, 43-48 **(2007)**

18. Kirk B.W., Feinsod M., Favis R., Kliman R. and Barany F., Single nucleotide polymorphism seeking long term association with complex disease, *Nucleic Acids Research,* **30,** 3295–3311 **(2002)**

19. Johannsson O., Loman N., Moller T., Kristofferson U., Borg A. and Olsson H., Incidence of malignant tumours in BRCA1 germline mutation carriers, *European Journal of Cancer,* **35,** 1248-1257 **(1999)**

20. Matullo G., Palli D., Peluso M., Guarrera S., Carturan S., Celentano E., Krogh V., Munnia A., Tumino R., Polidoro S., Piazza A. and Vineis P., XRCC1, XPD gene polymorphisms, smoking and (32)P-DNA adducts in a sample of healthy subjects, *Carcinogenesis,* **22,** 1437–1445 **(2001)**

21. Liu Q. and Sommer S.S., Pyrophosphorolysis-activatable oligonucleotides may facilitate detection of rare alleles, mutation scanning and analysis of chromatin structures, *Nucleic Acids Research,* **30,** 598–604 **(2009)**