Journal of
Proteomics & Bioinformatics

ISSN: 0974-276X

Malik et al., J Proteomics Bioinform 2011, 4:12
DOI: 10.4172/jpb.1000201

Research Article Open Access

In Silico Gene Expression Based Analysis on Claudin Family Members Association with Human Thyroid Cancer

Shaukat Iqbal Malik[1]*, Zoya Khalid[2] and Sheema Sameen[2]

[1]*Department of Bioinformatics, Mohammad Ali Jinnah University, Islamabad, Pakistan*
[2]*Department of Bioinformatics and Biotechnology, International Islamic University, Islamabad, Pakistan*

Abstract

Thyroid cancer is one of the major types of cancers worldwide. A large amount of effort has been put in order to find the genetic basis of this cancer. Many genes have been reported before out of which claudin family is one of the example. The claudin protein family consists of more than 20 members which are made up of key structural rudiments inside the tight junction. The association of claudin-1 with the thyroid cancer has already been predicted experimentally. In order to investigate the genetic reason of human thyroid cancer computationally, the systematic analysis of claudin gene family has been carried out. To fulfill this task the bioinformatics methodologies are combined for the assessment of claudin gene family expressions. Results obtained showed and verified the association of CLDN1 member with the thyroid cancer.

Abbreviations: PTC: Papillary Thyroid Carcinoma; SAGE: Serial Analysis of Gene Expression; TPM: Tags per Million; B-RAF: v-Raf murine sarcoma viral oncogene homolog B1; RAS Rat Sarcoma

Introduction

Thyroid cancer is the most common tumor of endocrine system. It has three major subtypes out of which papillary thyroid carcinoma accounts as the major type. The molecular genetics analysis of thyroid cancer has revealed the presence of B-RAF and RAS point mutations [1]

Papillary thyroid carcinoma is well differentiated thyroid cancer. Experimentally it has been shown that constitutive activation of MAP (Mitogen Activating Protein) kinase effectors play a major role in causing papillary thyroid carcinoma. In addition to that the activating mutations of the tyrosine receptor kinases RET and NTRK and point mutations of the intracellular signaling effectors RAS and BRAF are found to be associated with this. The chromosomal recombination is the cause of RET activation that ultimately fall out with a fusion protein made up of the intracellular tyrosine kinase domain of RET coupled to the N-terminal fragment of a heterologous gene, giving rise to the RET/PTC oncoproteins [2]. In comparison with the childhood and adult onest papillary thyroid carcinoma it has been reported that in childhood papillary thyroid cancer, despite of history of radiation exposure, RET/PTC rearrangements are a key event, while in adult-onset papillary thyroid cancer, the most common molecular event is BRAF point mutation, instead of the RET/PTC rearrangements [3].

The previous analysis on BRAF gene shows that BRAF mutation may be a key genetic factor for the metastatic progression of papillary thyroid carcinoma. In addition to that it has also been reported that this gene mutation is a major risk factor for loco regional lymph node metastasis and has potential utility as a surrogate marker [4]. As 80% of all the thyroid cancers diagnosed are of papillary carcinoma type ,this study will mainly focus on this sub type of human thyroid cancer.

The claudin protein family consists of more than 20 members which are made up of key structural rudiments inside the tight junction [5]. Tight Junctions have the capability to maintain the cell polarity which may regulate cell proliferation and differentiation [6]. The claudin (*CLDN*) genes encode a family of proteins which play a major role in the configuration of tight junction as well as associated with the proper functioning of these junctions. In the past decades, it has become evident that *CLDN* gene expression is commonly mutated in several human cancers [7,8].

There is a likelihood that along with the activation of MAP kinase effectors, there are other gene family members as well that might be involved in the pathogenesis of papillary thyroid cancer. For that purpose a high throughput in silico analysis has been made to systematically evaluate the gene expression of claudin gene family members. It has been evident that mutations in claudin are associated with various human cancers. For this the whole claudin family is systematically analyzed to check its association with human thyroid cancer. The wet lab analysis has previously been done which shows the involvement of claudin-1 as a genetic change in causing papillary thyroid cancer [9-11].

Methods

The methodology includes the following bioinformatics tools:

Gene finder

With the help of this gene finder tool a particular gene or more than one gene can be found. This tool is available at CGAP http://cgap.nci.nih.gov/Genes/GeneFinder. By selecting the papillary thyroid cancer and claudin gene family as search criteria the tool will provide the list of claudin gene family members which are thought to be involved in papillary thyroid cancer.

SAGE and Virtual Northern Blot Analysis

SAGE is a web based tool that is freely available at NCBI and on

***Corresponding author:** Shaukat Iqbal Malik, Department of Bioinformatics, Mohammad Ali Jinnah University, Islamabad, Pakistan, E-mail: drsimalik@jinnah.edu.pk

Received November 07, 2011; **Accepted** December 05, 2011; **Published** December 12, 2011

Citation: Malik SI, Khalid Z, Sameen S (2011) In Silico Gene Expression Based Analysis on Claudin Family Members Association with Human Thyroid Cancer. J Proteomics Bioinform 4: 278-283. doi:10.4172/jpb.1000201

CGAP database. SAGE provides the quantitative and concurrent analysis of a large number of transcripts and also makes available applicable means for the evaluation of expressed genes in different tissue states. SAGE measures the number of tags that contains the adequate information to distinctively identify a transcript. There are two SAGE libraries available, one for normal tissue and one for papillary thyroid tissue. The chosen tumor tissue was follicular variant as its library is available at SAGE genie [11]. SAGE Anatomic Viewer which is available at http://cgap.nci.nih.gov/SAGE/AnatomicViewer [12] was used for obtaining the tags. The reliable tags obtained from SAGE will then mapped to the unigene clusters. http://www.ncbi.nlm.nih.gov/unigene/.The authentic unigene clusters which were matched to SAGE tags were then picked. These tags were then used for the analysis of gene expression in normal and malignant tissue.

Microarray analysis

The two microarray datasets GDS1732 and GDS1665 available at GEO website is taken http://www.ncbi.nlm.nih.gov/geo/ [13-16].The first data set contains the expression profiling of 7 classical papillary thyroid carcinoma (PTC) samples and 7 paired normal thyroid tissue samples (normal group) while the second data set contains classical papillary thyroid carcinoma (PTC) tumors from 9 patients and 9 control samples. This analysis includes the statistical parameter t-test, and the clustering parameters to confirm the clauidin1 association with papillary thyroid cancer.

T-Test: It separates the significant and non-significant genes on the basis of P-value .The P-value less than 0.005 ($p<0.005$) are considered to be noteworthy. Further volcano plot marked the position of the susceptible gene as either significant or non-significant.

Clustering parameters: The genes which are found to be significant from the T-test analysis are further employed to clustering parameters. These parameters include the three algorithms which are Hierarchical clustering algorithm, K-Means clustering and Self Organizing Tree Algorithm.

Results

Gene finder

The gene finder tool categorized those family members which are more recurrently found to be linked with the papillary thyroid cancer on the basis of evaluation of gene expression between normal and malignant thyroid tissue. The gene members found by the gene finder tool are then chosen for further analysis. It is summarized in Figure 1.

SAGE and virtual northern blot analysis

The reliable tags of 7 claudin genes were extracted as recognized by the gene finder tool. The Tags per million counts (TPM) and the fold change was computed for each gene respectively. The results are summarized in Table 1.

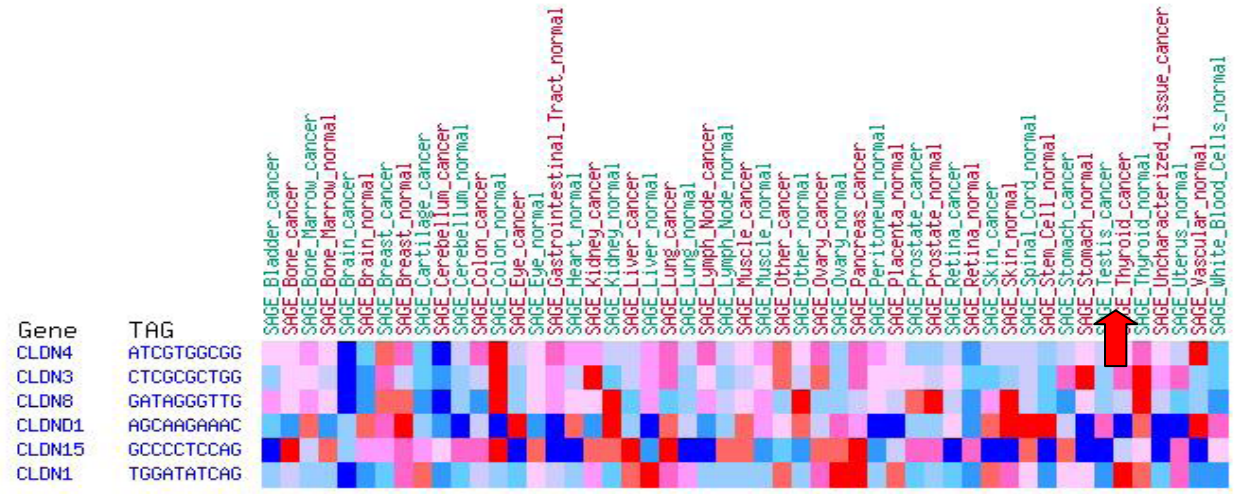

Figure 1: Short SAGE Map.

Gene ID	SAGE Anatomic Viewer			Virtual Northern Blot Analysis		
	Sequence ID	SAGE tags	Unigene Clusters	Normal (TPM)	Cancer (TPM)	Fold Change/Fold Down
CLDN1	NM_021101	TGGATATCAG	Hs.439060	12.075	43.31	3.58
CLDND1	NM_001040199	AGCAAGAAAC	Hs.531371	8	1	-8 /8
CLDND2	NM_152353	CAGAGCACCG	Hs.720536	0	0	0
CLDN8	NM_199328	GATAGGGTTG	Hs.162209	1	0	0
CLDN4	NM_001305	ATCGTGGCGG	Hs.647036	20	6	-3.33/3.33
CLDN3	NM_001306	CTCGCGCTGG	Hs.647023	156	28	-5.58/5.58
CLDN15	NM_001185080	GCCCCTCCAG	Hs.38738	1	1	1

Table 1: SAGE and Virtual Northern Blot Analysis.

Microarray analysis of claudin gene family

For the Micro array analysis the comparison is carried out between the two phenotypes in both the data sets. The two data sets GDS1732 and GDS1665 available at GEO website is taken. The Table 2 listed down the results obtained from microarray analysis which clearly shows CLDN1 as the most significant gene according to the two fold change difference. Further the results obtained from t-test and clustering parameters also confirmed the above mentioned outcomes (Table 3, Figure 2, 3, 4 and 5 respectively).

Discussion

In the past few years a large amount of effort has been done in

GDS1732				GDS1665			
Gene ID	**Normal**	**Cancer**	**Fold Change/Fold Down**	**Gene ID**	**Normal**	**Cancer**	**Fold Change/Fold Down**
CLDN1	188.83	654.2	3.46	CLDN1	6.22	16.89	2.71
CLDND1	892.802	664.034	-1.345/1.345	CLDND1	8.81	8.1	-1.08/1.08
CLND8	990.2	696.85	-1.422/1.422	CLND8	8.83	9	1.01
CLDND2	190.36	134.018	-1.42/1.42	CLDND2	6.51	6.47	-1.01/1.01
CLDN4	123.791	150.692	1.21	CLDN4	7.7	7.8	-1.01/1.01
CLDN3	264.2	470.13	1.779	CLDN3	10.1	9.83	-1.02/1.02
CLDN15	0.0	175.615	0	CLDN15	7.01	6.89	-1.02/1.02

Table 2: Microarray Analysis in normal and papillary thyroid tissue cells.

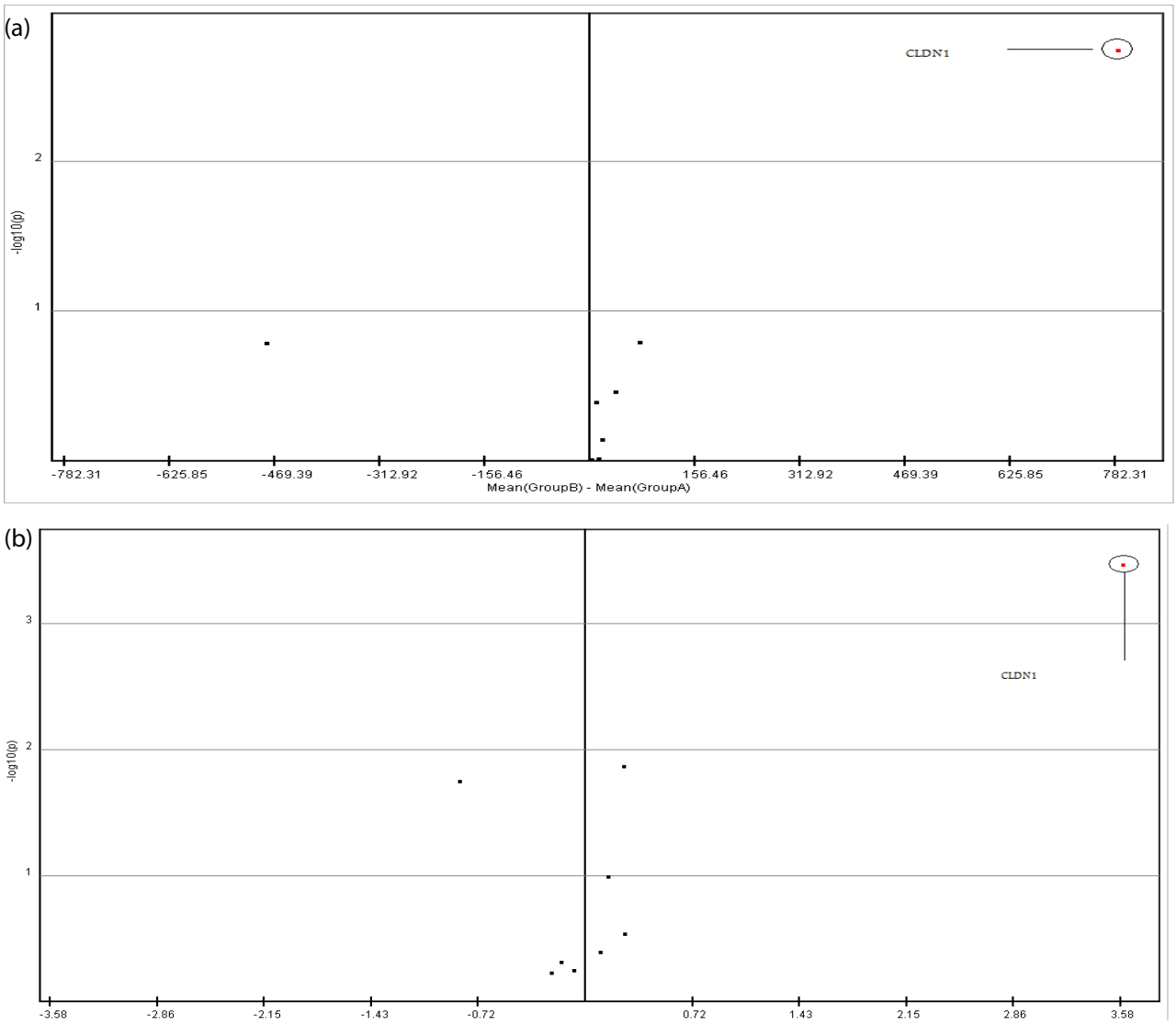

Figure 2: Log p-value versus mean difference between the two phenotypes: Red color shows the more significant genes. The black color is indicating the non significant gene. (A) GDS1732 (B) GDS1665

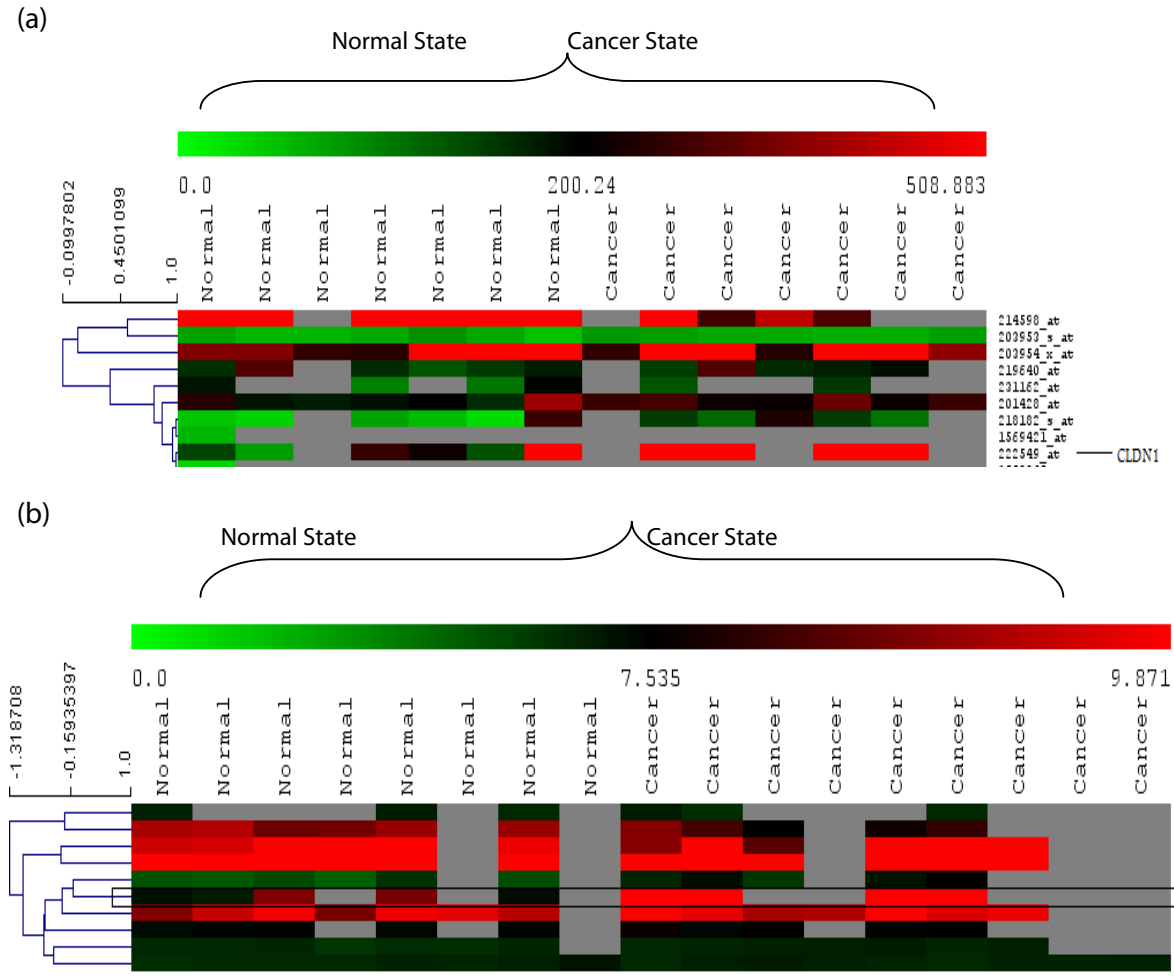

Figure 3: Hierarchical Clustering diagrams of the two datasets (a) GDS1732 (b) GDS1665. Red color indicates the highly expressed genes, Green color shows down regulating genes and Gray color indicates the missing values in the datasets.

Probe ID	Gene Symbol	P-value		Expression Level
218182_s_at	CLDN1	0.001	0.003	High (Up Regulated)

Table 3: Significant P-values in both the datasets.

finding the genetic cause of thyroid cancer. All the practice done was based on experimental analysis. These experimental technologies had revealed some gene family members which are deemed to be associated with the human thyroid cancer.

In this study instead of experimental analysis the in silico analysis approach has been adopted for the assessment of gene expression analysis in human thyroid carcinoma. To implement this approach practically the claudin gene family members has been acquired. Up to our information for the very first time, the systematic investigation of claudin family members has been carried out for gene expression analysis in papillary thyroid cancer a major type of thyroid carcinoma. The methodology adopted will help in scanning the gene expression patterns in less time and money. The results obtained from SAGE and virtual Northern Blot analysis showed clearly that out of these 7 genes only CLDN1 is found to be over expressed by taking the fold change value > 2 fold as significant. The other genes CLDND1, CLDN3, CLDN4, are found to be expressed in normal thyroid tissue. While there is no significant difference of expression level found in CLDND2, CLDN4, CLDN15 genes.

In addition to this the microarray analysis also determined the over expression of CLDN1 member in papillary thyroid carcinoma. The claudin gene whose expression was high in both the data sets is considered to be as the valid results and this could be judged by comparing the fold change value. As mentioned above fold change value which is greater than two fold is considered to be noteworthy. This fold change difference is the minimum threshold value to compute the difference in expression level. In accordance that the t-test also shows CLDN1 as positive significant gene, which in turn means that it is up regulating in papillary thyroid carcinoma. Furthermore all the three means of clustering analysis predicted that CLDN1 is over expressed in case of papillary thyroid cancer as its expression status is higher in cancer datasets as compared to normal one.

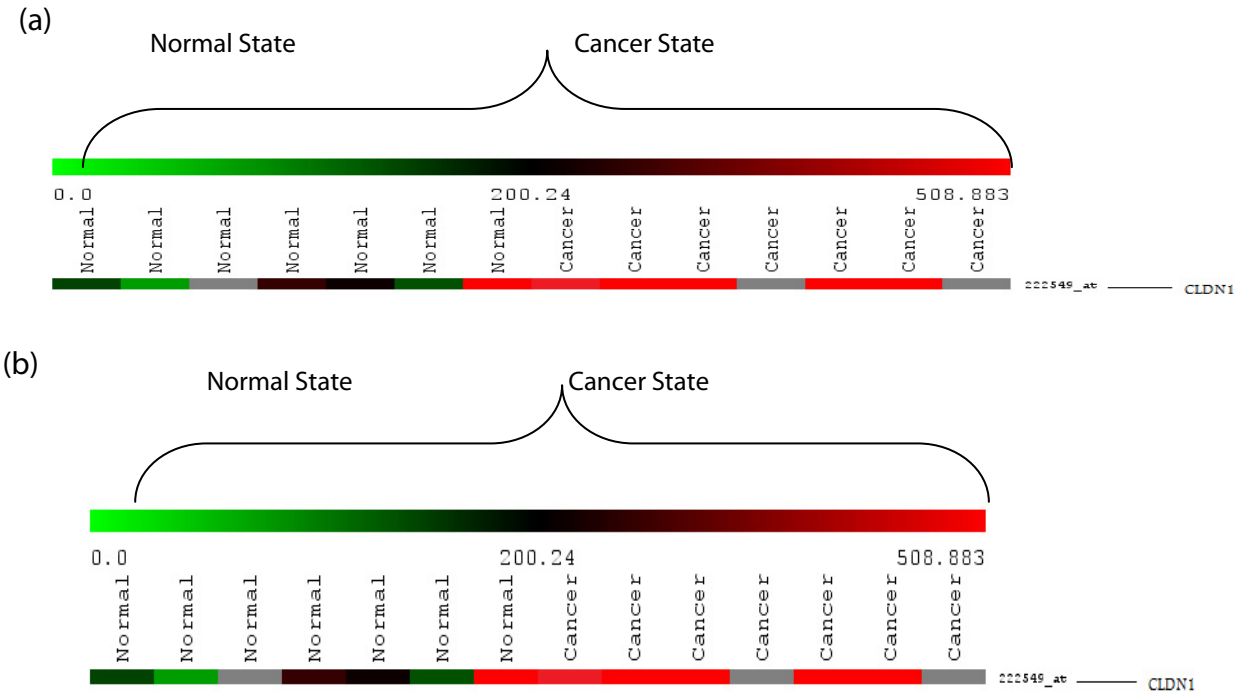

Figure 4: K-Means Clustering diagram of the two datasets (a) GDS1732 (b) GDS1665.

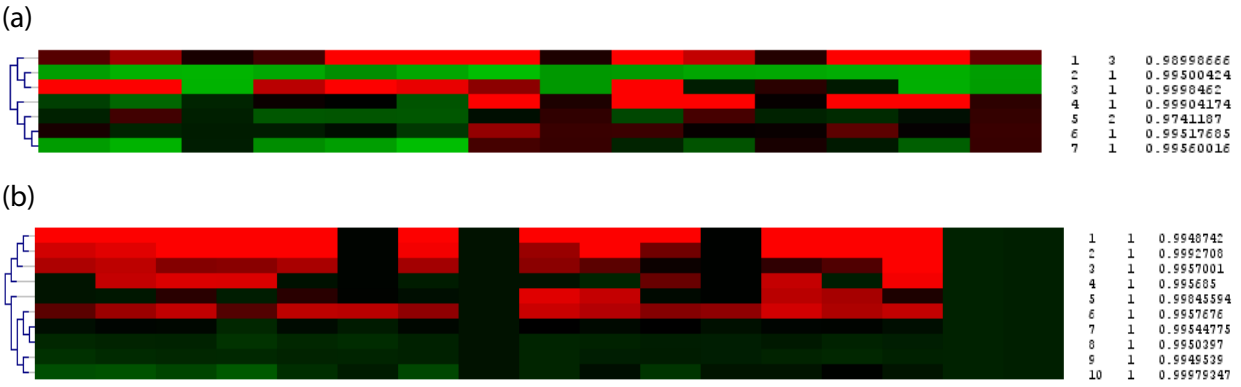

Figure 5: SOTA performed. Heat Map Analysis: Clusters are showing up regulation and down regulation in case of PTC. Red color spectrum shows up regulation while green color spectrum shows down regulation and Grey portion shows missing values in the datasets. (a) GDS1732 (b) GDS1665.

As stated above the association of CLDN1 with papillary thyroid carcinoma has been predicted experimentally before. Our computational approach has verified that CLDN1 act as a genetic change in causing papillary thyroid cancer.

CLDND1, CLDN3, and CLDN4 these three genes are found to be down regulated by the SAGE analysis but in microarray analysis no clear significance is found in both of the datasets.

The in silico bioinformatics approaches are thought to be reliable and accurate as it involved the data for analysis from reliable sources and these are based upon the DNA sequencing. Combined serial analysis can act as good starting point in disease gene discovery. The use of in silico gene mining approaches provides an excellent scaffold for the initial identification of key genes and gene clusters whose expression is altered in disease tissue which provides a road map to assist the biologists.

Conclusion

This study verified the association of claudin-1 as a genetic cause in causing papillary thyroid cancer. By this it can be inferred that bioinformatics methodologies are convincing approaches for the evaluation of gene expression. The results will further be verified by experimental approaches but this study provides an initial point for the biologists to find the new cancer insights.

References

1. Nikiforova MN, Nikiforov YE (2008) Molecular genetics of thyroid cancer: implications for diagnosis, treatment and prognosis. Expert Rev Mol Diagn 8: 83-95.

2. James A. Fagin (2005) Genetics of papillary thyroid cancer initiation: implications for therapy. Trans Am Clin Climatol Assoc 116: 259-269.

3. Hamatani K, Eguchi H, Ito R, Mukai M, Takahashi K, *et al.* (2008) RET/PTC rearrangements Preferentially Occurred in Papillary Thyroid Cancer among Atomic Bomb Survivors Exposed to High Radiation Dose. Cancer Res 68: 7176-7182.

4. Nikiforova MN, Kimura ET, Gandhi M, Biddinger PW, Knauf JA, *et al.* (2003) BRAF Mutations in Thyroid Tumors Are Restricted to Papillary Carcinomas and Anaplastic or Poorly Differentiated Carcinomas Arising from Papillary Carcinomas. J Clin Endocrinol Metab 88: 5399–5404.

5. Swisshelm K, Macek R, Kubbies M (2005) Role of claudins in tumorigenesis. Adv Drug Deliv Rev 57: 919-928.

6. Colato C (2011) Claudin-1 immunohistochemical expression in papillary thyroid carcinoma: a candidate marker? International Journal of Endocrinology and Metabolism.

7. Oliveira SS, Morgado-Díaz JA (2007) Claudins: multifunctional players in epithelial tight junctions and their role in cancer Cell Mol Life Sci 64: 17-28.

8. Morin PJ (2005) Claudin proteins in human cancer promising new targets for diagnosis and therapy. Cancer Res 65: 9603-9606.

9. Németh J, Németh Z, Tátrai P, Péter I, Somorácz A et al. (2010) High Expression of Claudin-1 Protein in Papillary Thyroid Tumor and its Regional Lymph Node Metastasis. Pathol Oncol Res 16: 19-27.

10. Fluge Ø, Bruland O, Akslen LA, Lillehaug JR, Varhaug JE (2006) Gene expression in poorly differentiated papillary thyroid carcinomas. Thyroid 16: 161–175.

11. Pauws E, Veenboer GJ, Smit JW, de Vijlder JJ, Morreau H, et al. (2004) Genes differentially expressed in thyroid carcinoma identified by comparison of SAGE expression profiles. FASEB J 18: 560-561.

12. Liang P (2002) SAGE Genie: a suite with panoramic view of gene expression. Proc Natl Acad Sci U S A 99: 11547-11548.

13. Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J ,et al. (2003) ArrayExpress--a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res* 31: 68-71.

14. Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, et al. (2005) NCBI GEO: mining millions of expression profiles—database and tools. *Nucleic Acids Res* 33: D562-D566.

15. Barrett T, Edgar R (2006) Gene Expression Omnibus : microarray data storage, submission, retrieval, and analysis. *Methods Enzymol* 411: 352-369.

16. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, et al. (2007). NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res* 35: D760-D765.