A High Throughput Computational Analysis of Claudin Gene Family in Human Ovarian Cancer

Dr. Shaukat I. Malik[1], S. Sameen[2], and Z. Khalid[2]

[1]Department of Bioinformatics, Mohammad Ali Jinnah University (MAJU) New Campus, Islamabad, Pakistan
[2]Department of Bioinformatics & Biotechnology, International Islamic University (IIU), Islamabad, Pakistan

Abstract- *Gene expression analysis is very instrumental in understanding the pathogenesis of the disease. To enhance the understanding of the molecular basis of the disease there is a need to extract the buried patterns in gene expression profiles. This paper is intended to provide a computational approach for the analysis of claudin gene family association with the pathogenesis of ovarian cancer. Our analysis verified some major members of claudin family as either up regulated or down regulated. It shows the differential expression of cldn3, cldn4 and cldn7 in ovarian cancer. In addition to that the up regulation of cldn16 and down regulation of cldn5 in human ovarian cancer has also been observed.*

Keywords: Ovarian cancer, Claudin, Expression analysis.

1 Introduction

Ovarian cancer is the sixth most common cause of cancer death among woman worldwide [1]. Environmental and genetic factors are both important in ovarian carcinogenesis. This disease predominantly affects postmenopausal women causing approximately 13,300 deaths each year and for over half of all deaths from genital cancer. The highly lethal nature of this tumor is related to the absence of symptoms in the majority of women with early stages of the disease and it is the leading cause of mortality due to gynecological malignancy. In the past two decades, much progress has been made in identifying genes involved in the development of ovarian cancer. These identified genes are useful in understanding the pathogenesis of ovarian cancer and defining its molecular signature. They can also serve as biomarkers for early diagnosis and targets for drug development. Claudin gene family is implicated with various types of cancers [2] [3] [4] [5]. This family consists of 23 tight junction proteins [6]. The correct arrangement of all claudin genes is very necessary to perform its function which is the formation of tight junctions. Any problem in its gene arrangement causes cancers. Association of ovarian cancer with some members of the claudin family has already been reported before e.g. cldn3 [7], cldn4 [8] and cldn7 [9] [10] [11] [12]. The function of claudins is highly tissue specific because claudin3 and claudin4 was observed in ovarian cancer but not in ovarian cystadenomas [13]. Here we will computationally analyze the whole caludin gene family association with ovarian carcinoma.

2 Methods

2.1 Gene Finder tool

The Gene Finder is a tool that identifies one gene or list of genes, based on selected search criteria. This tool is available at Cancer Genome Anatomy Project (CGAP) official website http://cgap.nci.nih.gov/Genes/GeneFinder. By choosing the search criteria as ovarian cancer and claudin gene family it showed all of the ovarian cancer related genes of claudin family.

2.2 SAGE Genie

The SAGE Genie is a gene expression database that reliably matches SAGE tags, 10 or 17 nucleotides in length, to known genes. It not only produces the list of tags but also provide the frequency of occurrence of these tags in each normal and cancerous tissues by scanning all of the given expression libraries. All publicly available data to date was used for the analysis of gene expression of claudin gene family. SAGE anatomic viewer http://cgap.nci.nih.gov/SAGE/AnatomicViewer [14] was used for collection of tags. Both NlaIII and Sau3A tags were mapped to UniGene clusters http://www.ncbi.nlm.nih.gov/unigene/. The reliable UniGene clusters matched to claudin tags were adopted to determine the levels of expression of claudin gene family in both normal and ovarian cancer libraries. The list of tags and matched unigene clusters is provided in table 2.

2.3 Virtual Northern

Virtual northern available at CGAP allows the user to observe the expression of a specific gene in all SAGE and EST libraries. Five libraries of ovarian carcinoma and two of normal ovarian expression were studied in northern blot analysis for the expression patterns of all of the gene

Symbol	Name	Sequence ID
CLDN1	Claudin 1	NM_021101
CLDN10	Claudin 10	NM_182848 NM_001160100 NM_006984
CLDN15	Claudin 15	NM_001185080 NM_014343
CLDN16	Claudin 16	NM_006580
CLDN3	Claudin 3	NM_001306
CLDN4	Claudin 4	NM_001305
CLDN5	Claudin 5	NM_001130861 NM_003277
CLDN6	Claudin 6	NM_021195
CLDN7	Claudin 7	NM_001307 NM_001185023 NM_001185022
CLDND1	Claudin domain containing 1	NM_001040199 NM_019895 NM_001040182 NM_001040181 NM_001040183 NM_001040200

members of claudin gene family. The difference of greater than two fold was considered significant.

2.4 Microarray analysis

Two microarray datasets having normal and cancerous ovarian cancer tissues available on Gene Expression Omnibus (GEO) http://www.ncbi.nlm.nih.gov/geo/ was used. The dataset GSE6008 contain 99 individual ovarian tumors and 4 individual normal ovary samples contributed by Hendrix ND [15] and the second dataset GSE4122 is contributed by Tate DL and co workers with 32 cancerous tumors and 14 controls. These datasets was used for further verification of SAGE and northern blot analysis results. Statistical analysis was done to analyze the microarray expression. Two major statistics applied to available data was t-test and significance analysis of microarrays (SAM).

3 Results

3.1 Gene Finder Results

Gene finder provided the list of only those claudin gene family members which found to be more frequently involved in the expression of ovarian cancer on the basis of number of tags available in CGAP libraries of ovarian cancer as compared to normal ovarian libraries. These selected gene members from claudin gene family are shown in table 1. Only these selected genes were chosen for further analysis.

Table 1: Gene finder results of claudin gene family members involved in the expression of ovarian cancer

3.2 SAGE and virtual Northern blot analysis of Claudin genes expression in ovarian cancer

There are two normal ovarian libraries and five SAGE libraries of ovarian cancer tissues available in GEO. This was also observed by using tool SAGE Absolute Level Lister (SALL) http://cgap.nci.nih.gov/SAGE/SALL?ORG=Hs available at CGAP website. The reliable tags of all selected claudin gene family members were then extracted from SAGE Genie by as shown in table 2. Only those genes were picked which have atleast > 2 fold difference. From 10 genes 7 genes were found to have greater than 2 fold difference. The cldn1, cldn 10 and cldnd1 have no significant differences or they found to have almost same result that's why they were excluded from the list. The virtual northern results confirm the involvement of cldn3, cldn7, cldn4, cldn15, cldn16, cldn5 and cldn6 in the ovarian cancer.

GENE ID	UNIGENE CLUSTER	SAGE TAG	NORMAL (TPM)	OVARIAN CANCER (TPM)
CLDN3	Hs.647023	CTCGCGCTGG	0.0	77
CLDN7	Hs.513915	TATAGTCCTC	0.0	37
CLDN4	Hs.647036	ATCGTGGCGG	0.0	91
CLDN15	Hs.38738	GCCCCTCCAG	4	9
CLDN16	Hs.251391	TTGCCATCCT	0.0	4
CLDN6	Hs.533779	TTTTGTTAGT	0.0	28
CLDN5	Hs.505337	GACCGCGGCT	0.0	14

Table 2: SAGE Anatomic Viewer and Northern blot analysis results:

3.3 Microarray analysis of Claudin genes expression in ovarian cancer

The involvement of above selected genes of claudin family was then verified by Microarray analysis. Two datasets GSE6008 and GSE4122 from GEO contains the gene expression information related to ovarian cancer. All of the above mentioned genes can be located in these data sets. The results obtained from both datasets (table 4) shows that cldn4, cldn7, cldn16 and cldn3 are highly over-expressed in ovarian cancer while cldn5 is down regulated. Cldn6 and cldn15 showed a very different behavior, as cldn6 is found to have over expression and down regulation of cldn15 in dataset GSE4122 while in GSE6008 no significant difference is detected. These findings through fold change analysis were further verified through t-test and SAM. The t-test confirmed the cldn3, cldn4, cldn7, cldn16, cldn15 and cldn5 as significant genes and cldn6 was the only non significant gene so it was excluded from further analysis. The differential expression of these significant genes was also detected in a volcano plot in fig 1. Further mining of selected members of

claudin family was done through SAM which separated cldn4, cldn3, and cldn7 as positive significant genes and cldn5 as negative significant. This is also shown in SAM graph fig 2.

Table 4: Microarray results of dataset GSE6008 & GSE4122

CLAUDIN GENES	REFERENCE ID	FOLD CHANGE IN GSE 6008	FOLD CHANGE IN GSE4122	P-VALUE	FALSE DISCOVERY RATE
CLDN4	201428_AT	> 2	> 5	0.0018636247	0.003261343
CLDN7	202790_AT	> 2	> 3	2.9067648E-6	6.7824512E-6
CLDN3	203953_S_AT	> 3	> 6	0.0	0.0
CLDN5	204482_AT	< 2	< 3	0.009433004	0.011005172
CLDN15	219640_AT	NO CHANGE	NO CHANGE	1.7732685E-8	6.2064395E-8
CLDN16	220332_AT	> 1.5	> 3	0.004344086	0.0060817203

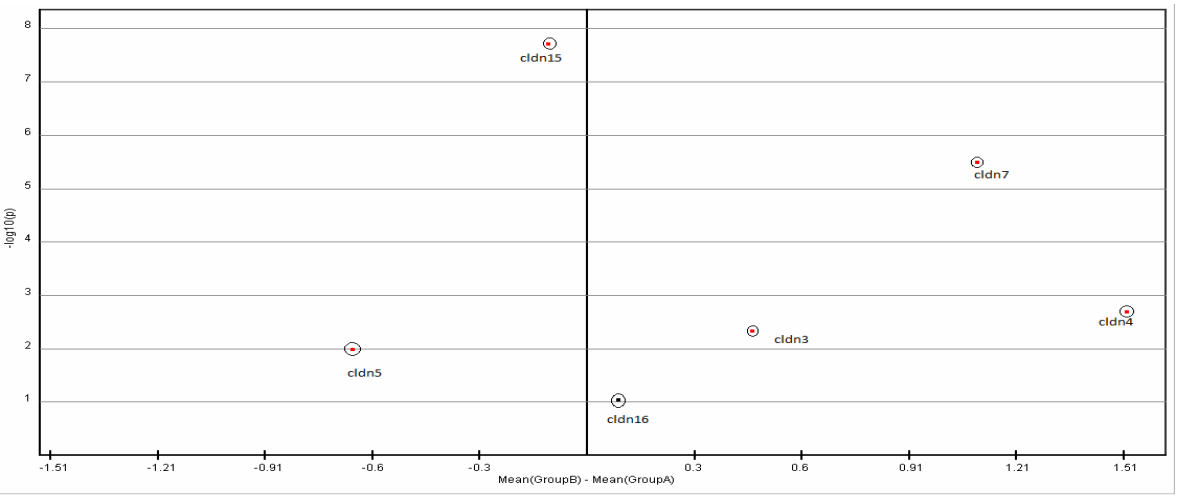

FIGURE 1: VOLCANO PLOT

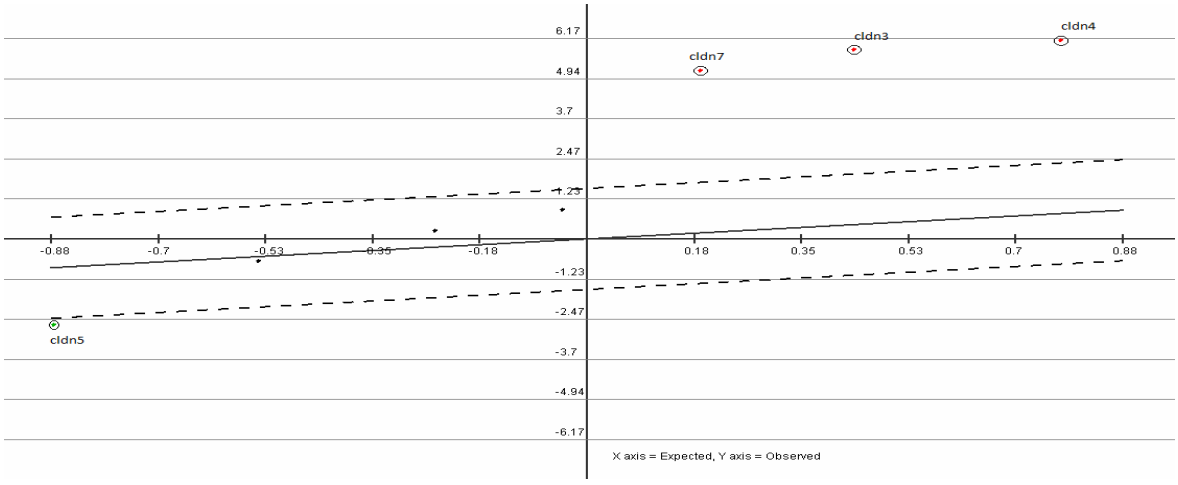

FIGURE 2: SAM GRAPH

4 Discussion

Claudins are tight junction proteins and their involvement in several cancers implicated their role in tumor development. Some of the claudin family members association with the ovarian cancer has already been identified in vitro but this is the first in silico analysis of complete claudin gene family's association specifically with the ovarian cancer. The approach used in this paper is very reliable because the in silico methods of detecting SAGE tags and northern blot analysis are very reliable gene expression methods because they are based upon DNA sequencing. Secondly the microarray analysis was done on the data available on GEO and all this data is experimentally produced data so the chances of error are minimized.

As it is described earlier that three members of claudin family has already been reported for their involvement in expression of ovarian cancer, our results not only verified the significant up regulation of these genes but we also observed the over expression of cldn16 in cancer state. Cldn16 showed up regulation in both SAGE and Microarray analysis. Although the expression of cldn16 is less than the cldn4, cldn3 and cldn7 but its involvement in cancer can never be neglected.

Another interesting finding is the down regulation of cldn5 in microarray data sets, which is very amazing as far as we observe the role of claudins. The tight junction formation ability of claudins makes us to believe their up regulating role in tumor formation but here the association of cldn5 as significant downregulated gene in ovarian cancer is the most surprising thing which reveals the fact that may be our knowledge about the claudins is still very limited and there are many other unraveled truths about the claudins that have to be identified yet. But the SAGE analysis of cldn5 revealed the totally different results by showing its up regulation in cancerous ovary. These findings about the cldn5 makes it role suspicious in ovarian cancer which must be analyzed in vitro.

5 Conclusion

Our work verified the previously known association of claudin members with ovarian cancer .In addition to that our systematic methodology has disclosed the high gene expression level of cldn16 in ovarian cancer which might play an important role in the cancer cell differentiation and proliferation. Furthermore cldn5 shows down regulated behavior verified by microarray analysis. This needs to be further confirmed, there may be a chance that cldn5 would serve as a novel biomarker for the treatment of ovarian cancer.

6 References

[1] Cannistra SA. "Cancer of the ovary"; N Engl J Med, Vol 351, Page number 2519–29, 2004.

[2] Morin PJ. **"**Claudin proteins in human cancer: promising new targets for diagnosis and therapy"; Cancer Res, Vol 65, Page numbers 9603-9606, 2000.

[3] Swisshelm K, Macek R, Kubbies M. "Role of claudins in tumorigenesis"; Adv Drug Deliv Rev, 57, 919-928, 2005.

[4] Nakanishi K, Ogata S, Hiroi S, et al. "Expression of Occludin and Claudins 1, 3, 4, and 7 in Urothelial Carcinoma of the Upper Urinary Tract"; Am J Clin Pathol, Vol 130, Page numbers 43-49, 2008.

[5] Dhawan P, Ahmad R, Chaturvedi R, et al. "Claudin-2 expression increases tumorigenicity of colon cancer cells: role of epidermal growth factor receptor activation"; Oncogene, (7 March 2011) | doi:10.1038/onc.2011.43.

[6] sukita S, Furuse M. "Pores in the wall: claudins constitute tight junction strands containing aqueous pores"; J Cell Biol, Vol 149, Page numbers 13-16, 2000.

[7] Heinzelmann-Schwarz VA, Gardiner-Garden M, Henshall SM., et al. "Overexpression of the Cell Adhesion Molecules DDR1 and Ovarian Cancer Claudin 3, and Ep-CAM in Metaplastic Ovarian Epithelium"; Clin Cancer Res, Volume 10, Page numbers 4427-4436, July 6, 2004.

[8] Boylan K.L.M., Misemer B, DeRycke MS., Andersen JD, et al. "Claudin 4 is differentially expressed between ovarian cancer subtypes and plays a role in spheroid formation"; International Journal of Molecular Sciences, Volume 12, Issue 2, Pages 1334-1358, February 2011.

[9] Hough CD, Sherman-Baust CA, Pizer ES, Montz FJ, Im DD, Rosenshein NB, Cho KR, Riggins GJ, Morin PJ. "Large-Scale Serial Analysis of Gene Expression Reveals Genes Differentially Expressed in Ovarian Cancer**";** Cancer Research, Vol 60, Page numbers 6281-6287, 2000.

[10] Rangel LBA, Agarwal R, D'Souza T, Pizer ES, Alò PL, Lancaster WD, Gregoire L, Schwartz DR, Cho KR, Morin PJ. "Tight Junction Proteins Claudin-

3 and Claudin-4 Are Frequently Overexpressed in Ovarian Cancer but Not in Ovarian Cystadenomas";Clin Cancer Res, Vol 9, Page number 2567-2575, 2003.

[11] Hewitt KJ**,** Agarwal R and Morin PJ. "The claudin gene family: expression in normal and neoplastic tissues"; BMC Cancer, 6:186, 2006.

[12] Bignotti E, Tassi RA., Calza S, et al. "Differential gene expression profiles between tumor biopsies and short-term primary cultures of ovarian serous carcinomas: Identification of novel molecular biomarkers for early diagnosis and therapy".; Gynecologic Oncology, Volume 103, Issue 2, Pages 405-416, November 2006.

[13] Rangel LBA, Agarwal R, D'Souza T, et al. "Tight junction proteins claudin-3 and claudin-4 are frequently overexpressed in ovarian cancer but not in ovarian cystadenomas"; Clin Cancer Res, Vol 9, Page numbers 2567–75, 2003.

[14] Boon K, Osorio EC, Greenhut SF et al. "An anatomy of normal and malignant gene expression"; PROC NATL ACAD SCI, Volume 99, Issue 17, Page numbers 11287-92, Aug 20, 2002.

[15] Hendrix ND, Wu R, Kuick R, Schwartz DR et al. "Fibroblast growth factor 9 has oncogenic activity and is a downstream target of Wnt signaling in ovarian endometrioid adenocarcinomas"; Cancer Res, Vol. no. 66, Issue no. 3, Page numbers (1354-62), Fe