Regulation of Protocadherin-15 In Glioma Progression

NILESH DEOKATE

A dissertation submitted for the partial fulfilment of the BS-MS dual degree in Biological sciences.

Indian Institute of Science Education and Research Mohali
April 2020

Certificate of Examination

This is to certify that the dissertation titled **"Regulation of Protocadherin-15 in Glioma Progression"** submitted by **Nilesh Deokate** (Reg. No. MS15108) for the fulfilment of **BS-MS dual degree programme** of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Sudip Mandal Dr. Lolitika Mandal Dr. Sabyasachi Rakshit
(Supervisor)

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Dr. Sabyasachi Rakshit at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography

Nilesh V Deokate

April, 2020

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Sabyasachi Rakshit

April, 2020

Acknowledgements

I would like to express my sincere and deepest gratitude to my supervisor, **Dr. Sabyasachi Rakshit**, for his guidance, remarks and encouragement throughout the duration of my stay at IISER Mohali and my master's thesis. Without his persistent help, the goal of this project would not have materialized.

I would like to thank **Dr. Malay and Dr. Varinder** for endless assistance and motivation, without which this would not have been possible. I would also thank Sai for helping me throughout this project.

I would also thank my lab member Gayatri, Jagadish, Surbhi, Nisha, Tanuja and Gaurav for teaching various molecular biology techniques and keeping lab environment fun and enjoyable.

Last but not the least, I would like to thank my parents and my sister for always believing in me and encouraging me to explore science.

List of Figures

1. Chapter 1

2. Chapter 2

3. Chapter 3

Abbreviations

Abbreviations	*Definitions*
2-HG	2-Hydroxyglutarate
bp	Base Pairs
CNV	Copy Number Variations
CSC	Cancer Stem Cells
CTCF	CCCTC-binding Factor
GBM	Glioblastoma Multiforme
HGG	High Grade Glioma
LGG	Lower Grade Glioma
MT	Mutant
TET	Ten-Eleven Translocation
WT	Wild Type

Preface

Due to high mortality rate around the world glioma is a hot topic for cancer researchers. Although Lower grade glioma (LGG) and Glioblastoma Multiforme (GBM) are glial cell cancer, their mode of progression and survival rate is significantly different than each other. To understand their mode of progression, it is important to study these gliomas at molecular level. In this thesis we show how Protocadherin-15, a non-classical cadherin family protein, is differentially expresses between LGG and GBM tumors. We also highlighted how its expression is regulated at genomic and epigenetic level.

Contents

Chapter 1

Introduction

Gliomas are the primary tumors of the central nervous system and known to cause more years of life loss than any other tumors. Tumors of the central nervous system account for around 2% of all cancers with around 3 lakh new cases diagnosed worldwide in the year 2018 (Globocan 2018)[1]. The incidence of central nervous system (CNS) tumors in India ranges from 5 to 10 per 100,000 population with an increasing trend and accounts for 2% of malignancie[2]. Majority of CNS tumors arise in glial cells that support and nourish the neurons and are known as Gliomas. The World Health Organization system (WHO) has first classified gliomas into four grades – Grade I to IV. Grades I and II are known as Low-Grade Gliomas (LGG), Grade III is called as High-Grade Glioma (HGG) and Grade IV, the malignant form of gliomas are called as Glioblastoma Multiforme (GBM). The GBMs are categorized into two types: primary and secondary GBMs . The primary GBMs develop rapidly de novo without any clinical evidence of lower stages, whereas secondary GBMs are transformed from lower stages (LGG to HGG) of gliomas. Primary GBMs account for 90% of GBMs and occur in elderly subjects (mean age 62 years). Secondary GBMs which account less in number, however, appear at a younger age (mean age 45 years)[3,4]. Genomic and morphological analysis of LGG and secondary GBM have shown significant differences. However, the key factors that regulate the differences and convert a least malignant with better prognosis cancer to the most devastating one are still not apparent. In-depth knowledge on the tumorigenic process of LGGs and deciphering the key factors that trigger the conversion of LGG to GBM via HGG will undoubtedly open up new directions in the diagnosis and prognosis of GBM and improve the overall survival of all GBM patients.

One important characteristic of GBM (both primary and secondary) is loss of 10q genomic region in patients[5]. The 10q region harbors many genes like PTEN, a tumor suppressor gene and its loss is considered to aggravate the carcinogenesis. Protocadherin-15 (PCDH15), a long-chain non-classical cadherin-class of proteins is also present in 10q genomic region. Based on mRNA expression as estimated from TCGA, PCDH15 expression is significantly upregulated in LGG-II than normal astrocytes, and steeply downregulated as the tumor progresses from LGG-II to GBM via LGG-III. The overall expression of PCDH15 is significantly low in GBM. Moreover, the subjects having higher PCDH15 expression showed better survival than those having lower PCDH15 expression. Therefore,

the loss of *10q* leads to reduced PCDH15 mRNA expression in GBM. In this study we aimed to decipher the role and regulation of PCDH15 in pathology of gliomas using bioinformatics analysis.

Chapter 2

Regulation of Protocadherin-15 during glioma progression

Methods

PCDH15 expression data for glioma patients was obtained from **GEPIA**[6] and Cbioportal[7,8] web server. GEPIA provides gene expression data of cancer and paired normal tissues. GEPIA uses data from TCGA (The Cancer Genome Atlas) and cancerous tissue gene expression and GETX data for normal tissue gene expression. Glioma patient's overall survival data with respect to PCDH15 expression was obtained from **cbioportal** database. To identify correlated genes Top upregulated and downregulated genes were extracted from GEPIA datasets. PCDH15 copy number alteration and methylation data was obtained from cbioportal. CpG island in the promoter region of PCDH15 was identified by using emboss **cpgplot**[9] server. **CTCFBSDB 2.0**[10,11] server was used to identify the potential CTCF binding sites present in PCDH15 sequence**.**

Statistical analysis

Unpaired two-tailed *t*-test, non-parametric Mann–Whitney test, ANOVA, and survival analysis using Log-rank (Mantel–Cox) test were done with graphpad prism 5 (Graph Pad Software, San Diego, CA, USA). Two-sided chi-square/Fisher tests were used to compare the proportion of patients in TCGA with regard to PCDH15 expression. A *P*-value of < 0.05 was denoted as a statistically significant difference.

Results

PCDH15 expression decreases as glioma progresses to higher grade

Due to vast differences in glioma progression and aggressiveness among primary and secondary GBM, we wanted to check the differential expression of genes in LGG and GBM. We used GEPIA overexpressed and under expressed genes data of LGG and GBM tumors. We selected genes which are differentially expressed between these tumors and then plot their expression. We found that PCDH15 is highly expressed in LGG tumors while its expression is downregulated in GBM tumors (Fig. 2.1a).

Tumor eventually progresses from lower stages to highly aggressive advanced stages. To identify the change in PCDH15 expression as tumor progresses, we plot PCDH15 expression with respect to tumor grade. We observed that PCDH15 expression decreases gradually and significantly. Further, during early stage of tumor formation in secondary GBM, the PCDH15 expression is very high. Overall, the expression trend indicates that PCDH15 might be involved in glioma initiation but not in glioma progression (fig.2.1b).

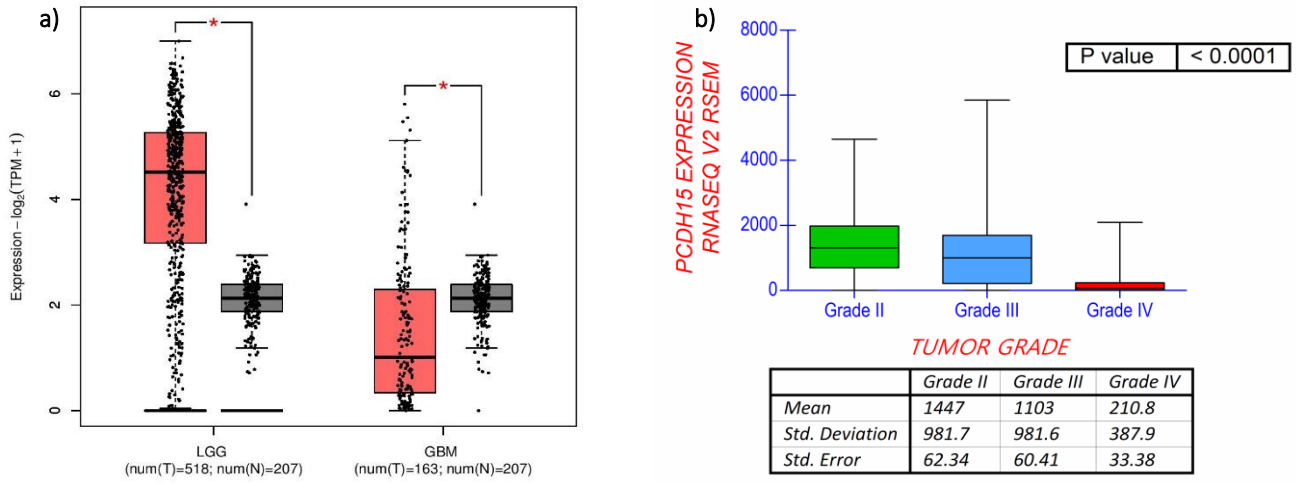


Figure 2.1: Expression pattern of PCDH15 in glioma. a) Differential expression of PCDH15 in LGG and GBM patients compared to control subjects. b) PCDH15 expression in different stages of glioma.

Patients having high expression of PCDH15 shows better overall survival

We analysed the overall survival of glioma patients with respect to the PCDH15 expression in LGG patients. Those patients having PCDH15 expression above median value were grouped as high and below median were grouped as low. From survival plot we found that patients having high PCDH15 expression showed better survival (3200 days) compared to the patients having low PCDH15 expression (1585 days) (Fig 2.2a). Similarly, we plot survival curve for GBM patients, where we did not find any significant correlation between PCDH15 expression and GBM patient survival (Fig 2.2b). This could be due to overall low expression of PCDH15 in all GBM patients than the normal healthy expression.

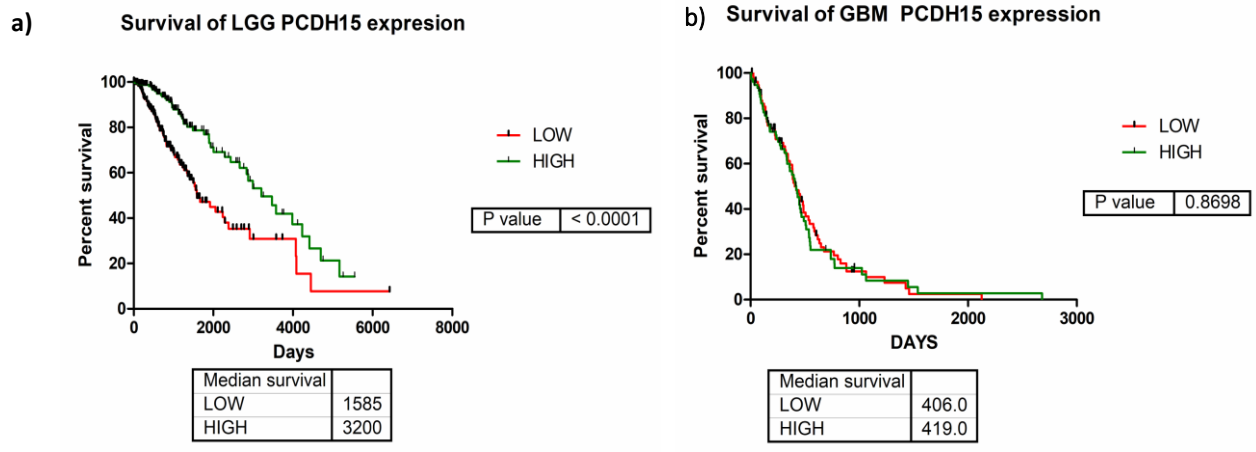
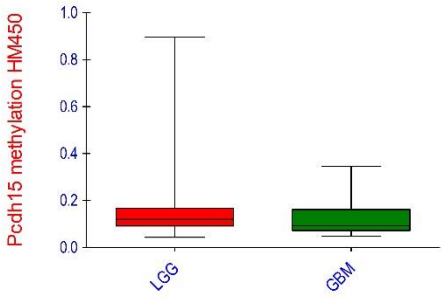


Figure 2.2: Effect of PCDH15 expression on glioma patients survival. Survival curve for LGG (a) and GBM (b) patients with respect to PCDH15 expression.

Methylation does not affect the PCDH15 expression in glioma

Regulations of cell-cell junction by modulating the expression of cadherins is common in solid cancers. Among cadherins, neuroepithelials usually express non-classical cadherins predominantly. As mentioned previously, PCDH15 is expressed in astrocytes and regulated differentially in different stages of glioma. We, therefore, asked what regulates the expression of PCDH15 in LGG and GBM tumors. Predominantly the expression of cadherins are regulated by DNA methylation in most cancers.[12] We, therefore, analysed the methylation data of LGG and GBM patients for PCDH15 gene using TCGA cbioportal. PCDH15 expression was correlated with methylation level of PCDH15 locus in LGG and GBM tumors using TCGA HM450 β-value. Higher β-value corresponds to a hyper methylation of locus while lower beta value is associated with hypo methylation of DNA locus. In LGG and GBM tumors, we did not notice any such correlation between PCDH15 expression and DNA methylation β-value (Fig. 2.3). Both tumors show hypo methylation of PCDH15 locus indicating that the decreased expression of PCDH15 in GBM is not controlled by methylation of PCDH15 locus.



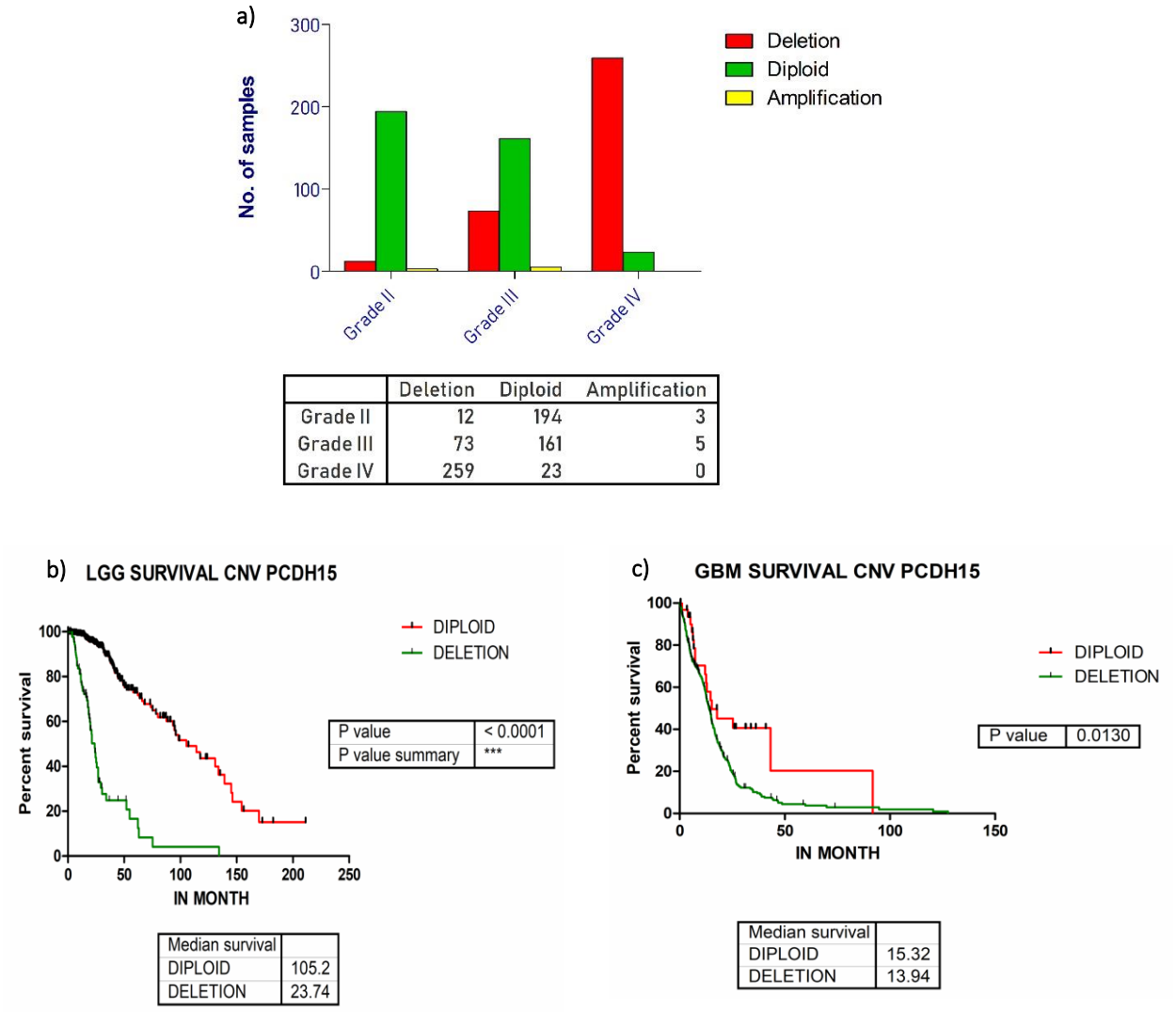
	LGG	GBM
Mean	0.1438	0.1169
Std. Deviation	0.08232	0.05927

Figure 2.3: Overall methylation status of PCDH15 gene in LGG and GBM tumors. DNA methylation levels of PCDH15 in LGG and GBM patients with respect to PCDH15 expression. Number of LGG patients is 516 and number of GBM patients is 60

Shallow deletion of PCDH15 locus is responsible for reduced PCDH15 expression in GBM.

Since DNA methylation was not responsible for PCDH15 expression alteration in LGG and GBM, we looked for other reasons for altered PCDH15 expression. Genomic alteration such as gene deletion, amplification and rearrangement also causes change in gene expression in various type of cancers[13,14]. As tumor progresses to the more advanced stages, the chromosomal abbreviations start to occur more frequently which leads to Copy number variation (CNV) such as gene duplication and

gene deletion[15]. These copy number variations alter the gene expression profile. So we checked for copy number variation difference between LGG and GBM tumor and we found that shallow deletion of PCDH15 gene takes place as tumor grade increases (fig.2.4a). This implied that the decreased PCDH15 expression in higher grade glioma is due to shallow deletion of PCDH15 locus occurring in patients. This CNV only explains the downregulation of PCDH15 in GBM but, LGG patients neither showed shallow deletion of PCDH15 gene locus nor gain in copy number so what causes high PCDH15 expression in LGG remains unknown. Further, we also observed that the patients having the genomic alterations showed worst tumor prognosis (fig. 2.4b). The LGG patients having normal diploid PCDH15 locus shows better overall survival (105 months) compare to the patients having shallow deletion of PCDH15 locus (23.7 months).

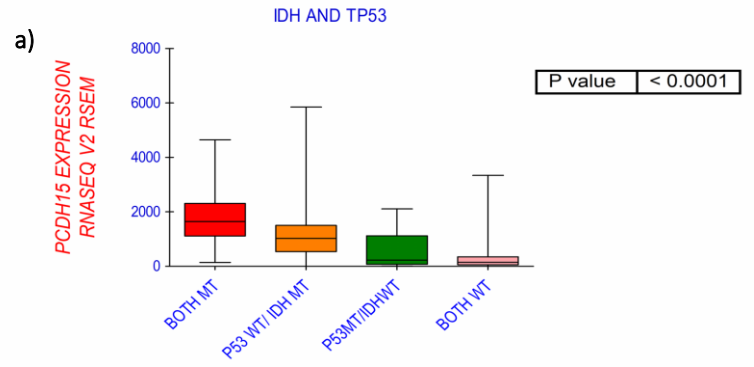


	Deletion	Diploid	Amplification
Grade II	12	194	3
Grade III	73	161	5
Grade IV	259	23	0

Figure 2.4: Copy number variations of PCDH15 locus. a) Copy number Variations in PCDH15 locus in different stages of glioma. b and c) overall survival of LGG and GBM patients compared to the copy number variation in PCDH15 gene.

IDH1 and *TP53* mutations regulate PCDH15 expression in LGG

In LGG tumours the IDH1 and TP53 mutations occur predominantly. Mutant IDH1 causes hyper-methylated histones and DNA which alters gene expression pattern while mutant TP53 shows gain of function and increases non-specific genes expression than normal functional TP53[16,17]. Therefore, we checked whether these mutations cause any direct or indirect effect on PCDH15 expression. LGG patients PCDH15 expression as well as IDH1 and TP53 mutation status data was extracted from cbioportal. We plot the PCDH15 expression of LGG patients with respect to their IDH1 and TP53 mutation status. We observed that the IDH1 mutant and TP53 mutant patients showed high expression of PCDH15 as compared to patients having wild type IDH1 and TP53 (fig. 5). IDH1 and TP53 mutations coexisted in many LGG patients. To check which mutation exerts the dominant effect on PCDH15 expression, we plot the graph between PCDH15 expression and IDH1 and TP53 status of LGG patients. The patients having IDH1 mutation alone showed high expression of PCDH15 compare to those patients having only TP53 mutation (fig 2.5a). This suggested that compared to TP53 mutations, in IDH1 gene mutations have more impact on PCDH15 expression.



	BOTH MT	P53 WT/ IDH MT	P53MT/IDHWT	BOTH WT
Mean	1797	1135	648.1	330.1
Std. Deviation	907.9	831.4	675.9	527.3
Std. Error	59.60	65.32	163.9	53.81

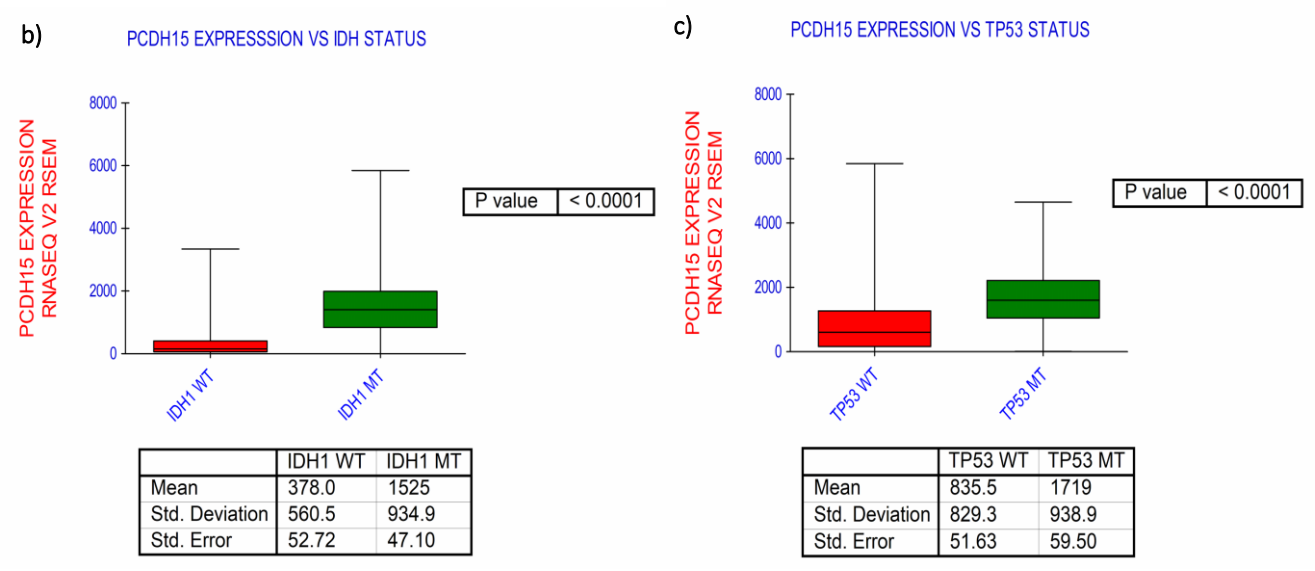


Figure 2.5: PCDH15 expression levels w.r.t. mutations in IDH1 and TP53 in glioma patients. a) PCDH15 expression w.r.t. mutations in idh1 and tp53 in LGG patients. b) PCDH15 expression w.r.t. mutant idh1 status in LGG patients. C) PCDH15 expression w.r.t. mutant tp53 status in LGG patients.

Disrupted CTCF binding regulates PCDH15 expression in IDH1 mutant gliomas

To decipher the mechanism of *IDH1* mutation mediated regulation of the PCDH15 expression we did literature study. We found that The *IDH1* mutation is neomorphic, i.e., IDH1 mutant produces a new oncometabolite 2-Hydroxyglutarate (2-HG) which interferes with iron-dependent hydroxylases, including the TET (Ten-Eleven Translocation) family of 5′-methylcytosine hydroxylases. These TET enzymes are involved in the removal of DNA methylation. Due to compromised activity of TET enzymes in *IDH1* mutant gliomas, these patients show glioma hyper methylation phenotype. The CpG islands in *IDH1* mutant tumours show higher methylation compared to the *IDH1*-WT tumours[18]. Recent study shows that if CTCF (CCTC-binding Factor) binding site is present in CpG island then in *IDH1* mutant tumour due to hyper methylation of CpG island the CTCF binding get disrupted[19]. This Disruption of CTCF binding causes remodelling of 3D chromatin structure and alters the gene expression pattern.

To check if there is CTCF binding present in upstream of PCDH15 gene we used CTCFBSDB 2.0 database (fig. 2.6a). We also looked for CpG island near PCDH15 locus. For this, we take first 3500bp nucleotides sequence of PCDH15 from NCBI and run it through emboss cpgplot to identify the CpG island present in sequence. We found that the CpG island of length 382 bp existed between nucleotide 521 to 902 bp. CTCFBSDB 2.0 also predicted the CTCF binding site in the same region (fig. 2.6b). This data indicated that in LGG type of gliomas due to IDH1 mutation CTCF binding site at PCDH15 locus get disrupted which leads chromatin remodelling and overexpression of PCDH15.

a)

Motif PWM	Motif Sequence	Input Sequence Name	Motif Start Location	Motif Length	Motif Orientation	Score
EMBL_M1	CGCCTCCCGCTGGA	TGGAAATAGTTGCAGAATCCTCAGAAAATATCACTTATAGAAGGGTATTAATATAAGAAATACAACTTCC	777	14	-	12.9425
EMBL_M2	AGAACTGCC	TGGAAATAGTTGCAGAATCCTCAGAAAATATCACTTATAGAAGGGTATTAATATAAGAAATACAACTTCC	417	9	+	12.1136
REN_20	TCTGACAGGAGGTGGAGCTC	TGGAAATAGTTGCAGAATCCTCAGAAAATATCACTTATAGAAGGGTATTAATATAAGAAATACAACTTCC	7574	20	+	11.4147
MIT_LM2	AGTCCACAGGATGCCACTA	TGGAAATAGTTGCAGAATCCTCAGAAAATATCACTTATAGAAGGGTATTAATATAAGAAATACAACTTCC	5171	19	+	13.3628
MIT_LM7	CTGACAGGAGGTGGAGCTCA	TGGAAATAGTTGCAGAATCCTCAGAAAATATCACTTATAGAAGGGTATTAATATAAGAAATACAACTTCC	7575	20	+	16.5881
MIT_LM23	CTGACAGGAGGTGGAGCTCA	TGGAAATAGTTGCAGAATCCTCAGAAAATATCACTTATAGAAGGGTATTAATATAAGAAATACAACTTCC	7575	20	+	14.5339

b)

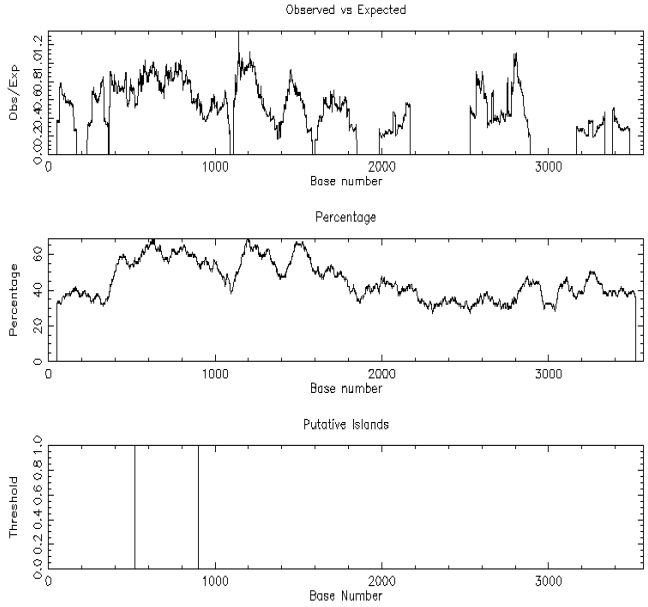


Figure 2.6: Predicted CpG island and CTCF binding sites in PCDH15 gene. a). CTCF binding motifs identified in the PCDH15 gene using CTCFBSDB 2.0 database. b) CpG island identified in PCDH15 gene using cpgplot tool.

Chapter 3

Conclusion

GBM is the most lethal among all cancers, with a median survival rate of less than 12 months. The standard treatment protocols are complex, and consists of surgical resection followed by radiation therapy and concurrent chemotherapy. Even then, the rate of recurrence of GBM is high, and no notable improvement in the survival rates of GBM patients have been observed in population statistics in the last three decades (Book on Glioblastoma chapter 8 by Ahmad Faleh Tamimi and Malik Juweid). It has therefore been an important topic of research in all fields of cancer biology including epidemiology, cell-biology, meta-analysis, gene-expression, and more. Even WHO has recently restructured the classifications gliomas based on gene-expression and mutation data. It has, therefore, become imperative to identify the major genes that are significantly regulated in gliomas and decipher their roles using multidisciplinary experimental approaches.

Cells rely strongly on cell-to-cell adhesion and cell-to-extracellular interactions for tissue integrity. Cadherins proteins are among the proteins mediating cell-to-cell adhesions. Tumour cells undergo alterations in expression of cadherins that lead to reduced cell to cell adhesion thereby promoting metastasis. Our analysis from TCGA indicated that PCDH15, a non-classical cadherin-class of cell-adhesion protein, is most significantly regulated in gliomas than normal astrocytes, higher in low-grade glioma (LGG) and lowest in GBM (last grade glioma). As during early stage of glioma the increasing PCDH15 expression suggest that PCDH15 work as oncogene and required for tumor initiation. At same time the decrease in PCDH15 expression as glioma progresses, better overall survival of PCDH15 overexpressing patients and shallow deletion of PCDH15 locus indicates that PCDH15 might be a tumor suppressor gene. Therefore, more in-vitro and in-vivo studies will required to revel the true nature of PCDH15 in glioma.

In this thesis using bioinformatics study we shows that in LGG type of tumor the predominantly occurring IDH1 mutation disrupts the CTCF binding which leads to chromatin remodelling and thus increases the PCDH15 expression. In HGG and GBM tumor the shallow deletion of PCDH15 locus at genomic level is responsible for the downregulation of PCDH15 expression.

Notably, PCDH15 plays a significant role in developing the kinocilium and stereocilia in the inner ear. An orthologue of PCDH15, Cad99C, in Drosophila melanogaster significantly contributes to the development of microvilli in follicle cells[20]. The high expression of PCDH15 is also associated with the number as well as the morphologies of the processes in astrocytes[21]. It is, therefore, imperative to hypothesize that the high-expression of PCDH15 in LGG helps the astrocytes to

maintain well-developed processes that lack in GBM. The processes help astrocytes in LGG to maintain distinct cell-cell junctions and so the vestibular integrity, thus making them less hypoxic, less stiff, least proliferative, and least invasive than GBM. Interestingly, the hypoxic and the perivascular nature are the niches for CSC in GBM.

Bibliography

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424. doi:10.3322/caac.21492
2. Dasgupta A, Gupta T, Jalali R. Indian data on central nervous tumors: A summary of published work. *South Asian J Cancer*. 2016;5(3):147. doi:10.4103/2278-330x.187589
3. Ohgaki H, Kleihues P. Population-Based Studies on Incidence, Survival Rates, and Genetic Alterations in Astrocytic and Oligodendroglial Gliomas. *J Neuropathol Exp Neurol*. 2005;64(6):479-489. doi:10.1093/jnen/64.6.479
4. Ohgaki H, Dessen P, Jourde B, et al. Genetic pathways to glioblastoma: A population-based study. *Cancer Res*. 2004;64(19):6892-6899. doi:10.1158/0008-5472.CAN-04-1337
5. Ohgaki H, Kleihues P. Genetic pathways to primary and secondary glioblastoma. *Am J Pathol*. 2007;170(5):1445-1453. doi:10.2353/ajpath.2007.070011
6. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Web Serv issue Publ online*. 2017;45. doi:10.1093/nar/gkx247
7. Cerami E, Gao J, Dogrusoz U, et al. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2(5):401-404. doi:10.1158/2159-8290.CD-12-0095
8. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6(269). doi:10.1126/scisignal.2004088
9. Abio Madeira F′, Mi Park Y, Lee J, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Web Serv issue Publ online*. 2019;47. doi:10.1093/nar/gkz268
10. Ziebarth JD, Bhattacharya A, Cui Y. CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization. doi:10.1093/nar/gks1165
11. Bao L, Zhou M, Cui Y. CTCFBSDB: a CTCF-binding site database for characterization of vertebrate genomic insulators. *Nucleic Acids Res*. 2008;36:83-87. doi:10.1093/nar/gkm875
12. Berx G, van Roy F. Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harb Perspect Biol*. 2009;1(6). doi:10.1101/cshperspect.a003129
13. Chakravarthi BVSK, Nepal S, Varambally S. Genomic and Epigenomic Alterations in Cancer. *Am J Pathol*. 2016;186(7):1724-1735. doi:10.1016/j.ajpath.2016.02.023
14. Krohn A, Diedler T, Burkhardt L, et al. Genomic deletion of PTEN is associated with tumor progression and early PSA recurrence in ERG fusion-positive and fusion-negative prostate cancer. *Am J Pathol*. 2012;181(2):401-412. doi:10.1016/j.ajpath.2012.04.026

15. Shlien A, Malkin D. Copy number variations and cancer. *Genome Med*. 2009;1(6):62. doi:10.1186/gm62
16. Raineri S, Mellor J. IDH1: Linking Metabolism and Epigenetics. *Front Genet*. 2018;9. doi:10.3389/fgene.2018.00493
17. Rivlin N, Brosh R, Oren M, Rotter V. Mutations in the p53 tumor suppressor gene: Important milestones at the various steps of tumorigenesis. *Genes and Cancer*. 2011;2(4):466-474. doi:10.1177/1947601911408889
18. Turcan S, Rohle D, Goenka A, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature*. 2012;483(7390):479-483. doi:10.1038/nature10866
19. Flavahan WA, Drier Y, Liau BB, et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature*. 2016;529(7584):110-114. doi:10.1038/nature16490
20. D'Alterio C, Tran DDD, Au Yeung MWY, et al. Drosophila melanogaster Cad99C, the orthologue of human Usher cadherin PCDH15, regulates the length of microvilli. *J Cell Biol*. 2005;171(3):549-558. doi:10.1083/jcb.200507072
21. Li J, Khankan RR, Caneda C, et al. Astrocyte-to-astrocyte contact and a positive feedback loop of growth factor signaling regulate astrocyte maturation. *Glia*. 2019;67(8):1571-1597. doi:10.1002/glia.23630