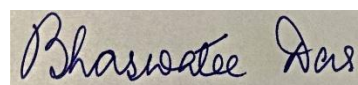


Dedicated to

"Maa and Pitai"

Declaration

I hereby declare that the thesis entitled “**Studies on immune related genes with genomic instability in cancer.**” is an authentic work carried out by me under the supervision of Prof. Shashi Baruah, Department of molecular Biology and Biotechnology, Tezpur University, Assam-784028. No part of this has been presented for any other degree or diploma earlier.

A rectangular box containing a handwritten signature in dark ink. The signature is written in a cursive style and reads "Bhaswatee Das".

Place: Tezpur

(Bhaswatee Das)

Date: 17-03-2025



TEZPUR UNIVERSITY
(A Central University by an Act of Parliament)
Napaam, Tezpur-784028
District: Sonitpur, Assam, India

CERTIFICATE OF SUPERVISOR

This is to certify that the thesis entitled “**Studies on immune related genes with genomic instability in cancer.**” submitted to the School of Sciences, Tezpur University in requirement of partial fulfilment for the award of the degree of Doctor of Philosophy in Molecular Biology and Biotechnology is a record of research work carried out by **Ms. Bhaswatee Das** under my supervision and guidance.

All help received by her from various sources have been duly acknowledged. No part of this thesis has been submitted elsewhere for award of any other degree.

(Signature of supervisor)

Date:

Place: Tezpur

Name: Dr. Shashi Baruah

Designation: Professor

School: Sciences

Department: Molecular Biology and Biotechnology, Tezpur University.

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Prof. Shashi Baruah, for her unwavering support and encouragement throughout my Ph.D. journey. Her invaluable guidance, advice, and motivation have not only played a crucial role in shaping my research work but have also empowered me to find my voice and take a stand for myself in life. Her mentorship has been instrumental in my personal and professional growth, and I am truly grateful for her presence in my academic journey.

I am deeply thankful to the Honorable Vice-Chancellor of Tezpur University for granting me the opportunity to pursue my Ph.D. and carry out my research at this esteemed institution.

I extend my heartfelt appreciation to my Doctoral Committee members – Prof. S. K. Ray, Dr. Aditya Kumar – for their invaluable suggestions and insightful inputs during my Ph.D. tenure.

I am grateful to the Head, Department of Molecular Biology and Biotechnology at Tezpur University for his constant support and encouragement throughout my Ph.D. journey. I am also grateful to all the faculty members and non-teaching staff of the department for their assistance and cooperation.

I am grateful for the financial support provided by Tezpur University and the University Grants Commission (UGC) (696/(PWD) (CSIR-UGC NET DEC. 2016)), New Delhi.

I would like to express my gratitude to my lab seniors – Neelanjana Ba, Saurav Da, Debashree Ba, and Akash Da for their continuous support and encouragement throughout my Ph.D. tenure.

I am thankful to my wonderful labmates, Mayuri, Dalal, Chourajit, and Sushmita, for creating a supportive and cheerful environment in the lab. I am forever thankful to my friends- Anupama, Krishna, Archana and Subrata for making this journey so cheerful and jolly. They all have been there with me in all my ups and downs throughout my Ph.D. journey and I adore them a lot for being so awesome human being.

I would like to acknowledge and appreciate my Ph.D. mates – Panchi, Amit, Nitisha, Bhabna, PK, Joydeep and Khalifa – for their companionship and support throughout this journey.

I am forever indebted to Pitai (Mr. Khanindra Kumar Das), Maa (Mrs. Runu Das), my little brother Mr. Jyotishman Das for their unwavering support, motivation, unconditional love, and blessings. I cannot forget to mention my cats Babakon, mimi, Chotu who makes my existence worthwhile and bring so much happiness into my life. I am deeply grateful to Mr. Ripunjoy Borah for being there through all the challenges and joys.

With heartfelt gratitude, Bhaswatee Das

LIST OF FIGURES

	<u>CHAPTER I</u>	Page No.
Figure 1.1	A diagram showing strategies used by tumor for immune suppression	5
Figure 1.2	A schematic illustration depicting three phase of cancer immunoediting	6
Figure 1.3	Antigen processing and presentation pathway downloaded from KEGG pathways database	9
	<u>CHAPTER II</u>	
Figure 2.1	A diagram representing all 14 hallmarks of cancer	13
Figure 2.2	A schematic representation of the features of G-quadruplex (G4) structures in DNA that includes a depiction of four guanine-rich tracts (G-tracts) that form planar G-tetrads through Hoogsteen base-pairing	22
	<u>CHAPTER IV</u>	
Figure 4.1.1	Bar graphs depicting the number of samples with type of variations in the 24 cancers	35-39
Figure 4.1.2	Bar graphs depicting the number of mutated samples with polymorphisms, insertions and deletions in the 24 cancers	40-42
Figure 4.1.3	Bar graphs depicting the type of nucleotide changes in the 24 cancers.	43-45

Figure 4.1.4	Top 10 frequently mutated genes in the 24 different types of cancer	46-50
Figure 4.1.5	Copy number alterations in frequently mutated genes of the respective cancers.	53-59
Figure 4.1.6	Putative G4 locations and immune related genes along the human genome.	59
Figure 4.1.7	Box plot with anova ($p < 2.2e-16$) for comparing G4 density (G4/1000bp) in genes of immune genes, housekeeping genes and proto-oncogenes	60
Figure 4.1.8	Box plot with anova ($p < 2.2e-16$) for comparing G4 density (G4/1000bp) in promoters of immune genes, housekeeping genes and proto-oncogenes	61
Figure 4.1.9	Box plot with anova ($p < 2.2e-16$) for comparing G4 density (G4/1000bp) in promoter of different immune classes to proto-oncogenes	62
Figure 4.1.10	Box plot with anova ($p < 2.2e-16$) for comparing G4 density (G4/1000bp) in genes of different immune classes to proto-oncogenes	63
Figure 4.2.1	5year Kaplan Meier survival curves for mutation in PIK3CA in different cancer types	69-71
Figure 4.2.2	5year Kaplan Meier survival curves for mutation in TG in different cancer types	73-74
Figure 4.2.3	Kaplan Meier survival curves for mutated versus wild type samples of cancers in 5 year survival time in 4 antigen presenting genes (1. HLA-A, 2. HLA-B, 3. HLA-DRB1, 4. CIITA) in our study.	76-82
Figure	Kaplan Meier survival curves for 4 antigen	84-100

4.2.4	presenting genes in our study (1.HLA-A, 2.HLA-B, 3. HLA-DRB1, 4. CIITA) considering expression.	
Figure 4.3.1	Pathways over-represented in PIK3CA mutated (a) UCEC, (b) LGG dataset	104
Figure 4.3.2	The complement and coagulation pathway, showing differentially expressed genes in the PIK3CA mutated dataset (a) UCEC, (b)LGG.	108
Figure 4.3.3	Difference in fractions of immune cell infiltration in mutated and non-mutated PIK3CA dataset of (a) UCEC, (b) LGG	110
Figure 4.3.4	Difference in fractions of immune cell infiltration in mutated and non-mutated TG dataset of BRCA	114
Figure 4.3.5	Figure 4.3.6: Differential gene expression of (a)RFX ((i)RFXANK, (ii)RFXAP, (iii)RFX5), (b) NFY((i)NFYA, (ii)NFYB, (iii)NFYC), (c)CREB, (d)HLA-E and (e)HLA-G, (f)HLA-A, (g)HLA-B, (h)HLA-DRB1(i)CIITA in tumor versus normal samples of LGG	115-117
Figure 4.3.6	Kaplan Meier survival curves for in (a)RFX ((i)RFXANK, (ii)RFXAP, (iii)RFX5), (b) NFY((i)NFYA, (ii)NFYB, and (iii)NFYC), (c)CREB1, (d)HLA-E and (e)HLA-G	119-120
	<u>CHAPTER V</u>	
Figure 5.1	A diagram representing changes in immune microenvironment with PIK3CA mutations in UCEC	134

Figure 5.2	A diagram representing changes in immune microenvironment with PIK3CA mutations LGG	134
Figure 5.3	A diagram representing changes in immune microenvironment with alterations in expression levels of HLA-A, HLA-B, HLA-DRB1, CIITA in LGG	135

LIST OF TABLES

	<u>CHAPTER III</u>	Page No.
Table 3.1	Details of the datasets downloaded from cBioportal	26-29
	<u>CHAPTER IV</u>	
Table 4.1.1	List of frequently mutated immune related genes divided into three categories according to biological function.	51
Table 4.1.2	Immune-related genes with frequent mutations and CNA both in the particular cancer	52-53
Table 4.1.5	Number of immune related genes with and without G4 both along the gene and in promoter	64
Table 4.1.6	Number of G4 with and without mutations in promoters and along the gene of immune related genes	64
Table 4.2.1	Cox-proportional hazard analysis for both mutated and non-mutated datasets of UCEC and LGG (a) Univariate (b) Multivariate	72
Table 4.2.2	Data from Kaplan Meier plots for survival considering mutations in the different cancer types for the four antigen presenting genes considered in the study	75
Table	Influence of change in expression levels of the 4	83-84

4.2.3	antigen presenting genes in our study on survival.	
Table 4.3.1	Differentially expressed genes of complement and coagulation pathway in PIK3CA mutated (A)UCEC and (B)LGG dataset.	106-107
Table 4.3.2	Correlation between differentially expressed genes and Tfh and M1 macrophages.	111
Table 4.3.3	Pathways over-represented in TG mutated BRCA dataset	112
Table 4.3.4	Correlation between differentially expressed genes and M1 macrophages in TG mutated BRCA dataset.	114
Table 4.3.5	Association between expression of the four antigen presenting genes in our study (a) HLA-A, (b) HLA-B, (c) HLA-DRB1, (d) CIITA and immune cell infiltration in LGG	121-122
Table 4.3.6	Correlation between 4genes (HLA-A, HLA-B, HLA-DRB1 and CIITA) and HLA-E	122
Table 4.3.7	Correlation among CIITA, RFXANK, RFXAP, RFX5, CREB, HLA-A, HLA-B, HLA-DRB1 and HLA-E. The cells marked in red could not attain significance	123

LIST OF ABBREVIATIONS

G4	G-quadruplexes
HLA	Human leukocyte antigen
NK	Natural killer cell
IL-10	Interleukin 10
TGF- β	Transforming growth factor-beta
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
PD-1	Programmed Cell Death Protein 1
PD-L1	Programmed Cell Death Ligand 1
Tregs	Regulatory T cells
TAM	Tumor associated Macrophages
MDSC	Myeloid derived suppressor cells
CSF1	Colony stimulating factor 1
CCL2	Chemokine (C-C motif) ligand 2
CCL22	Chemokine (C-C motif) ligand 22
CCL5	Chemokine (C-C motif) ligand 5
CXCL5	C-X-C motif chemokine 5
CXCL8	C-X-C motif chemokine 8
KRAS	Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
BRAF	B-Raf
CASP8	Cysteine-aspartic acid protease
CNA	Copy Number Alteration
TG	Thyroglobulin
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PIK3/AKT	Phosphoinositide-3-kinase–protein kinase B/Akt
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
MHC	Major Histocompatibility complex
HLA-A	Human leukocyte antigens A

HLA-B	Human leukocyte antigens B
HLA-C	Human leukocyte antigens C
HLA-DR	Human leukocyte antigens DR
HLA-DP	Human leukocyte antigens DP
HLA-DQ	Human leukocyte antigens DQ
TNF	Tumor necrosis factor
C3	Complement component 3
C4	Complement component 4
C5	Complement component 5
HLA-DRB1	Human leukocyte antigens DRB1
CIITA	CIITA class II major histocompatibility complex transactivator
NSCLC	Non-small cell lung carcinoma
TNF- α	Tumor necrosis factor alpha
IFN- γ	Interferon-gamma
HLA-G	Human leukocyte antigens G
HLA-E	Human leukocyte antigens E
IL-10	Interleukin 10
TGF- β	Transforming growth factor-beta
IDH1	Iocitrate Dehydrogenase 1
IDH2	Isocitrate Dehydrogenase 2
STING	Stimulator of Interferon Genes
B2M	beta-2 microglobulin
TME	Tumor microenvironment
TIME	Tumor immune microenvironment
CAR-T	Chimeric Antigen Receptor T-cell
TCGA	The Cancer Genome Atlas
ICGC	International Cancer Genome Consortium
ACC	Adrenocortical carcinoma
BLCA	Bladder Urothelial Carcinoma
BRCA	Breast invasive carcinoma
CHOL	Cholangiocarcinoma

COAD	Colorectal adenocarcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
HCC	Hepatocellular carcinoma
HNSC	Head and Neck Squamous cell Carcinoma
KICH	Kidney Chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
OV	Ovarian carcinoma
PRAD	Prostate adenocarcinoma
ESCA	Esophageal adenocarcinoma
THCA	Thyroid carcinoma
UCEC	Uterine Corpus Endometrial Carcinoma
GBM	Glioblastoma multiforme
LGG	Brain Lower Grade Glioma
PCPG	Pheochromocytoma and Paraganglioma
DLBC	Diffuse large B cell lymphoma
LAML	Acute Myeloid Leukemia
SARCOMA	Sarcoma
GISTIC	Genomic Identification of Significant Targets in Cancer
DEG	Differentially expressed genes
GEPIA	Gene Expression Profiling Interactive Analysis
COSMIC	COmmon Software Measurement International Consortium
TIMER	Tumor Immune Estimation Resource
TCR	T-cell receptor
BCR	B-cell receptor
PQRS	putative G4
M1	M1 macrophages
ACE2	Angiotensin Converting Enzyme 2
IDO1	Indoleamine 2,3-Dioxygenase 1

IDO2	Indoleamine 2,3-dioxygenase 2
CD8+	Cytotoxic T-cells
CD4+	Helper T-cells
B cell	Plasma cells
M0	Undifferentiated macrophages
M1	M1 macrophages
M2	M2 macrophages
DC	Dendritic cells
MAC	Membrane attack complex
KNG1	Kininogen 1
F2	Coagulation factor II, thrombin
FGA	Fibrinogen alpha chain
FGB	Fibrinogen beta chain
FGG	Fibrinogen gamma chain
SERPIND1	Serpin family D member 1
VTN	Vitronectin
CFHR5	Complement Factor H Related 5
C3P1	Complement Component 3 Precursor Pseudogene
CFH	Complement Factor H
CFHR3	Complement Factor H Related 3
F13A1	Coagulation Factor XIII, A Polypeptide
BDKRB2	Bradykinin Receptor B2
C6	Complement C6
C7	Complement C7
CR1	Complement C3b/C4b Receptor 1
C4BPB	Complement Component 4 Binding Protein Beta
RFXANK	regulatory factor X-associated protein-containing ankyrin repeat
RFXAP	Regulatory factor X-associated protein
NFYA	Nuclear transcription factor Y subunit alpha
NFYB	Nuclear transcription factor Y subunit beta
NFYC	Nuclear transcription factor Y subunit gamma

CREB	cAMP-response element binding protein
RFX5	Regulatory factor X5
Th1	Type 1 T helper cells
C3b	C3b is the larger of two elements formed by the cleavage of complement component 3
C5a	Complement C5a Anaphylatoxin
C5aR	Complement component 5a receptor 1