CONTENTS

Contents	Page No.
Abstract	i-iii
Declaration by candidate	iv
Certificate of Supervisor	v
Acknowledgment	vi-vii
Table of contents	viii-xiii
List of Figures	xiv-xxiii
List of Tables	xxiv-xxvi
List of Abbreviations	xxvii-xxix

Sl.No.	Table of Contents	Page No.
Chapter I: Introduction		
1.1	Gallbladder: Structure and function	1
1.2	Carcinoma of the gallbladder: Definition and pathophysiology	1-2
1.3	Etiological factors associated with GBC	2-4
1.3.1	1.3.1 Cholelithiasis and Cholecystitis	2-3
1.3.2	1.3.2 Gender disparity	3
1.3.3	1.3.3 Gallbladder pathologies	3
1.3.4	1.3.4 Environmental exposures	3-4
1.4	Clinical presentation and pathological features of GBC	4
1.4.1	Clinical manifestations	4
1.4.2	1.4.2 Pathological features of GBC	4-5
1.5	Management of GBC	6
1.6	Hypothesis	7
	Bibliography	8-9
Chapter II: Review of Literature		
2.1	Epidemiology of GBC	10-13
2.1.1	Global trends of GBC incidence	10-11
2.1.2	GBC incidence in India	11-12
2.1.3	Incidence of GBC in Assam	13

2.2	Gallstone disease and its role in GBC pathogenesis	13-14
2.3	Molecular pathogenesis of GBC	15-17
2.3.1	Genetic mutations	15
2.3.2	Epigenetic modifications in GBC	16
2.3.3	Microsatellite instability and Loss of Heterozygosity in GBC	16-17
2.4	Regulation of noncoding RNAs in GBC pathogenesis	18-24
2.4.1	Role of miRNA regulation in GBC	18-20
2.4.2	Mechanism of lncRNA regulation	20-21
2.4.3	Potential lncRNA biomarkers identified in GBC	21-23
2.4.3.1	Oncogenic lncRNAs in GBC	22-23
2.4.3.2	Tumor suppressive lncRNAs in GBC	23
2.4.4	ceRNA regulation in GBC	23-24
2.5	Transcriptional regulation: Transcription factors as key regulators in GBC	25-26
2.6	Role of systems biology in identifying pathological signatures in cancer	26-27
2.7	NGS-based transcriptomic studies for precision oncology	28-29
2.8	Transcriptomic-based studies in GBC	29-31
2.9	Gaps in research	31-32
	Bibliography	32-47
Chapter III: Materials and Methodology		
3.1	Materials	48-49
3.1.1	Reagents and kits	48
3.1.2	Computational tools and databases	48-49
3.1.3	In-house computational pipelines	49
3.2	Methodology	50-63
3.2.1	Study design and participants	51
3.2.2	Transcriptome sequencing	52
3.2.3	Validation through quantitative Real-Time (qRT)-PCR	52-53
3.2.4	Retrieval of publicly available transcriptomic datasets	54
3.2.5	Transcriptome data analysis and generation of raw expression counts of transcribed genes	54-55

3.2.5.1	Identification of annotated and novel lncRNAs using in-house developed end-to-end novel lncRNA (ETENLNC) identification pipeline.	55-57
3.2.5.2	Coding Potential Analysis (CPA)	56
3.2.5.3	BLASTn analysis	56
3.2.6	Identification of significant differentially expressed genes and lncRNA from transcriptomic datasets.	57-58
3.2.7	3.2.7 Functional enrichment and pathway analysis	58
3.2.8	3.2.8 Gene coexpression network analysis	58-60
3.2.8.1	3.2.8.1 Pearson's correlations analysis for identification of co-expressed gene modules	59
3.2.8.2	3.2.8.2 Module Preservation Analysis	59
3.2.8.3	3.2.8.3 Intramodular Connectivity Analysis	59-60
3.2.9	3.2.9 PPI network analysis	60-61
3.2.10	ceRNA regulatory network analysis	61
3.2.11	Prediction of RNA secondary structures	61
3.2.12	Transcriptional Regulatory Network (TRN) analysis for identification of potential regulatory TFs in GBC.	61-63
3.2.13	Calculation of EMT scores	63
3.2.14	Cross-validation of the expression of DEGs and lncRNA identified through transcriptomic data analysis.	63
	Bibliography	64-66
CHAPTER IV: Understanding the gallbladder cancer transcriptome and identifying potential genes, processes, and pathways associated with GBC pathogenesis.		
4.1	Introduction	67-68
4.2	Results	68-114
4.2.1	Case study 1: Comparative transcriptomic data analysis of hepatobiliary cancers (HBCs) - GBC, hepatocellular carcinoma (HCC), and intrahepatic cholangiocarcinoma (ICC)	69-70
4.2.1.1	Dataset information	71
4.2.1.2	Identification of unique and shared significant DEGs in GBC, HCC, and ICC	71-73
4.2.1.3	Identification of module eigengene and detection of non-preserved modules	74-76

4.2.1.4	Pathway enrichment analysis of DEGs identified from significant nonpreserved modules	76-77
4.2.1.5	Identification of hub genes from the nonpreserved module clusters	77-79
4.2.2	Case study 2: Transcriptome data analysis and identification of differential gene expression signatures in GBC samples compared to adjacent normal samples	80-81
4.2.2.1	Dataset Information	81
4.2.2.2	Identification of differential gene expression profile in GBC compared to normal samples.	82
4.2.2.3	Differential gene coexpression network analysis and identification of nonpreserved modules from GBC and normal coexpression network	83-85
4.2.2.4	Functional annotation and pathway associated with genes of the non-preserved modules	85-87
4.2.2.5	Hub gene identification from non-preserved modules.	87-90
4.2.3	Case study 3: Transcriptomic data analysis on GBC compared to GSD groups with different follow-up periods	91-92
4.2.3.1	Dataset information and retrieval	92
4.2.3.2	Identification of differential molecular signatures in GBC compared to GSD with three different follow-up periods	93-94
4.2.3.3	Overlapping and unique DEGs in GBC compared to GSD with different follow-up periods.	94-95
4.2.3.4	Protein-protein interaction analysis and hub gene identification	96-97
4.2.3.5	In-silico validation of selected hub genes for each case study from TCGA datasets	98-99
4.2.4	Transcriptome sequencing and data analysis of GBC and GSD patients from Assam, North-East India.	100-101
4.2.4.1	Collection of GBC and GSD tissue samples	102
4.2.4.2	Transcriptome sequencing and generation of the GBC transcriptome dataset	103-104
4.2.4.3	Identification of differential gene expression profile of GBC+GS and GBC patients from Assam	104-105

4.2.4.4	Identification of differentially expressed shared DEGs between GBC and GBC+GS groups.	105-106
4.2.4.5	Functional annotation and pathway enrichment analysis of the DEGs identified in GBC+GS and GBC groups.	107-109
4.2.4.6	Construction of PPI network modules and identification of hub genes in GBC+GS and GBC group	110-111
4.2.4.7	Identification of the crucial pathological pathways associated with gene sets in PPI network modules	112-113
4.2.4.8	Validation of the selected hub DEGS through qRT-PCR.	113-114
4.3	Discussions	115-118
4.4	Summary	118-119
	Bibliography	120-123
Chapter V: Analysing noncoding RNA linked to GBC pathogenesis and constructing a ceRNA regulatory network to identify potential noncoding regulatory signatures in GBC.		
5.1	Introduction	124-125
5.2	Results	126-142
5.2.1	Identification of differential lncRNA profile in GBC and GBC+GS groups	126-127
5.2.2	Identification of shared and specific lncRNA signatures in GBC and GBC+GS groups.	127-128
5.2.3	Screening of hub DE-lncRNAs and DE-nlncRNAs through the construction of mRNA-lncRNA and mRNA-nlncRNA correlation networks	129-130
5.2.4	Pathway enrichment analysis of the DE-mRNAs coexpressed with hub DE-lncRNA and DE-nlncRNA identified in GBC and GBC+GS cases.	131-133
5.2.5	Analysis of lncRNA-mediated interaction and construction of ceRNA regulatory networks.	133-138
5.2.6	Validation of the identified hub lncRNAs in independent datasets and through qRT-PCR	138-142
5.3	Discussions	143-147
5.4	Summary	147

	Bibliography	148-152
Chapter VI: Understanding the rewiring patterns of transcriptional regulation in GBC pathogenesis and identifying the crucial transcription factors involved in GBC development.		
6.1	Introduction	153-154
6.2	Results	155-167
6.2.1	Identification of differentially expressed TFs (DETFs)	155
6.2.2	Functional pathway enrichment of the DETFs identified in GBC and GBC+GS groups	156-157
6.2.3	Construction of TF-TG and TF-lncRNA regulatory networks and identification of potential hub TFs in GBC and GBC+GS groups.	158-161
6.2.4	Pathway enrichment analysis of the DEGs and DE-lncRNAs targeted by the hub TFs identified from the GBC and GBC+GS networks.	161-163
6.2.5	Cross-validation of the identified potential DETFs in independent publicly available GBC transcriptomic dataset and TCGA datasets of gastrointestinal cancers.	163-166
6.2.6	Validation of the selected hub TFs through qRT-PCR.	166-167
6.3	Discussions	168-170
6.4	Summary	170
	Bibliography	171-173
Chapter VII: Conclusion and Future perspectives		
7.1	Conclusion	174-176
7.2	Future perspective	177

	List of publications	178
	Appendix-I: Approval of Ethical Committee	xxx
	Appendix-II: Reprints of publications	xxxi