

Genetic Predisposition and Molecular Mechanistic Studies of Matrix Metalloproteinases in Gallbladder Carcinogenesis

By

Vinay J

Enrolment No. LIFE11201604010

National Institute of Science Education and Research (NISER)

Bhubaneswar

*A thesis submitted to the Board
of Studies in Life Sciences*

*In partial fulfillment of requirements for
the Degree of*

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE



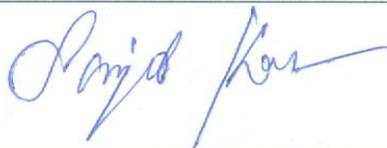
May, 2023

Homi Bhaba National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by **Mr. Vinay J** entitled "**Genetic Predisposition and Molecular Mechanistic Studies of Matrix Metalloproteinases in Gallbladder Carcinogenesis**" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman - Dr. Sanjib Kar



Date: 03.08.2023

Guide / Convener - Dr. Manjusha Dixit



Date: 03.08.2023

Examiner - Dr. Anshika Srivastava



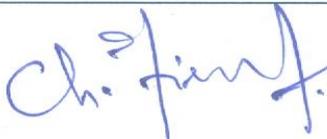
Date: 03.08.2023

Member 1 - Dr. Debasmita P Alone



Date: 03.08.2023

Member 2 - Dr. Tirumala K Chowdary



Date: 03.08.2023

Member 3 - Dr. P V Ramchander



Date: 3.8.2023

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.
I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 03/08/2023

Place: NISER, Jatani

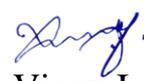


Dr. Manjusha Dixit
(Thesis Supervisor)

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfilment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.


Vinay J

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree/diploma at this or any other Institution/University.



Vinay J

CERTIFICATE

This is to certify that the thesis entitled "Genetic Predisposition and Molecular Mechanistic Studies of Matrix Metalloproteinases in Gallbladder Carcinogenesis", which is being submitted by Mr. Vinay J in partial fulfilment of the degree of Doctor of Philosophy (PhD) in Life Sciences of Homi Bhabha National Institute is a record of his own research work carried by him. He has carried out his research for the last seven years on the subject matter of the thesis under my supervision at National Institute of Science Education and Research (NISER), Bhubaneswar.

To the best of our knowledge, the matter embodied in this thesis has not been submitted for the award of any other degree.



Signature of the Candidate

Date: 03/08/2023

Vinay J
National Institute of Science
Education and Research (NISER)
Bhubaneswar



Signature of the Supervisor

Date: 03/08/2023

Dr. Manjusha Dixit
Associate Professor
National Institute of Science
Education and Research (NISER)
Bhubaneswar

Publications pertaining to the thesis:

1. **Vinay J**, Mishra D, Meher D, Dash S, Besra K, Pattnaik N, Singh SP, Dixit M. Genetic association of MMP14 promoter variants and their functional significance in gallbladder cancer pathogenesis. *J. Hum. Genet.* 66, 947–956 (2021). <https://doi.org/10.1038/s10038-021-00917-x>
2. **Vinay J**, Ananya Palo, Kusumbati Besra, Manjusha Dixit. Gallbladder cancer-associated genetic variants rs1003349 and rs1004030 regulate MMP14 expression by altering SOX10 and MYB binding sites. *Human Molecular Genetics*, 2023;, ddad077, <https://doi.org/10.1093/hmg/ddad077>
3. **Vinay J**, Shivaram Prasad Singh, Manjusha Dixit. Matrix metalloproteinase-7 (MMP7) promoter variants association and functional significance in gallbladder cancer. (Under Review)
4. **Vinay J**, Shivaram Prasad Singh, Manjusha Dixit. Matrix metalloproteinase-2 (MMP2) promoter variants association and functional significance in gallbladder cancer. (Under preparation)

Invited Book chapter:

1. Manjusha Dixit, **Vinay J**, Soham Choudhury. Targeted Therapies in Gallbladder Cancer: Current Status and Future Perspectives. In: Kumar Shukla, V., Pandey, M., Dixit, R. (eds) *Gallbladder Cancer*, (2023). Springer, Singapore. https://doi.org/10.1007/978-981-19-6442-8_16

Other publications:

1. Mahananda B, **Vinay J**, Palo A, Singh A, Sahu SK, Singh SP, Dixit M. SERPINB5 Genetic Variants rs2289519 and rs2289521 are Significantly Associated with Gallbladder Cancer Risk. *DNA Cell Biol.* 2021 May;40(5):706-712.
2. Sinha KK^{*}, **Vinay J**^{*}, Parida S, Singh SP, Dixit M. Association and functional significance of genetic variants present in regulatory elements of SERPINB5 gene in gallbladder cancer. *Gene.* 2022 Jan 15;808:145989.

* Joint first author.

Conference proceedings:

1. **Vinay J**, Kusumbati Besra, Niharika Pattnaik, Shivaram Prasad Singh, Manjusha Dixit. Genetic association and functional significance of matrix metalloproteinase-14 promoter variants rs1004030 and rs1003349 in gallbladder cancer pathogenesis. 2022 March 22, Journal of Gastroenterology and Hepatology, Volume37 Issue4 774-744, PP-0123.
2. **Vinay J**, Nachiketa Mohapatra, Shivaram Prasad Singh, Manjusha Dixit. Genetic association and functional significance of MMP7 promoter variants in gallbladder cancer pathogenesis. 2022, IACR, ISBN: 978-81-954625-2-0, 234-235, AP-A33.



Vinay J

**Dedicated to
My Beloved Family and
Teachers**

Acknowledgements

I owe my deepest gratitude to my supervisor Dr. Manjusha Dixit, who's invaluable support and contribution helped me in completing my doctoral work. Her optimism concerning my work, enthusiasm, friendly behaviour, motivation and freedom of conscience made me what I am today. I am thankful to the National Institute of Science Education and Research (NISER), Department of Atomic Energy (DAE) for granting the fellowship and all facilities during the PhD tenure. I would also like to express my appreciation to Prof. S. Panda, Director of NISER, and Prof. T. K. Chandrashekar, Founder Director of NISER, for their generous provision of laboratory facilities and instrumentation required for carrying out this work.

My sincere gratitude towards my Thesis Monitoring Committee members, Dr. Sanjib Kar, Dr. Tirumala Kumar Chowdary, Dr. Debasmita P Alone, and Dr. P V Ramchander for their critical and useful suggestions that immensely improved my thesis. I also owe my deepest gratitude to my collaborator Dr. Shivaram Prasad Singh, Dr. Kusumbati Besra and the study participants for providing the necessary samples. I would like to acknowledge Dr. Debakanta Mishra, Dr. Dinesh Meher, Dr. Suryakanta Parida, Dr. Ayashkanta Singh, Dr. Saroj Kanta Sahu, Sashibhusan Dash, and other staff members of collaborating hospitals for their assistance in coordinating the collection of patient samples. Special thanks to Dr. Niharika Pattanaik, Senior pathologist, AMRI Hospitals, and Dr. Nachiketa Mohapatra, Senior pathologist Apollo Hospital Pvt Ltd. for their help with clinical studies related to this project.

I am also grateful for the support provided by the Chairperson of SBS NISER and all other faculty members of SBS, NISER, throughout my project. I extend my thanks to my seniors, particularly Gargi di, for her guidance during the initial stages of my Ph.D. My special acknowledgement to my senior lab members Dr. Dinesh kumar and Dr. Khurshidul Hassan for their guidance and discussions related to this work.

I am thankful to my lab members Dr. Ankit Tiwari, Bratati Mukherjee, Saket A Patel, Ananya Palo, Neha Ramani, Kirti Sinha, Bishwahiree Mahananda, Pratush, Monali, Rehan, Abhishek, Prasannajit, and other younger member of the lab for their interactive discussion and maintain a healthy research environment.

I would like to extend a heartfelt thank you to my friends, Somlata Khamaru and Naresh Kumar G, who have helped me immensely, and constantly motivated me challenging times.

It is my great pleasure to thank my batchmates Ram Prasad Sahoo, Ramani Shyam, Aranyadip, Saptarsi Mitra, Ajay, Victor, Ankita and friends Prakash, Shiva, Vamsi, Suwendu, Srini for their immense love and support.

I am grateful to the academic and non-academic staff of NISER for their blessings, encouragement, and assistance. I would like to thank all my teachers and friends who have consistently supported and accepted me despite my flaws.

It is my greatest honour to have such motivating Mother and Father who never ceased to believe in me. Their unwavering support, unconditional love, and unwavering belief in my abilities have been the driving force behind my accomplishments. I am forever grateful for their sacrifices, guidance, and encouragement throughout my academic journey.

Lastly, I would like to express my heartfelt appreciation to all my cousins and close family members for their continuous support, love, and care. Their presence and unwavering belief in me have been a constant source of strength and motivation.

I would also like to take this opportunity to apologize if I have inadvertently omitted anyone in this brief acknowledgment. It in no way diminishes my gratitude towards them. I am truly thankful for the support and contributions of all those who have played a part in shaping my academic and personal growth.

Once again, my sincere thanks to everyone who has been a part of my journey and has contributed to my success.



Vinay J

CONTENTS

Introduction	0
1.1 Introduction.....	1
2 Review of Literature.....	7
2.1 Gallbladder cancer	8
2.2 Pathology of gallbladder cancer	10
2.2.1 Chronic inflammation-mediated pathway.....	11
2.2.2 Congenital abnormality in bile duct development model	13
2.2.3 Preneoplastic lesions of Gallbladder Carcinoma	14
2.3 Epidemiology of gallbladder cancer	15
2.3.1 Gallbladder cancer demographics	15
2.3.2 GBC Related Mortality	18
2.4 Etiology and risk factors	19
2.4.1 Gender and age.....	20
2.4.2 Gallstones.....	21
2.4.3 Obesity	22
2.4.4 Chronic inflammation	23
2.4.5 Primary sclerosing cholangitis	24
2.4.6 Genetic factors.....	24
2.5 Major pathways associated with gallbladder cancer.....	30
2.5.1 PI3K/AKT/mTOR Signalling Pathway.....	30
2.5.2 MAPK/ERK Signalling Pathway	31
2.5.3 EGFR pathway	32
2.6 Available and potential therapies for gallbladder cancer.....	33
2.7 Matrix metalloproteinases.....	36
2.7.1 Matrix metalloproteinase's role in tumorigenesis.....	37
2.7.2 Matrix metalloproteinase mode of action.....	39
2.7.3 Genetic association studies of MMP2, 7 and 14 in cancers	42

3	Materials and Methods	45
3.1	Materials.....	46
3.1.1	Study participants	46
3.1.2	Cell lines	48
3.1.3	Vector constructs, shRNAs, antibodies and primer sequences.....	48
3.2	Chemicals and reagents	48
3.3	Methodology	52
3.3.1	DNA extraction.....	52
4.1.1	Polymerase chain reaction.....	53
3.3.2	PCR gel purification	54
3.3.3	Sanger’s sequencing	55
3.3.4	Sequencing reaction cleanup	55
3.4	Cloning.....	56
3.4.1	DH5 α Ultra-competent cell preparation	56
3.4.2	Cloning of MYB and SOX10 shRNA	57
3.4.3	Cloning of Luciferase vectors.....	63
3.4.4	sgRNA design and CRISPR-Cas9 cloning.....	67
3.5	CRISPR Cas9 mediated genome editing.....	69
3.6	Cell culture and cell-based assays in current study.....	70
3.6.1	Culturing of human gallbladder cancer and HEK293T cell lines.....	70
3.6.2	Subculture, trypsinization and passage of cell lines	70
3.6.3	Storage and revival of mammalian cell lines.....	71
3.6.4	Stable and transient transfection of cell lines	71
3.6.5	Cell proliferation assay	72
3.6.6	Colony formation assay	72
3.6.7	Scratch wound healing assay	73
3.6.8	Transwell migration assay	73
3.6.9	Transwell Matrigel invasion assay	74

3.7	Luciferase assay	74
3.8	Electrophoretic mobility shift assay (EMSA).....	74
3.9	Chromatin immunoprecipitation assay (ChIP)	75
3.10	Immunohistochemistry	76
3.10.1	Poly-L-Lysine coating of slides	76
3.10.2	Immunohistochemistry staining	76
3.10.3	Immunohistochemistry scoring	76
3.11	Western blotting.....	77
3.11.1	Preparation of mammalian cell lysate	77
3.11.2	Protein estimation.....	77
3.11.3	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)..	77
3.11.4	Immunoblotting	78
3.11.5	Stripping and reprobing of blots.....	79
3.12	<i>In-silico</i> analysis	79
3.12.1	Differential Gene Expression Analysis of RNA-Seq data from GEO database 79	
3.12.2	Gene Set Enrichment Analysis (GSEA) for differentially expression genes in two gallbladder cancer expression data set from GEO	80
3.12.3	Visualization and interpretation of pathway enrichment analysis using Cytoscape EnrichmentMap tool	81
3.12.4	Somatic mutation and Copy number alteration analysis using cBioportal ...	81
3.12.5	Study of methylation status in TCGA data	82
3.12.6	Analysis of TCGA data using UCSC Xena browser.....	82
4	To find out the association of MMP14 promoter variants with GBC and its molecular mechanism.....	83
4.1	Introduction.....	85
4.2	Screening of MMP14 promoter SNPs in GBC.....	87
4.3	Discovery set allele and genotype frequency analysis.....	87
4.4	Analysis of Hardy-Weinberg equilibrium (HWE) and study power	88

4.5	Demographic profile of study subjects.....	89
4.6	The rs1004030 allele C is associated with an increased risk of gallbladder cancer	90
4.7	The variant rs1003349 is associated with gallbladder cancer	91
4.8	Linkage disequilibrium and haplotype analysis of MMP14 promoter variants ...	92
4.9	MMP14 expression in gallbladder cancer tissue is between the genotypes of variant rs1004030 and rs1003349.....	92
4.10	Risk variants rs1004030 and rs1003349 may elevate MMP14 expression.....	94
4.11	The risk alleles ‘G’ (rs1004030) and ‘C’ (rs1003349) provide binding sites for transcription factors SOX10 and MYB	97
4.12	Perturbation of SOX10 and MYB expression alters MMP14 expression.....	99
4.13	Knockout of rs1004030 and rs1003349 containing promoter region reduces MMP14 expression and abolishes the effect of SOX10 and MYB on MMP14 expression	102
4.14	SOX10 and MYB expression levels have an allele-specific effect on reporter gene expression.....	103
4.15	MMP14 increases tumorigenic properties of GBC cell lines.....	105
4.16	MMP14 expression in gallbladder cancer tissues samples	107
4.17	MYB inhibitor treatment reduces the tumorigenic properties in GBC cells.....	109
4.18	Deletion of the transcription factor binding site reduces the tumorigenic properties in GBC cells.....	111
4.19	Discussion	112
5	To determine the association of MMP7 promoter variants with GBC and its effect on gene expression.....	116
5.1	Introduction	118
5.2	Screening of MMP7 promoter SNPs.....	119
5.3	Discovery set allele and genotype frequency analysis	119
5.4	Analysis of Hardy-Weinberg equilibrium and study power	121
5.5	Demographic profile of study subjects.....	122

5.6	rs113823671 allele C is associated with an increased risk for gallbladder cancer	123
5.7	Allele A of variant rs17098318 is associated with Gallbladder cancer	124
5.8	Haplotype analysis	126
5.9	Promoter variant rs113823671 risk allele C shows increased luciferase activity	126
5.10	Gallbladder cancer tissue expression of MMP7 and its association with rs113823671 and rs17098318 genotypes	127
5.11	Discussion	130
6	To discover the genetic association of MMP2 promoter SNPs with GBC and genotype-phenotype correlation	133
6.1	Introduction	135
6.2	Screening of MMP2 promoter variants in GBC	136
6.3	Alleles and genotypes frequency in the discovery set	136
6.4	Analysis of Hardy-Weinberg equilibrium and study power	141
6.5	Demographic profile of study subjects	141
6.6	Association of the rs1961998763 variant with increased risk of gallbladder cancer	142
6.7	Allele 'T' of variant rs243865 is associated with increased risk of gallbladder cancer	143
6.8	Association of rs1961996235 variant with increased risk of gallbladder cancer	145
6.9	rs1391392808 allele 'T' is associated with an increased risk for gallbladder cancer	146
6.10	The variant v.55477735G>A is associated with an increased risk for gallbladder cancer	147
6.11	Association of rs1488656253 variant with increased risk of gallbladder cancer	148
6.12	Allele 'T' of variant rs17859816 is associated with Gallbladder cancer	148
6.13	Haplotype analysis	149
6.14	MMP2 expression in gallbladder cancer tissue and its association with promoter variants	151
6.15	Promoter variants rs243865 risk allele 'T' shows increased luciferase activity .	154

6.16	Discussion	156
7	To find out MMP associated pathways altered in GBC using publicly available databases	159
7.1	Introduction	161
7.2	Differential Gene Expression Analysis of RNA-Seq data from GEO database.	162
7.3	Enrichment analysis of DEGs in primary GBC patients	162
7.4	Enrichment analysis of DEGs in metastatic GBC patients	166
7.5	GSEA.....	170
7.6	The role of matrix metalloproteinases (MMP-2,-7 and -14) in the pathogenesis of GBC.....	176
7.7	Somatic mutations, copy number variation and methylation profile of MMP-2, -7 and -14, in GBC patients	177
7.8	Discussion	181
8	Summary and conclusion.....	185
9	BIBLIOGRAPHY.....	191
8.1	Bibliography.....	192
9	APPENDIX.....	217
8.2	APPENDIX I: VECTORS	218
8.3	APPENDIX II: shRNAs and sgRNAs.....	218
8.4	APPENDIX III: ANTIBODIES	219
8.5	APPENDIX IV: PRIMER SEQUENCES	219
8.6	APPENDIX V: KIT PROTOCOLS/ ASSAY PROCEDURES	221
8.7	APPENDIX VI: BUFFERS.....	226

SUMMARY

Gallbladder cancer (GBC) is the most common type of biliary tract cancer. According to the Globocan 2020 report, GBC is responsible for 10% of the global GBC burden and is the 16th leading cause of cancer-related fatalities in India. Studies indicate a concentration of cases in particular geographic and racial groups. In India, regions with high incidence rates include the north, central, east, and north-east, while states with low incidence rates include the west and south, showing a population-specific genetic susceptibility. The ECM remodelling proteins known as matrix metalloproteinases (MMPs) have important roles in tissue morphogenesis, angiogenesis, development, and wound healing. Its deregulation in cancer is frequently linked to the breakdown of basement membrane and cell-matrix adhesion molecules, which promotes tumour invasion and metastasis. The genetic predisposition and increased expression of MMPs have been reported in several malignancies, including gastric, hepatocellular, colorectal, ovarian, and oral cancer. We learned from the literature that MMP-1, -2, -3, -7, and -14 play a part in tumours initiation and progression. The matrix metalloproteinase family member-MMP14 is the most commonly studied member and is involved in a variety of biological processes, including as angiogenesis, proliferation, invasion, and basement membrane. By cleaving the LTBP-1 protein, MMP14 and MMP2 activate TGF- β in a CD44-dependent way, promoting tumorigenesis. The MMP14 plays a part in tumour initiation by activating latent TGF β 1 and RANLK, which is frequently elevated in cancers. According to reports, the functional promoter variants rs1003349 and rs1004030 facilitate the binding of Sp1 and RR1 and modulate the expression of MMP14. MMP7 is the smallest member of MMPs due to its lack of the C-terminal hemopexin domain, which is seen in other MMPs. MMP7 is essential for the shedding and bioavailability of several proteins, including insulin-like growth factor-binding proteins, tumour necrosis factor- α , Fas ligand, heparin-binding EGF-like growth factor, β 4-integrin, and E-cadherin. The MMP7 functional variations rs11568818 and rs11568819

have been linked to an increased risk of breast, gastric, and lung cancer. Promoter variant rs11568818 has been reported to bind CREB protein in an allele-specific manner. MMP2 improves IGF bioavailability by degrading IGFBP-3 and promotes cellular growth and proliferation. The MMP2 regulatory promoter variations rs243865 and rs2285053 have been linked to an elevated risk for lung, nasopharyngeal, and esophageal carcinoma. MMP2 promoter variations rs243865 and rs2285053 influence Sp1 binding in an allele-specific way, which has been verified in murine macrophage and aortic smooth muscle cell lines. Our work is focused on MMP14, MMP2, and MMP7 due to their roles in tumour initiation. We hypothesised changes in expression levels of these MMPs due to genetic predisposition may contribute to cancer initiation. Prior studies support the allele-specific effect on expression levels in other malignancies; however, the relevance and underlying mechanism of functional variations in GBC remain unknown.

Based on the above literature, we carried out a case-control study to understand the role of MMP-2, -7, and -14 promoter variants in the pathogenesis of GBC in the Eastern Indian population (Odisha). Along with the genetic association study of these MMPs, detailed functional studies were carried out by using various genetic, molecular, and biochemical assays to understand their role in GBC pathogenesis.

ABBREVIATIONS

GBC	Gallbladder cancer
SNPs	Single Nucleotide Polymorphisms
GWAS	Genome-Wide Association Studies
DNA	Deoxyribonucleic acid
OR	Odds Ratio
CI	Confidance Interval
LD	Linkage Disequilibrium
HWE	Hardy-Weinberg Equilibrium
PCR	Polymerase chain reaction
EMT	Epithelial to mesenchymal transition
ECM	Extracellular matrix
MMPs	Matrix metalloproteinases
EGF	Epidermal growth factor
FGF	Fibroblast growth factors
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
EGFR	Epidermal growth factor receptor
CHOL	Cholagiocarcinoma
HCC	Hepatocellular carcinoma
siRNA	Small interfering RNA
sgRNA	Small guided RNA
FFPE	Formalin-fixed paraffin-embedded
KD	Knock-down
SC	Scrambled
EV	Empty vector
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
qRT-PCR	Quantitative real-time polymerase chain reaction
TCGA	The Cancer Genome Atlas
MAPK	Mitogen-activated protein kinase
DMEM	Dulbecco's modified eagle medium
RPMI	Roswell park memorial institute
DPBS	Dulbecco's phosphate-buffered saline
DMSO	Dimethyl sulfoxide

FBS	Fetal bovine serum
PBS	Phosphate-buffered saline
ERK	Extracellular signal-regulated kinase 1/2
AKT	Protein kinase B
CREB	cAMP-response element binding protein
IGFBP-3	Insulin-like Growth Factor Binding Protein 3
IGF	Insulin-like Growth Factor
RANLK	Receptor activator of nuclear factor κ B ligand
TGF- β	Transforming Growth Factor beta
CD44	Cluster of differentiation 44
Sp1	Specificity protein 1
SOX10	SRY-Box Transcription Factor 10
MYB	Myeloblastosis viral oncogene homolog
shRNA	Short hairpin RNA
SREBP-1c	Sterol-regulatory element binding protein-1c
USF-1	Upstream Transcription Factor 1
RR1	Ribonucleotide Reductase 1

LIST OF FIGURES

Figure 2.1.1 Anatomy and clinical features of the gallbladder and invasive gallbladder adenocarcinoma.	9
Figure 2.2.1 Histologic types of invasive gallbladder carcinoma..	10
Figure 2.2.2 Sequential histological and molecular alterations in the pathology of gallbladder cancer linked to inflammation.	11
Figure 2.2.3 High-grade dysplasia/carcinoma in situ in the gallbladder types.	12
Figure 2.2.4 Types of gallbladder metaplasia.	12
Figure 2.3.1 The global gallbladder cancer incidence.	16
Figure 2.3.2 Estimated number of new gallbladder cancer cases in worldwide and Asia from 2020-2040 (both gender, age (0-85+)).	16
Figure 2.3.3 Age-adjusted incidence rates (AAR) of gallbladder cancer between Northern and Southern cities of India for both genders.	17
Figure 2.3.4 Age-adjusted incidence rates (AAR) of gallbladder cancer across all population-based cancer registries in India for both genders.	18
Figure 2.3.5 Estimated age-standardized mortality rates of GBC worldwide in 2020.19	
Figure 2.4.1 The gallbladder cancer risk increases after the age of 30.	21
Figure 2.4.2 The Gallstone prevalence and the occurrence of gallbladder cancer in high-risk populations.	22
Figure 2.6.1 The Major pathways associated with gallbladder cancer.	35
Figure 2.7.1 The Matrix metalloproteinases domain structure.	37
Figure 2.7.2 Matrix metalloproteinase modes of action.	40
Figure 3.4.1 Designing of shRNA for pLKO.1 TRC Cloning Vector.	58
Figure 3.4.2 pLKO.1 vector map.	60
Figure 3.4.3 Designing of luciferase assay inserts.	64
Figure 3.4.4 CRISPR-Cas9 target selection, sgRNA design and pSpCas9(BB) plasmid construction.	68
Figure 4.5.1 Gallbladder cancer genetic risk forest plot and linkage disequilibrium plot for MMP14 promoter variants rs1004030 and rs1003349.	90
Figure 4.9.1 The MMP14 protein expression in gallbladder cancer and surrounding normal tissue.	93
Figure 4.9.2 MMP14 protein expression levels in gallbladder cancer patients with different genotypes of variants rs1004030 and rs1003349.	94

Figure 4.10.1 Evaluation of the functional potential of promoter variants rs1004030 and rs1003349.	96
Figure 4.11.1 The risk alleles 'C' (rs1004030) and 'G' (rs1003349) provide binding sites for transcription factors SOX10 and MYB.	98
Figure 4.12.1 MYB and SOX10 expression levels regulate MMP14 expression.. .	100
Figure 4.12.2 MYB and SOX10 expression perturbation in the HEK293T and G415 cell lines.	101
Figure 4.13.1 The effect of SOX10 and MYB on MMP14 expression in promoter knockout cells.....	103
Figure 4.14.1 SOX10 and MYB expression levels have an allele-specific effect on reporter gene expression.....	104
Figure 4.15.1 MMP14 expression affects the tumorigenic characteristics of GBC cells.. ..	106
Figure 4.15.2 Stable expression of MMP14 in TGBC1TKB and G415 cell lines.. .	107
Figure 4.16.1 The expression and distribution of MMP14 protein in gallbladder tumour tissue and surrounding normal tissue.. ..	108
Figure 4.17.1 Effect of MYB inhibition and CRISPR-Cas9-mediated transcription factor binding site deletion on the tumorigenic properties of G415 cells.	111
Figure 4.17.2 The MMP14 expression level following MYB inhibition treatment in G415 cells.....	111
Figure 5.5.1 Gallbladder cancer genetic risk forest plot and linkage disequilibrium plot for <i>MMP7</i> promoter variants rs113823671, and rs17098318.	123
Figure 5.9.1 Allele-specific luciferase activity of variants rs113823671 and rs17098318.	127
Figure 5.10.1 The MMP7 protein expression in gallbladder tumour tissue and adjacent normal tissue.	129
Figure 5.10.2 MMP7 protein expression in gallbladder cancer patients with different genotypes of variants rs113823671 and rs17098318.	129
Figure 6.3.1 Linkage disequilibrium plot for <i>MMP2</i> promoter variants.	151
Figure 6.3.2 The MMP2 protein expression in gallbladder tumour tissue and adjacent normal tissue.. ..	152
Figure 6.3.3 MMP2 protein expression in gallbladder cancer patients with genotypes of its promoter variant rs17859816.	154
Figure 6.3.4 Allele-specific luciferase activity of MMP2 promoter variant rs243865.	155

Figure 7.3.1 The Bubble plot data of functional enrichment analysis of primary GBC patients from DAVID web database.....	164
Figure 7.4.1 Bubble plot analysis data from DAVID web database showing functional enrichment analysis of metastatic GBC patients.	168
Figure 7.5.1 Visualisations of gene set enrichment analysis (GSEA) using Cytoscape Enrichment Map.	174
Figure 7.5.2 Gene set enrichment analysis (GSEA). <i>Results of primary and metastatic GBC samples gene set enrichment analysis (GSEA).</i>	176
Figure 7.7.1 The OncoPrint plots showing frequency of Genetic alterations in various MMPs in gallbladder cancer.	179
Figure 7.7.2 Methylation status of MMP-2, -7 and -14 promoter region and its correlation with its mRNA expression..	181

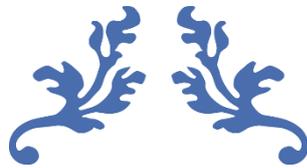
LIST OF TABLES

Table 2.4.1 Genetic predisposition studies of candidate genes in gallbladder cancer*	27
Table 2.7.1 MMPs induce carcinogenesis in mice with genetic modifications	38
Table 2.7.2 Genetic association studies of MMPs in various cancer	44
Table 3.1.1 Demographic profile of the gallbladder cancer patients and controls recruited in the study	47
Table 3.1.2 Gallbladder cancer patients clinical and histopathological characteristics	47
Table 3.1.3 List of cell lines used in the study and their cellular origin	48
Table 3.3.1 List of primers used in the study	53
Table 3.3.2 PCR reaction mix	54
Table 3.3.3 Thermal cycling condition for PCR reactions	54
Table 3.3.4 Sequencing PCR reaction mix	55
Table 3.3.5 Thermal cycling condition for sequencing PCR reactions	55
Table 3.4.1 Annealing of Oligos	58
Table 3.4.2 Digesting pLKO.1 TRC Cloning Vector	59
Table 3.4.3 T4 DNA ligase reaction mix	60
Table 3.4.4 PCR reaction mix	61
Table 3.4.5 Thermal cycling condition for PCR reaction	61
Table 3.4.6 Screening for insert clones by restriction digestion	62
Table 3.4.7 Sequencing PCR reaction mix	62
Table 3.4.8 Thermal cycling condition for sequencing PCR reactions	62
Table 3.4.9 Annealing of Oligos	64
Table 3.4.10 Digestion of pGL4.23 luciferase vector	64
Table 3.4.11 T4 DNA ligase reaction mix	65
Table 3.4.12 Colony PCR reaction mix	66
Table 3.4.13 Thermal cycling condition for colony PCR reactions	66

Table 3.4.14 Sequencing PCR reaction mix.....	66
Table 3.4.15 Thermal cycling condition for sequencing PCR reactions	67
Table 3.4.16 List of oligonucleotides used to generate sgRNAs for the generation of CRISPR-Cas9 based MMP14 promotor deletion construct	67
Table 3.4.17 Digestion of pSpCas9(BB) vector	69
Table 3.6.1 Reaction setup for Lipofectamine 3000 mediated transfection#.....	72
Table 3.11.1 The volumes and concentrations of components to prepare SDS-PAGE gels*78	
Table 4.3.1 Allelic and genotypic frequency distribution of MMP14 promoter variants in GBC and control subjects (Discovery set).....	88
Table 4.4.1 Sample size and HWE calculations for MMP14 promoter SNPs	89
Table 4.5.1 Demographic profile of the gallbladder cancer patients and controls recruited in the study	89
Table 4.7.1 Allelic and genotypic frequency distribution of rs1004030 and rs1003349 in GBC and control subjects	91
Table 4.8.1 Haplotype association of the variants rs1004030 and rs1003349 with gallbladder cancer.....	92
Table 5.3.1 Allelic and genotypic frequency distribution of MMP7 promoter variants in GBC and control subjects (Discovery set).....	121
Table 5.4.1 Sample size and HWE calculations for MMP7 promoter SNPs.	122
Table 5.5.1 Demographic profile of the gallbladder cancer patients and controls recruited in the study	122
Table 5.6.1 Allelic and genotypic frequency distribution of rs113823671 in GBC and control subjects	124
Table 5.7.1 Allelic and genotypic frequency distribution of rs17098318 in GBC and control subjects	125

Table 5.8.1 Haplotype association of the variants rs113823671 and rs17098318 with gallbladder cancer.	126
Table 5.9.1 Putative transcription factor binding at variants rs113823671 and rs17098318.	127
Table 6.3.1 Allelic and genotypic frequency distribution of MMP2 promoter variants in primer set 1 (Discovery set).....	138
Table 6.3.2 Allelic and genotypic frequency distribution of MMP2 promoter variants in primer set 2 (Discovery set).....	139
Table 6.3.3 Sample size and HWE calculations for MMP2 promoter SNPs.....	141
Table 6.3.4 Demographic profile of the gallbladder cancer patients and controls recruited in the study.....	142
Table 6.3.5 Allelic and genotypic frequency distribution of rs1961998763 in GBC and control subjects.....	143
Table 6.3.6 Allelic and genotypic frequency distribution of rs243865 in GBC and control subjects.....	144
Table 6.3.7 Allelic and genotypic frequency distribution of rs1961996235 in GBC and control subjects.....	145
Table 6.3.8 Allelic and genotypic frequency distribution of rs1391392808 in GBC and control subjects.....	146
Table 6.3.9 Allelic and genotypic frequency distribution of a novel variant NC_000016.10:g.55477735G>A in GBC and control subjects.....	147
Table 6.3.10 Allelic and genotypic frequency distribution of variant rs1488656253 in GBC and control subjects	148
Table 6.3.11 Allelic and genotypic frequency distribution of variant rs17859816 in GBC and control subjects	149
Table 6.3.12 Haplotype association of the MMP2 promoter variants with gallbladder cancer	150

Table 7.3.1 Gene ontological classifications of differentially expressed genes in primary GBC patients.....	165
Table 7.4.1 Gene ontological classifications of differentially expressed genes in metastatic GBC patients using DAVID web database.....	168
Table 7.5.1 Gene set enrichment analysis showing top gene-sets in GBC with primary samples, ranked according to NES.....	171
Table 7.5.2 Gene set enrichment analysis showing top gene-sets in GBC with metastatic samples, ranked according to NES	174



CHAPTER 1

Introduction



Vinay J
NISER, Bhubaneswar

1.1 Introduction

Gallbladder cancer (GBC) is the most frequent biliary tract malignancy and the digestive tract's fifth most common malignant neoplasm. It is often diagnosed at a later stage due to its metastatic nature. GBC manifests a unique combination of various predisposing factors; these include ethnicity, genetic predisposition, and other risk factors such as female gender, gallstone, chronic inflammation, and congenital abnormalities. Its frequency of incidence varies in different geographic regions and ethnic groups (1). High-risk regions include Chile, India, Pakistan, China, and Japan (2). The incidence of gallbladder cancer among American Indians, Alaskan native people, Eastern European people, and North Indians is relatively higher than other races (3). The clustering of cases suggests possible genetic predisposition within the population. Hence, identifying population-specific genetic risk factors is an integral aspect of untangling the complex pathophysiology of GBC.

GBC is a highly lethal disease that poses a significant threat because of its poor prognosis and tendency to spread quickly to nearby lymph nodes and organs, such as the liver and bile ducts (4). The spread of GBC is responsible for most cancer-related deaths (5, 6). Studies have shown that around 50% of GBC patients have liver metastases and lymph node involvement (7, 8). During metastasis, tumour cells undergo epithelial-mesenchymal transition (EMT), where cells of epithelial origin acquire features crucial for invasion and metastasis, including enhanced motility, invasiveness, and the ability to degrade components of the extracellular matrix (ECM), thereby altering the tumour microenvironment (9). Where, the matrix metalloproteinases (MMPs) have been shown to play a significant role in altering the tumour microenvironment and promoting tumorigenesis (10). MMPs deregulation is often associated with the disintegration of basement membrane and cell-matrix adhesion molecules, which facilitates tumour invasion and metastasis (11).

In addition to EMT, MMPs affect various physiological functions like wound healing, growth, development, and immune responses (11, 12). Also, they have role in apoptosis, cell migration, and angiogenesis (13, 14). The MMP's functionality often complements classical tumour properties leading to invasion, immune system avoidance, and metastasis. Many reports emphasizes over expression of MMPs in various cancers such as, melanoma (15), gliomas (16), breast (17, 18), lung (19, 20), ovarian (21), colon (22), pancreatic (23), hepatocellular (24), and gastric carcinoma (25) and it is often correlated with tumour metastasis.

Interestingly, several studies have showed the role of MMPs in cancer initiation (26). MMP7 overexpression in transgenic knockout mice demonstrated its role in early-stage mammary tumorigenesis by promoting angiogenesis through the degradation of sFlt-1, a VEGF inhibitor (27, 28). Additionally, the expression of MMP7 by mutant APC tumoursuppressor protein has been found to induce early-stage mammary tumorigenesis by activating Wnt/ β -catenin signalling (29-31). Likewise, MMP2 promotes tumorigenesis by increasing IGF bioavailability through IGFBP-3 degradation, promoting cellular growth and proliferation (32). In pancreatic cancer, MMP2 and MMP9 contribute to the angiogenic switch, which is a crucial element in tumorigenesis (33). Taken together, these studies underscore the importance of MMPs in tumorigenesis. Except for one isolated study in 2009 (34), which prospected the differential expression of MMP2, MMP9 and MMP14 in GBC, till now no studies have been carried out to gain further insight into the possible role of MMPs in GBC initiation and metastasis. Investigating early tumorigenic factors, such as MMPs, in GBC is critical because it can help understand the population-specific genetic predisposition and establish the molecular mechanisms underlying tumour invasion and metastasis. Odisha, is a part of North and East Indian states which have high frequency of GBC; therefore, we focused on identifying the MMPs having role in tumour initiation. It is possible that genetic polymorphism in MMPs can affect their expression level and in turn alter the risk for GBC development. MMP14, MMP7

and MMP2 are widely studied to have tumorigenic roles in multiple cancers (26). Interestingly, they have role both at tumour initiation and progression (26, 29, 33), therefore we selected them to explore their role in GBC.

The role of MMP14 in tumour initiation and progression was first reported in breast cancer (35). Where, overexpression of MMP14 in a transgenic mice model promote tumorigenesis by remodeling of the ECM and inducing mammary gland adenocarcinoma formation. The MMP14 transgenic mice displayed abnormalities in 82% of female mammary glands. The author further verified the abnormalities as lymphocytic infiltration, fibrosis, hyperplasia, alveolar structure disruption, dysplasia, and adenocarcinoma (35). The MMP14 is often upregulated in cancers such as oral, breast (36), lung (37), liver (38), ovary (39), colon, bladder, and gastric cancers (40) and have role in the tumour initiation, invasion, and metastasis. Also, MMP14 has a significant role in tumorigenesis by activating latent TGF β 1 (10) and RANLK (41), thereby modulating tumor microenvironment by involving cancer-associated fibroblasts (CAFs) and tumor-associated macrophages to facilitate tumor progression in cancer (42).

MMP7 (matrilysin-1) is a small secretory proteolytic enzyme and known for its tumorigenic role in various cancers (43). The overexpression of MMP7 in the transgenic knockout mice demonstrates its role in early-stage mammary tumorigenesis (28). MMP7 degrades soluble VEGF receptor-1 (sVEGFR-1/sFlt-1) to promote angiogenesis in human umbilical vein endothelial cells (HUVECs), which is an endogenous VEGF inhibitor that sequesters VEGF and blocks its access to VEGF receptors. The degradation of soluble VEGFR-1 then liberates VEGF from the endogenous trap and allows its access to membrane receptors on endothelial cells, a step required for VEGF-driven angiogenesis (27). The MMP7 expression is associated with poor prognosis and often upregulated in many cancers such as, colon, liver, ovary, and pancreatic cancer (44-46) and show positive correlation with invasiveness of many cancer (47).

However, it's worth noting that, till date no studies available on MMP7's role in GBC pathogenesis.

Similar to MMP7, there are many reports suggest the role of MMP2 in tumour initiation and other key aspects of cancer such as, angiogenesis and metastasis (26). To study the same MMP2 knock out mice were used, where the tumour size, colony number in lungs and tumour induced angiogenesis was significantly reduced when injected with BL16 melanoma cells and Lewis lung carcinoma cells (13). Similarly, the role of MMPs in development of metastatic prostatic neuroendocrine cancer were evaluated by using transgenic mice expressing SV40 large T antigen in prostate neuroendocrine cells under transcription regulation of cryptidin2 gene (CR2-TAg), which is used to develop prostatic neuroendocrine cancer mice model. Where the MMP2 and MMP7 knock out showed decreased tumour burden, lung metastasis, blood vessel density, and increased survival rate, which emphasize the role of these MMPs in development of metastatic prostatic neuroendocrine cancer (48). Overall, these studies emphasize the role of these MMPs in various stages of cancer such as, tumour initiation, progression, and metastasis. As we already discussed, population specific genetic predisposition and aggressive metastasis are the important aspects of GBC pathogenesis. An understanding the genetic predisposition and underlying molecular mechanisms of these matrix metalloproteinases may be important for devising therapies aimed at preventing spread of tumour cells to nearby vital organs (6).

A comprehensive literature review shows many gaps in understanding the role of these MMPs in GBC. Firstly, despite the high prevalence of GBC in Odisha (Eastern Indian state) (49, 50), currently, no studies have been done to elucidate population-specific genetic predisposition. Secondly, even after late-stage diagnosis and poor prognosis, no comprehensive study is available to understand the genetic predisposition and underlying molecular mechanisms of tumour cell invasion and metastasis. Thirdly, except for an isolated report by Kirimlioğlu et al. from Turkey, no studies are available on MMP's expression in GBC (51). Lastly, single study

by Sharma *et al.* showed the association of genetic variants in MMP2, MMP7 and MMP9 as a risk factor for GBC (52). However, the underlying mechanism of these variants in GBC has yet to be explored. We thought it was important to address the aforementioned lacunae to gain more insight into the role of MMPs in GBC pathogenesis.

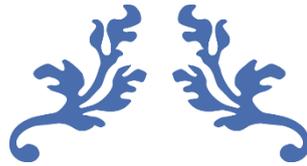
In the first chapter, we have presented our findings on the genetic association between MMP14 promoter variants and gallbladder cancer. Additionally, we have elucidated the functional role of the genetic variants rs1003040 and rs1003349 in regulating MMP14 expression, as well as assessed MMP14 expression in GBC through immunohistochemistry. Using the JASPAR database, we conducted an in-silico promoter analysis and identified transcription factors MYB and SOX10 binding sites at loci rs1003349 and rs1004030, respectively. We confirmed these findings through various assays, such as electrophoretic mobility shift assays, CHIP assay, luciferase reporter assays, and CRISPR-cas9-based deletion of the 119 bp genomic region surrounding the locus. We further evaluated the potential role of SOX10 and MYB in regulating the expression of MMP14 by Western blotting and conducted various cell-based assays, including wound healing, colony formation, proliferation, and invasion in two GBC lines, G415 and TGBC1TKB. Lastly, we carried out an in-silico correlation analysis of SOX10, MYB, and MMP14 expression in GBC using GEO dataset.

In the second chapter of the thesis, we presented our findings on the genetic association of MMP7 promoter variants with gallbladder cancer. Additionally, we investigated the putative role of MMP7 promoter variants rs113823671 A>C and rs17098318 G>A on the expression of MMP7. Additionally, we utilized immunohistochemistry to assess the level of MMP7 expression in GBC. To better understand the role of the associated SNPs in MMP7 expression, we conducted reporter luciferase assays and genotype-phenotype correlation studies.

In the third chapter, our study focused on examining the genetic association of MMP2 promoter variants with gallbladder cancer. Also, we carried out immunohistochemistry to assess the level

of MMP2 expression in GBC. Furthermore, we conducted genotype-phenotype correlation studies and reporter luciferase assays to shed light on the potential role of the associated SNPs in MMP2 expression.

In the final chapter of our thesis, we aimed to gain a better understanding of the biological pathways associated with MMPs in GBC by conducting pathway enrichment analysis using publicly available databases. To accomplish this, we examined gene expression profiles of GBC patients and normal samples from the GEO datasets. Using various analytical tools, such as DESeq2, GSEA, DAVID Bioinformatics Resources, Cytoscape, EnrichmentMap, UCSC Xena browser, and cBioportal, we identified and analyzed the molecular pathways and networks related to MMPs. Furthermore, we used the SMART App to explore the epigenetic contributions of MMP14, MMP7, and MMP2 in GBC.



CHAPTER 2

Review of Literature



Vinay J
NISER, Bhubaneswar

2.1 Gallbladder cancer

Gallbladder cancer is the most common gastrointestinal malignancy, with early liver and bile duct involvement leading to jaundice and quick organ metastasis (2). Its frequency varies in different geographic regions and within ethnic groups. It is more prevalent in Uttar Pradesh, Bihar, West Bengal, and Odisha state (also called the GBC belt, an endemic region with one of the highest incidences in the world) (53). These clustering of GBC incidence may be due to population specific genetic predisposition and various environmental risk factors.

An overview of the histopathologic characteristics and pre-cancerous conditions of GBC suggests that preneoplastic lesions may play a crucial role in the development of the disease (54). It was first reported by Yamamoto *et al.* that approximately 80% of GBCs follow a progression from dysplastic mucosa to carcinoma in situ and invasive carcinoma (55), with the primary sites of occurrence being the fundus (60%), body (30%), or neck (10%) of the gallbladder (56). The pathophysiology of GBC suggests that metaplastic alterations in the gallbladder mucosa, particularly intestinal or pseudopyloric kinds, are caused by long-term cholelithiasis and cholecystitis. These metaplastic alterations have been identified as precancerous conditions for GBC and provide further insight into the disease's pathophysiology (56, 57). Such metaplasia (particularly of the intestine type) causes epithelial dysplasia and carcinoma-*in situ*, and typically, it takes 15 years for dysplasia to turn into advanced gallbladder carcinoma (58). Unfortunately, the progression is usually swift and silent, indicating a poor outlook. Early detection and surgical resection are critical for a good prognosis. While approximately 50% of patients have lymph node metastases, less than 10% of patients have tumors that can be surgically removed at the time of surgery (59). The lack of a serosal layer next to the liver in the gallbladder allows for hepatic invasion and progression to distant metastatic tumors, contributing to the poor prognosis (2). Despite surgery, most patients

develop metastatic disease (60), emphasizing the need for effective primary prevention strategies to address early metastasis.

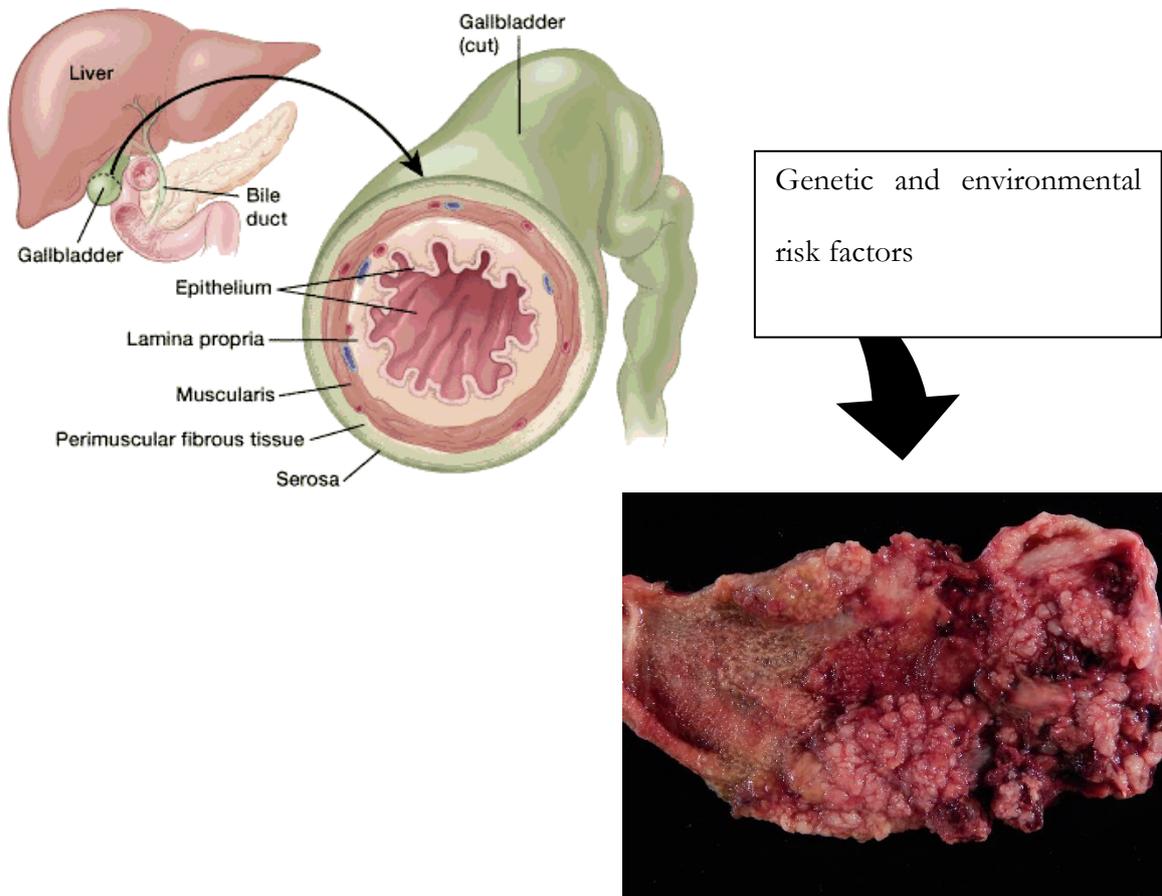


Figure 2.1.1 Anatomy and clinical features of the gallbladder and invasive gallbladder adenocarcinoma. *Gallbladder cancer gross image represents invasive adenocarcinoma of multifocal nodular/papillary proliferation in the fundus and body of the gallbladder. (The figure is adapted with permission from Akki et al., Gallbladder carcinoma. pathologyoutlines.com. Accessed November 21st, 2022)*

2.2 Pathology of gallbladder cancer

The pathobiology of gallbladder cancer is poorly understood. The papillary or tubular histologic subtypes make up the majority of gallbladder adenocarcinomas (85-97%), followed by sarcoma, neuroendocrine carcinoma, adenosquamous cell carcinoma, anaplastic carcinoma, and squamous cell carcinoma (61) (**Figure 2.2.1**).

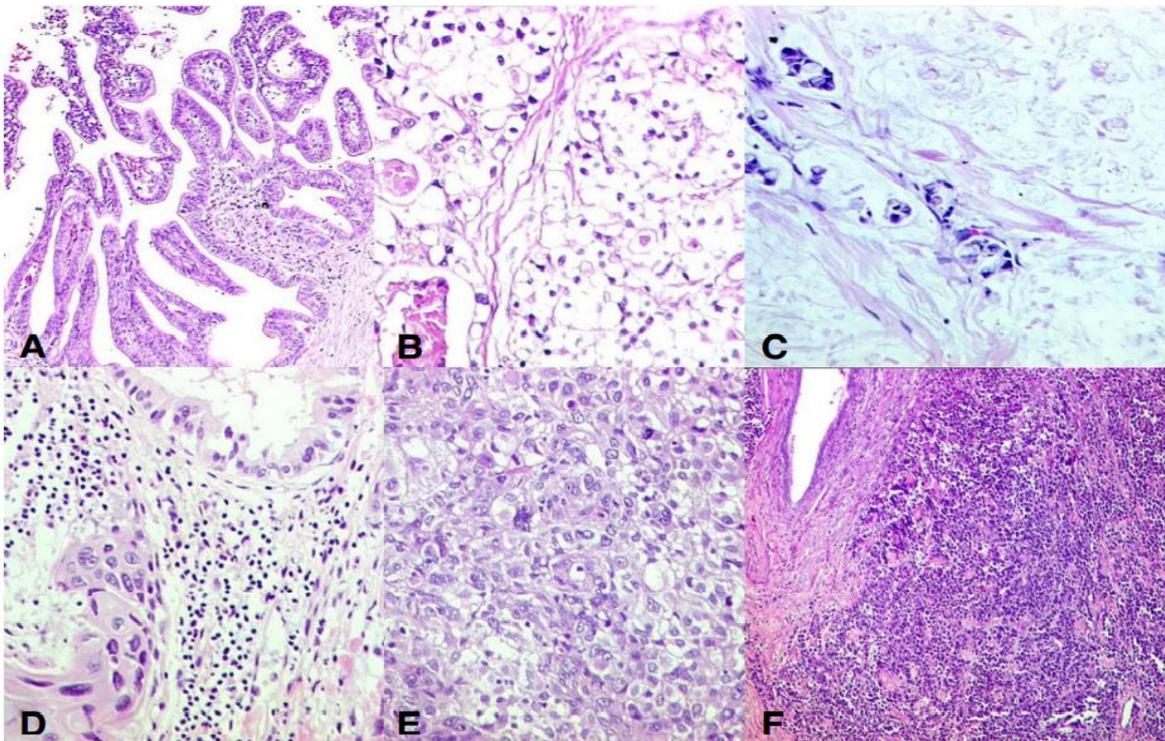


Figure 2.2.1 Histologic types of invasive gallbladder carcinoma. (A) *Papillary type*, (B) *Clear cell adenocarcinoma*, (C) *Mucinous type* (D) *Adenosquamous type*, and (E) *Undifferentiated carcinoma with pleomorphic giant cell type*. The figure is adapted with permission from Bal et al., 2015 (62).

There are two important pathways associated with gallbladder cancer pathogenesis. The most common one is the chronic inflammation-mediated pathway, and the second one is a somewhat less common pathway mostly observed in the population of Japan called congenital abnormality in bile duct development.

2.2.1 Chronic inflammation-mediated pathway

The gallbladder undergoes multistage histo-pathological and molecular changes starting with chronic inflammation to give rise to gallbladder cancer. The chronic-inflamed gallbladder often leads to the formation of preneoplastic lesions, leading to the metaplasia-dysplasia-carcinoma sequence (63).

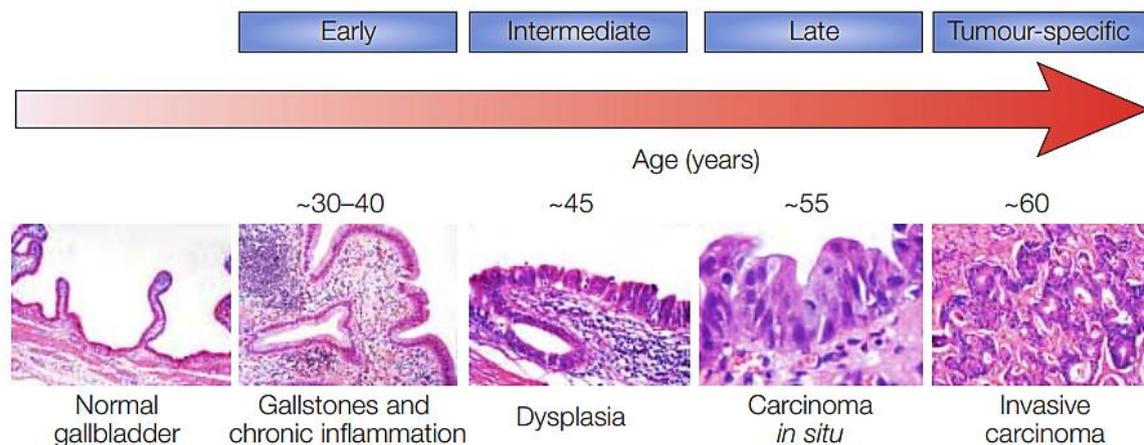


Figure 2.2.2 Sequential histological and molecular alterations in the pathology of gallbladder cancer linked to inflammation. An invasive carcinoma's initiation follows a well-defined series of consecutive flat-epithelial, premalignant alterations, including dysplasia and progression to carcinoma *in situ*. The picture also displays each histological change's median age at diagnosis. The figure is *Pathology of Gallbladder Carcinoma: Current Understanding and New Perspectives*. The figure is adapted with permission from Barreto et al., 2014 (63).

Wistuba and Albores-Saavedra presented the first illustration of the dysplasia-carcinoma cascade based on successive histological and molecular alterations in the etiology of gallbladder carcinoma due to gallstones and inflammation (64). This illustration was further enhanced by Wistuba and Gazdar, who gave the median ages at diagnosis for each histological alteration. The multi-step pathogenetic sequence described by Wistuba and Gazdar is diagrammatically shown in **Figure 2.2.2** (65).

Two main hypotheses can explain how most epithelial tumors, particularly glandular or adenocarcinomas, undergo malignant transformation. They are the adenoma-carcinoma sequence and the dysplasia-carcinoma sequence.

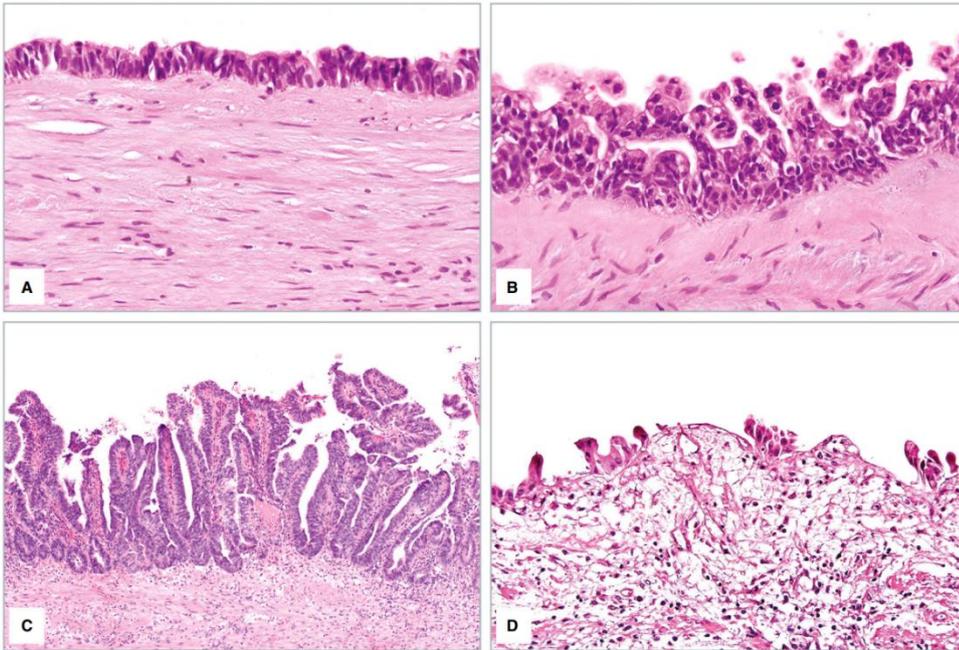


Figure 2.2.3 High-grade dysplasia/carcinoma in situ in the gallbladder types. (A) flat, (B) micropapillary, (C) tall papillary, or (D) denuding/clinging type. The figure is adapted with permission from Rao et al., 2021(66).

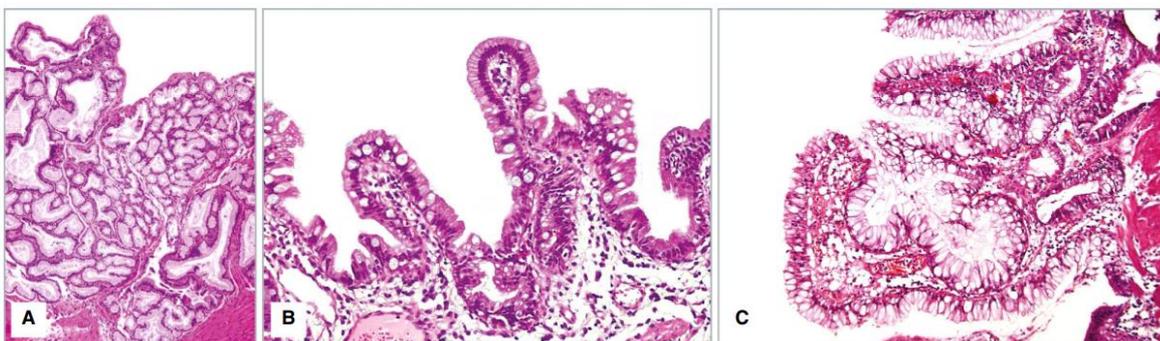


Figure 2.2.4 Types of gallbladder metaplasia. A 'Polypoid pseudo-pyloric gland type', B 'Intestinal type', and C 'Foveolar type'. The figure is adapted with permission from Rao et al., 2021 (66).

The first hypothesis is based on changes to the vesicular mucosal epithelium, where metaplasia typically manifests as an adaptive response following persistent irritation or inflammation. Repeated episodes of chronic cholecystitis may produce gastric-type pseudo-pyloric metaplasia or intestinal metaplasia. About 50% of cases of chronic cholecystitis show some form of metaplasia. The most frequent type of pseudo-pyloric metaplasia is pyloric (containing both neutral and acidic mucins) or foveolar (with neutral mucins). Goblet cells with acid mucin, which may also have neuroendocrine and Paneth cells, are a characteristic of intestinal metaplasia (67). This metaplasia develops dysplasia, leading to carcinoma-in-situ and invasive carcinoma (**Figure 2.2.2** and **Figure 2.2.3**).

Given that >80% of gallbladder cancers are adenocarcinomas, the significance of these cellular and molecular alterations leading to the 'metaplasia–dysplasia–carcinoma' sequence cannot be undermined. This is paradoxical because gallstones-mediated chronic inflammation serves as the primary hypothesized triggering cause. The inflammation alone should have resulted in more squamous differentiation than adenocarcinoma (68).

The second hypothesis suggests that a benign glandular tumor, like an adenoma, can turn cancerous over time (69, 70). Experimental and clinical evidence supports both hypotheses; however, each pathway has a different significance and meaning in certain cases. Different circumstances and times would require these injuries to evolve and transition into more aggressive forms. Finally, evidence suggests that the adenoma-carcinoma pathway is not the most important in the GBC and exhibits molecular changes distinct from those seen in the dysplasia-carcinoma pathway (71).

2.2.2 Congenital abnormality in bile duct development model

A congenital deformity known as an abnormal junction of the pancreaticobiliary duct causes the pancreatic duct to empty into the biliary system outside the duodenal wall. This abnormality

can easily be found in cholangiography using magnetic resonance cholangiopancreatography, endoscopic retrograde cholangiopancreatography, or endoscopic ultrasound (EUS) imaging. The abnormality is scarce in western nations, but it is more evident in Asians, especially those of Japanese descent (72). A common channel that is too lengthy could undermine the sphincter of Oddi's ability to act as a gatekeeper, allowing pancreatic secretions to reflux into the bile ducts and gallbladder and potentially causing malignant alterations in the mucosa (73). Although women are more likely to develop this anomaly, it is not associated with gallstones. This junction anomaly occurs in around 10% of patients with gallbladder cancer (59). The disease here typically presents histologically as a papillary carcinoma (74). The patients with congenital abnormality of the bile duct show incidence of lymph node metastasis in 71% of the cases at the time of initial diagnosis or surgery. Nevertheless, prophylactic cholecystectomy should be considered in patients with an abnormal pancreaticobiliary duct junction (75).

2.2.3 Preneoplastic lesions of Gallbladder Carcinoma

Biliary tract cancers follow the same stepwise tumour growth pattern as most gastrointestinal tumors. The pre-invasive stage is followed by a clearly defined and physically distinct invasive stage. Characterization of pre-invasive lesions is of utmost importance since it allows for developing screening methods for high-risk populations and preventing cancer in its earliest stages. Pre-invasive lesions also provide crucial links in our understanding of carcinogenesis. An increase in knowledge of pancreatic precursor lesions in recent years has reignited a keen interest in its gallbladder counterparts. It is evident that similar to pancreatic cancer, biliary cancers are preceded by two different types of pre-invasive intraepithelial neoplasia (the flat or non-tumoral type and the tumoral or mass-forming type) (76). Numerous premalignant lesions have been reported to exist in the gallbladder. Over time, a huge vocabulary of confusing terms ('pyloric gland adenoma,' 'papillary adenoma,' 'tubulo-papillary adenoma,' 'biliary adenoma,' 'intestinal adenoma,' 'transitional adenoma,' 'papillary neoplasm,' 'papillary carcinoma,' and

‘intracystic papillary neoplasm’) for overlapping morphological features has developed as a result of pathologists' inconsistent nomenclature usage and the lack of experience with these lesions (77-82).

2.3 Epidemiology of gallbladder cancer

2.3.1 Gallbladder cancer demographics

GBC is known to have varying incidence rates across the globe, which suggests that certain geographical regions and ethnicities may have a higher predisposition towards the disease. Countries like Japan, Chile, India, Pakistan, Bangladesh, Nepal, Bhutan, Slovakia, and Bolivia have reported higher rates of GBC occurrence, while central Europe and western countries like the USA, Canada, Australia, UK, and New Zealand have reported lower rates (83). The estimated number of new GBC cases in worldwide and Asia are shown in **Figure 2.3.1** and **Figure 2.3.2**

GBC is the most common cancer among Chilean women, far more common than breast and cervical cancer. It is presently the most common gastrointestinal cancer in Shanghai, China, where its frequency is growing (84).

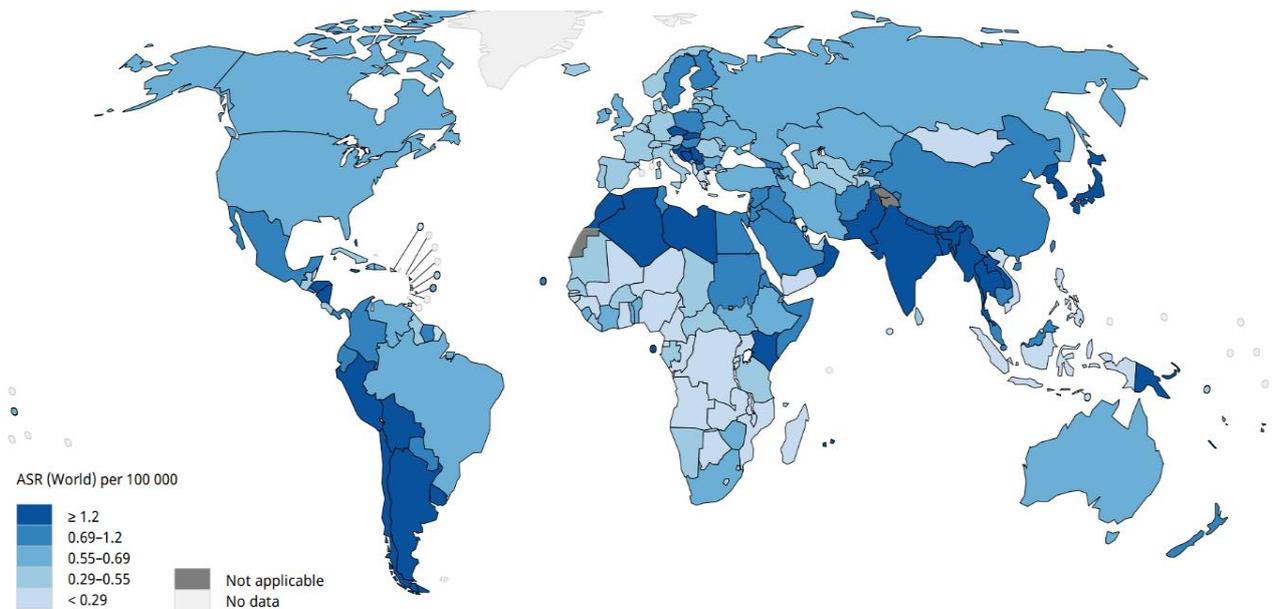


Figure 2.3.1 The global gallbladder cancer incidence. (The figure is adapted with permission from *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer. Available from: <https://gco.iarc.fr/today>, accessed November 21st, 2022.)

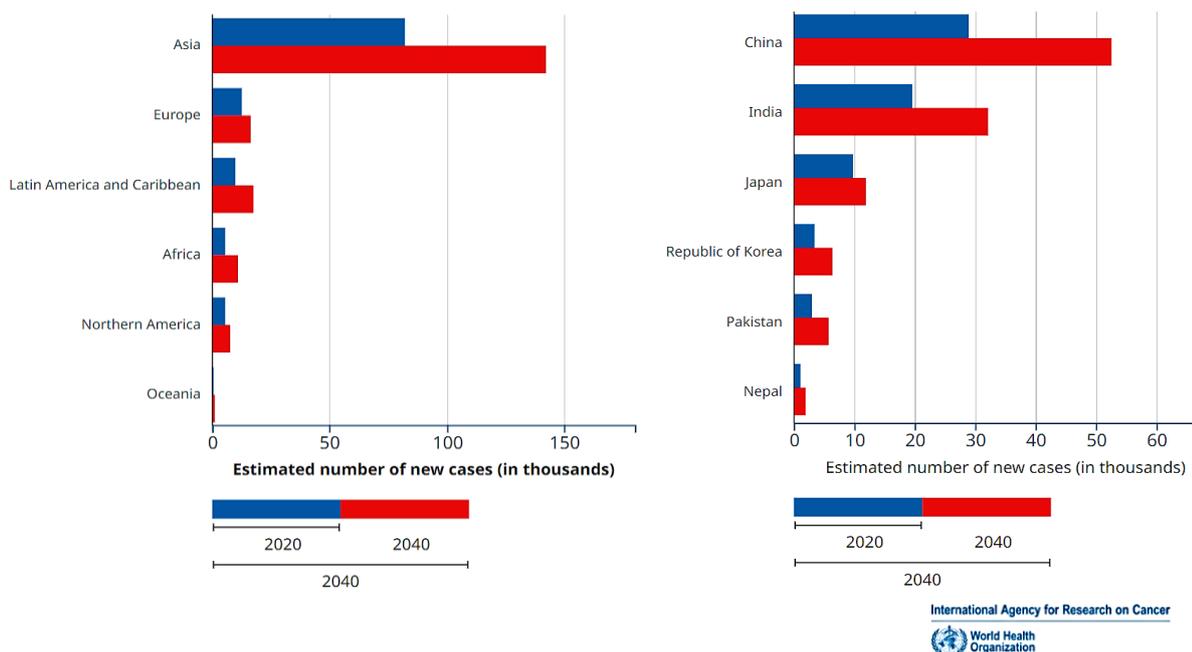


Figure 2.3.2 Estimated number of new gallbladder cancer cases in worldwide and Asia from 2020-2040 (both gender, age (0-85+)). (The figure is adapted with permission from *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer. Available from: <https://gco.iarc.fr/tomorrow>, accessed November 21st, 2022.)

The incidence of GBC is particularly high in Northern and North-Eastern India, along with other states like Uttar Pradesh, Bihar, Odisha, and West Bengal, which comprise the GBC belt (85). As per GBC reports in North Indian cities, Women are two to six time more affected than men and usually develops in individuals over 50 to 65 years old (86). According to six cancer registries of Indian Council of Medical Research (ICMR) (1990-96), the incidence rate of GBC in South Indian region is ten times lower than North Indian region. As shown in **Figure 2.3.3**, the incidence rate in northern cities such as, Delhi (3.7 for men and 8.9 for women) and Bhopal (1.6 for men and 2.5 for women) are significantly higher compare to southern cities such as, Chennai (0.5 for men and 0.8 for women) and Bangalore (0.6 for men and 0.7 for women) (87, 88). As per, ICMR registries (2006-2008), the highest GBC incidence is observed in major Indian cities such as, Bhopal, Kolkata, Delhi, Dibrugarh, and Mumbai. Also, it ranks amongst first 10 most frequent cancers in the region (**Figure 2.3.4**) (89). Overall, there are notable gender, geographic, and ethnic differences in the global incidence, which points to a significant role for genetic and environmental variables in its pathogenesis.

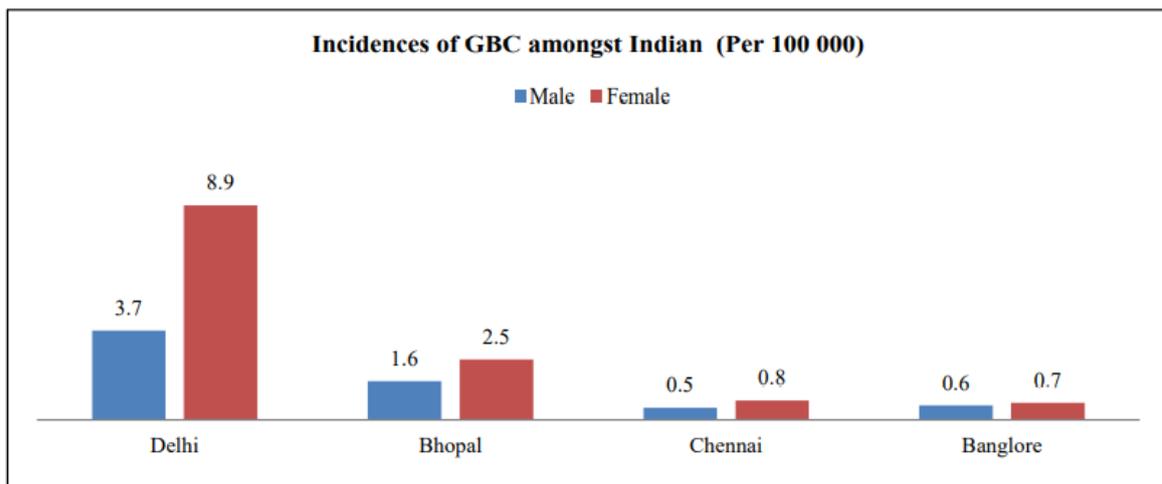


Figure 2.3.3 Age-adjusted incidence rates (AAR) of gallbladder cancer between Northern and Southern cities of India for both genders. *Source: Adapted from reference (87)*

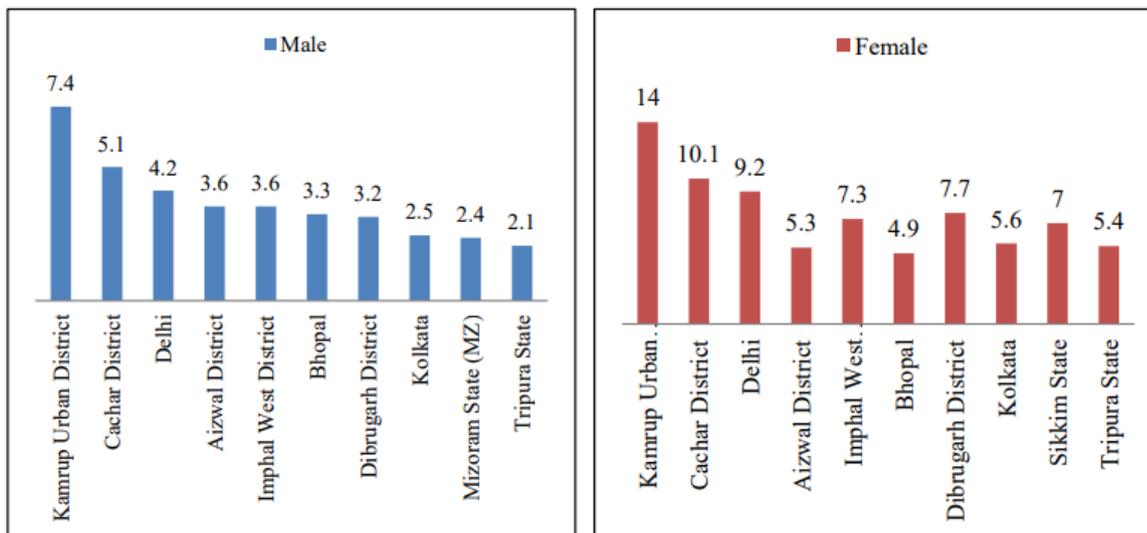


Figure 2.3.4 Age-adjusted incidence rates (AAR) of gallbladder cancer across all population-based cancer registries in India for both genders. *Source: Adapted from reference (88)*

2.3.2 GBC Related Mortality

Gallbladder cancer has very poor prognosis. In the initial stage of cancer, it does not have specific symptoms, and it is usually found by accidental diagnosis with other gastrointestinal diseases. One reason for the high mortality rate in GBC could be the manifestation of visible symptoms in an advanced cancer stage (90). The median survival time for individuals with advanced GBC is six months, with a 5% five-year survival rate for stage IV and a 32% five-year survival rate for lesions limited to the gallbladder mucosa (2, 91). Countries with the highest mortality rates due to GBC are Bolivia, Bangladesh, Chile, Nepal, Bhutan, Pakistan, and India (**Figure 2.3.5**)(83).

Gallbladder cancer accounts for 1.2% of all diagnosed cancers and 1.7% of all cancer-related deaths (92), showing that Gallbladder cancer accounts for more mortality than the incidence rate (92). The absence of a serosal layer next to the liver, which enhances the likelihood of metastasis through hepatic connective tissue, is likely responsible for the higher mortality rate (2, 91). According to Global Cancer Statistics 2020, there were approximately 115,949 new

gallbladder cancer cases, with 41,062 men and 74,887 women. In total, 84,695 people died from it in 2020, of which approximately 30,265 were men and 54,430 were women (83). The age-standardized incidence rate for males and females are approximately 0.9 and 1.4, while age-standardized death rate, it is 0.7 and 1.0 cases per 100,000. Similarly, the 2020 report also indicated that women are three times more likely to die from gallbladder cancer than men (83). Overall, these reports suggest that the incidence of gallbladder cancer and mortality rates worldwide show a more significant gender discrepancy.

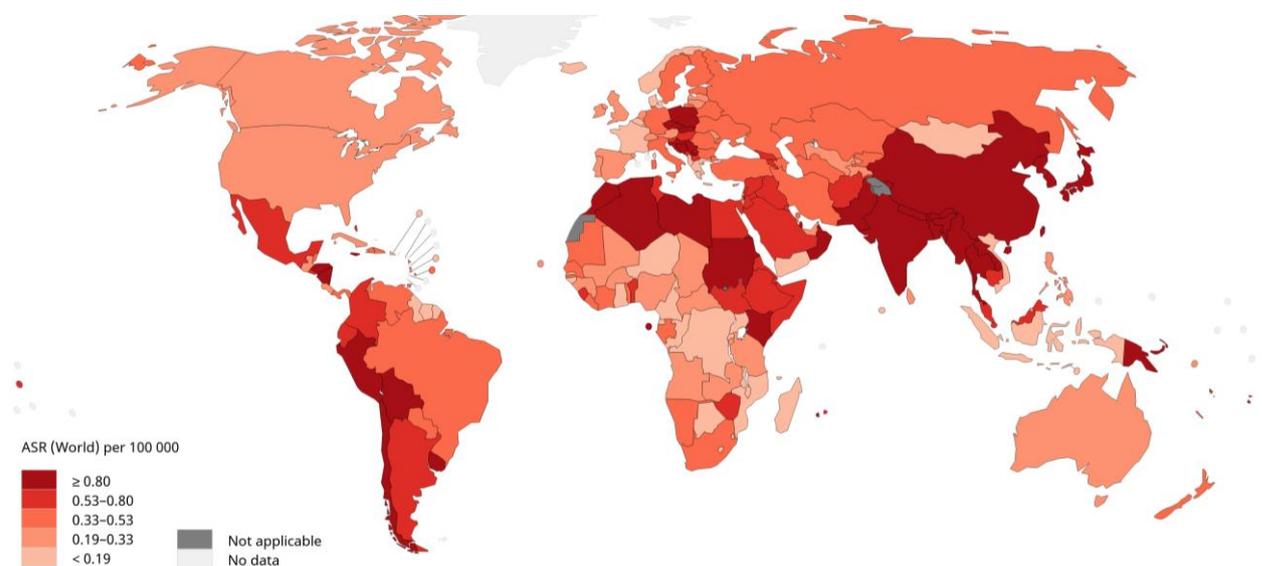


Figure 2.3.5 Estimated age-standardized mortality rates of GBC worldwide in 2020. (The figure is adapted from *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer. Available from: <https://gco.iarc.fr/today>, accessed November 21st, 2022.)

2.4 Etiology and risk factors

Gallbladder cancer is known to be caused by various environmental and genetic risk factors. Chronic inflammation, exposure to certain chemicals such as radon gas and dichloro diphenyl trichloroethane (53), and heavy metals such as chromium, lead, arsenic, and zinc (93) are among the environmental factors that have been linked to the development of GBC. Additionally, some dietary factors have also been found to contribute to its development (53, 94). The clustering

of cases in specific geographical locations further supports the notion of environmental, ethnic, and genetic risk factors in the pathogenesis of gallbladder cancer. Furthermore, the disease is more prevalent in females, potentially due to hormonal changes (95). The limitations of available epidemiological research, such as small sample sizes, differences in clinical and pathological criteria, and difficulty in determining exposure to any presumptive risk factor, have hindered our understanding of the disease's causes (96). Identifying risk factors is crucial for understanding the pathogenetic mechanisms driving geographic and ethnic diversity, as well as for developing prevention and treatment strategies.

2.4.1 Gender and age

The prevalence of GBC is higher among women compared to men, with the disparity being most evident in northern India, Pakistan, and among American Indian females. Women are affected at a rate of two to six times more than men (96). The correlation between GBC and high parity provides further evidence of the role of female sex hormones. Although the impact of oestrogen receptor and progesterone receptor expression on gallbladder cancer is unknown, females with GBC display higher co-expression of both receptors compared to males, indicating a possible target for therapeutic intervention (97).

A study conducted in Massachusetts General Hospital involving 402 patients with gallbladder cancer revealed that the median age of the patients was 72 years, and 72.3% of them were female (98). Similarly, a retrospective study conducted over a 10-year period at the MSK Cancer Centre involving 435 GBC patients reported that the median age of patients was 67 years (99). However, in India, the average age at which GBC is diagnosed is lower than that in Western countries, with patients being diagnosed at an average age of 53.39 ± 11.12 years compared to 70 ± 10 years in the West (100). The incidence of GBC increases with age, with the risk starting from 30 years old, thus emphasizing the importance of early detection even for younger patients (**Figure 2.4.1**)(101).

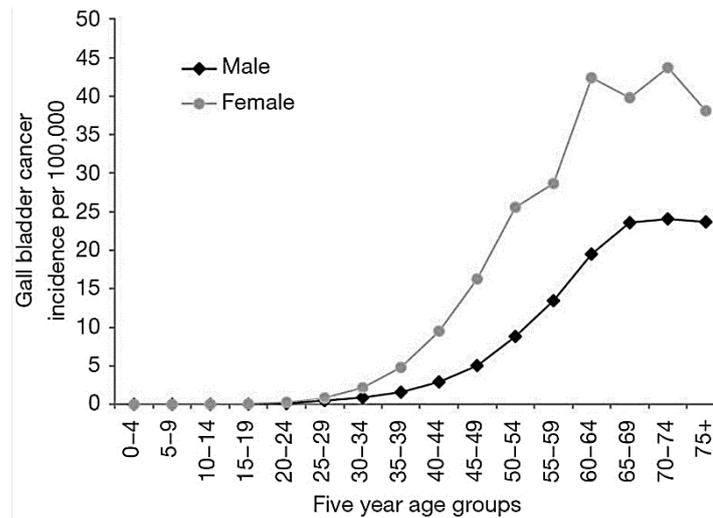


Figure 2.4.1 The gallbladder cancer risk increases after the age of 30. This graph shows the incidence from Indian population for various age and gender groups. Adapted from reference (95).

2.4.2 Gallstones

Gallstones have been identified as a major risk factor for GBC development, as the majority (85%) of individuals with GBC also have gallstones. The prevalence of gallstone disease is higher in certain indigenous communities in North and South America, and there is a significant association between the prevalence of gallstone disease and GBC incidence (**Figure 2.4.2**) (102-104). Studies have shown that patients with gallstone disease have a significantly higher chance of developing GBC, with a Shanghai-based case-control study reporting a 34-fold increase in risk (105). According to an autopsy data-based study from Chile, patients with gallstones have a seven-fold higher risk of developing GBC than patients without gallstones.

The risk increases as cholelithiasis lasts longer and gallstones get larger (106). In fact, the relative risk of GBC in people with gallstones is 2.4 if the diameter is 2.0-2.9 cm, but it rises to 10.1 if it is larger than 3.0 cm (105, 107). However, in a recent Swedish study, the overall incidence of gallbladder cancer in patients with gallstones was found to be 0.5% (108), implying

that other genetic or environmental factors may also contribute to the development of the disease (109, 110).

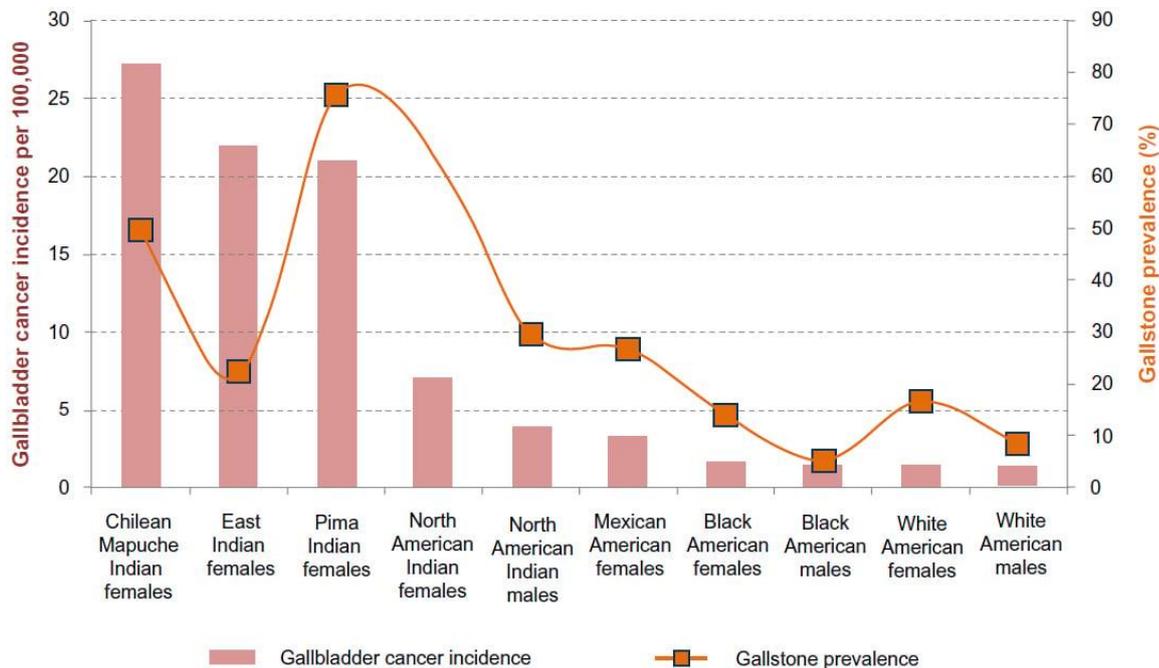


Figure 2.4.2 The Gallstone prevalence and the occurrence of gallbladder cancer in high-risk populations. Gallstones are a significant risk factor for gallbladder cancer because both the incidence and prevalence of gallstones are higher in some ethnic groups. Adapted from reference (2).

2.4.3 Obesity

Obesity, defined as having a BMI of 30 or higher, has been found to be associated with an increased risk of gallbladder cancer, with a relative risk of 1.88 (95% CI: 1.66-2.13) (2, 53). A multicentre study conducted in Europe also revealed that obesity raises the incidence of gallbladder cancer, with an adjusted relative risk of 2.1 (95% CI: 1.2-3.8) observed between the highest and lowest BMI quartiles (104). Moreover, there is a 1.06-fold increase in risk for GBC for every unit increase in BMI (111). Interestingly, research conducted in India indicates that individuals with GBC have lower BMIs compared to patients with gallstone disease or healthy control subjects. However, this may be due to weight loss caused by secondary cancers rather than a causative relationship between low BMI and GBC (112). Nevertheless, further

comprehensive studies are necessary to better comprehend the correlation between obesity and other potential factors that may contribute to the inverse relationship across different populations.

2.4.4 Chronic inflammation

The chronic inflammation of the gallbladder may be caused by various factors such as calcium deposition in the gallbladder wall (porcelain gallbladder), chronic bacterial infections, and primary sclerosing cholangitis (113). The extent of porcelain gallbladder being identified in diagnostically confirmed GBC is less than 1% of total gallbladder cases. The controversy presents whether porcelain gallbladder a true risk factor or not. Some reports are stating, on average, 25% of cases are associated with gallbladder cancer (113, 114), while more recent study suggest no association (115). In cases of partial calcification, spotted or dappled calcification, which are known to be premalignant, it is recommended to remove the glandular areas of the mucosal lining prophylactically (116). The hardening of the muscularis lining does not seem to be a sign of gallbladder cancer (117).

Chronic bacterial infections can also cause irritation and inflammation in the gallbladder and are considered a risk factor for GBC. Many bacteria have been identified by PCR or by culturing specimens from patients with cholecystitis and cholelithiasis, such as 'Escherichia coli, Klebsiella pneumoniae, Citrobacter freundii, Salmonella spp., Helicobacter spp., Enterobacter spp., Enterococcus spp., Pseudomonas aeruginosa, Bacteroides fragilis, Staphylococcus aureus, Proteus spp. and Acinetobacter spp' (118-125). Carriers of *S. typhi* are 8-12 times more likely to develop gallbladder cancer, with a 6% probability (126). Gallbladder cancer is also associated with *Helicobacter bilis*, with risk ratios of 6.5 in Japanese and 5.86 in Thai patients (127). Patients with primary sclerosing cholangitis (PSC) are often at a higher risk for gallbladder cancer, as dysplasia occurs in 37% and adenocarcinoma occurs in 14% of

gallbladders from primary sclerosing cholangitis (PSC) patients due to their general tendency for biliary carcinoma (128).

2.4.5 Primary sclerosing cholangitis

A chronic liver illness called PSC causes the bile ducts both inside and outside of the liver to swell, scar, and finally become constricted or blocked. This results in bile accumulation in the liver, leading to further liver damage. PSC is a chronic fibroinflammatory condition that has been linked to cancer development through persistent inflammation (129). A high frequency of inflammatory, metaplastic, and neoplastic changes has been found in the gallbladders of patients undergoing liver transplants for PSC. In fact, gallbladders fully embedded in end-stage PSC exhibit higher rates of pyloric metaplasia, intestinal metaplasia, dysplasia, and invasive adenocarcinoma than the general population (128). Adenocarcinomas in PSC patients develop from a background of flat type mucosal dysplasia as shown in **Figure 2.2.3**, indicating a metaplasia-dysplasia-carcinoma sequence similar to that seen in sporadic gallbladder carcinogenesis. Also, “field effect or chronic inflammation” in the intrahepatic and extrahepatic biliary tree is supported by the tight relationship between gallbladder neoplasia and biliary neoplasia in PSC patients (128). Therefore, those with PSC should undergo annual abdominal ultrasound screening for lesions in order to monitor for gallbladder cancer. Lesions more than 0.8 cm should be treated with a cholecystectomy (130).

2.4.6 Genetic factors

Gallbladder cancer is a result of a combination of genetic susceptibility and environmental risk factors. Genetic factors play a significant role in influencing the frequency of gallbladder cancer (2). However, there is limited understanding of the molecular alterations involved in the pathogenesis of gallbladder cancer compared to other neoplasms (65). Research on molecular

alterations in gallbladder cancer patients has primarily been done in Chile and Japan due to the high incidence of this neoplasm in these countries (108).

Somatic mutations: Analysis of somatic mutations revealed that the majority of gallbladder cancer patients exhibited alterations in certain canonical pathways. Specifically, TP53 alterations were present in 74% of samples, followed by cell cycle alterations in 46% of samples (64) and RTK-RAS alterations in 46% of samples (131). Other common alterations included TGF β in 28% of samples and NOTCH in 9% of samples (132, 133). Further research demonstrated that common alterations in GBC were observed in TP53, SMAD4, NOTCH1, ERBB2, PIK3CA, MET, and PTEN genes (64, 133). Studies investigating the genetic evolution of gallbladder adenomas to invasive carcinomas have shown that mutations in CTNNB1, KRAS, and PIK3CA may be initiating events during carcinogenesis and could play a significant role in the progression of cell from premalignant to malignant states (134).

KRAS-activating point mutations are the most frequently observed among GBC patients, particularly at codons 13 and 61 (135). Although KRAS mutations are less common in western countries, they have different incidence rates in Japan (136, 137). Studies show that KRAS mutations in gallbladder cancer are more common in anomalous pancreatobiliary duct junction patients in Japan due to pancreatic juice reflux (138-141). In GBC, TP53 anomalies are frequently observed, and the incidence of these mutations does not seem to vary by region (142, 143).

Genome-wide association studies (GWAS): Several GWAS have identified genetic variants associated with GBC. For instance, rs11887534 in the ABCG8 gene and rs12882491 in the TRAF3 gene were found to be associated with GBC in an admixed Chilean Latinos with Mapuche Native American ancestry (144). A large-scale GWAS conducted in the Indian population identified genetic variants rs1558375 and rs4148808 in ABCB4, and rs17209837 in ABCB1 to be significantly associated with GBC (145). Additionally, GWAS in the Japanese

population found rs7504990 in the DCC gene to be associated with gallbladder cancer (146). The results of these GWAS highlight population-specific genetic predisposition as an important risk factor in GBC, especially in high-risk groups such as, India (145), Chile, Latin America (144), and Japan (147).

Candidate gene-based association studies: A study on the north Indian population has revealed that genetic variants of DNA repair pathway genes, such as ERCC2 rs1799793 Asp312Asn, MSH2 rs2303426 IVS1+9G>C, rs2303425 -118T>C, and OGG1 rs2072668 748-15C>G, play a role in the development of gallbladder cancer (148). Other studies from China and India have also shown that genetic variants of inflammatory pathway genes, including TGFb1 (rs1800469)-509C>T, TNF- α (rs1800629) -308G>A, IL6 (rs1800795) 236C>G, IL1B (rs16944) -1060T>C, and IL8 (rs10805066) -13985C>G, are associated with gallbladder cancer (149-153). Additionally, several candidate gene-based genetic association studies have suggested the involvement of genes related to lipoprotein metabolism (154-156), MMPs (52, 100), DNA repair (148, 157), SERPINs (158, 159), fatty acid metabolism, steroidogenesis, and inflammatory pathways (1). For instance, in a study of the Indian population found that SNPs in MMP2 (rs2285053, rs243865), MMP7 (rs11568818), and MMP9 (rs17577) are associated with GBC (52). An overview of candidate gene-based genetic studies reported in gallbladder cancer till date is provided in **Table 2.4.1**.

Table 2.4.1 Genetic predisposition studies of candidate genes in gallbladder cancer*

Gene	Polymorphism	Population	Ref.
DNA repair pathway genes			
XPC	rs2228000, rs2228001	China	(160)
ERCC2	rs1799793, rs13181	North Indian	(148)
MSH2	rs2303426, rs2303425	North Indian	
OGG1	rs2072668	North Indian	
TP53	rs1042522	Chilean, Hungary, Japanese	(161-163)
XRCC1	rs1799782	North Indian Shanghai, China	(164, 165)
APEX1	rs3136820	Shanghai, China	(164)
RAD23B	rs1805335, rs1805329	Shanghai, China	(166)
FEN1	FEN1-69G>A	China	(167)
Hormonal pathway genes			
CCKAR	rs1800857	North Indian	(168)
CCK and CCKAR	rs2071011G>C, rs915889C/T, rs3822222C/T, rs1800855T/A	Shanghai, China,	(169)
ESR1	rs2234693, rs3841686, rs2228480, rs1801132, rs9340799	Shanghai, China, North India	
ESR2	rs1256049		
PGR	Ins/Del		
AR	(CAG)n	Shanghai, China	(171)
COMT	rs4818	Shanghai, China	(172)
CYP1A1	rs2606345		
CYP1B1	rs10012		
CYP19A1	rs1065778, rs700518, rs2304463, rs700519, rs1065779, rs4646		
HSD3B2	rs1819698, rs1361530		
HSD17B3	rs2066479		
HSD17B1	rs2830		
SHBG	rs6259		
SRD5A2	rs523349		
RXR-a	rs1536475		
RXR-b	rs1805343, rs2744537, rs2076310	Shanghai, China	
INS	rs689		
PPARD	rs2016520		
PPARG	rs3856806		
Inflammatory pathway genes			
CR1	rs2274567, rs12144461	North Indian	(174)
IL1RN	86-bp VNTR, rs689466	North Indian	(151)
PTGS2	rs20417, rs5275	North Indian	(175)

IL1B	rs16944	Shanghai, China north Indian	(175, 176)
IL10	rs1800871, rs1800872	Shanghai	(150, 151)
IL-8	rs10805066	China	(149)
EGF	rs4444903	North Indian	(153)
TGFb1	rs1800469	Shanghai, north Indian	(150, 152, 153)
TNF-?	rs1800629		
IL6	rs1800795		
IL8	rs10805066	China	(149)
MMP-2	rs2285053, rs9340799	North Indian	(52)
MMP-7	rs11568818, rs2250889		
MMP9	rs17576, rs17577		
TIMP2	rs8179090		
Metabolic pathway genes			
MTHFR	rs1801133	Indian	(177)
APOB	rs17240441	Indian	(154)
NAT2	rs1799929, rs1799930, rs1799931	Indian	(178)
GSTT1	Null polymorphism	Indian	(179)
GSTP1	rs1695		
CYP17	rs743572	Shanghai Indian	(180, 181)
GSTM1	Null polymorphism	Indian, Chilean Hungary Japanese	(161, 162, 179)
CYP1A1	rs4646903, rs1048943	Indian, Chilean Hungary Japanese	(161, 163, 172, 182)
Cyp1a1 cyp1b1	rs5930	Shanghai	(183)
LDLR	rs6413504, rs14158		
LPL	rs263		
ALOX5	rs2029253		
ApoB	rs693	Indian Chilean	(154, 184)
ABCG8	rs11887534	North Indian Shanghai China	(106, 136)
CETP	rs708272, rs1800775	Chilean Shanghai China	(156, 184)
LRPAP1	rs11267919	North Indian Shanghai China	(155, 156)
CYP7A1	rs3808607, rs3824260	North Indian	(185)
CYP17	rs743572	North Indian	(180, 181)
ApoB	rs676210, rs673548, rs520354, rs1367117, rs440446	Shanghai	(183)
CYP2C19	rs4244285, rs4986893	Japanese	(186)
ADRB3	rs4994	North Indian	(187)
Apoptosis pathway			

CASP8	rs3834129, rs1045485, rs3769818	North Indian	(188)
Nuclear Receptors			
Lxr-alpha, Beta	rs7120118, rs35463555, rs2695121	North Indian	(189)
Cancer Stem cell gene			
CD44	rs13347, rs353639, rs187116, rs187115	North Indian	(190)
NANOG	rs11055786	North Indian	(191)
ALCAM	rs1157		
EpCAM	rs1126497		
SOX-2	rs11915160		
OCT.4	rs3130932		
NANOG	rs11055786		
Prostate stem cell antigen			
PSCA	rs2294008, rs2978974	India, Japan	(192, 193)
miRNA			
hsa-miR-146a	rs2910164	North Indian	(194)
hsa-mir-196a2	rs11614913	North Indian	
hsa-mir-499	rs3746444	North Indian	
miR-27	rs895819	North Indian population	(195)
miR-570	rs4143815	North Indian population	
miR-181	rs12537	North Indian population	
GWAS-associated genes			
DCC	rs7504990	Japan	(147)
	rs2229080	North Indian	(196)
	rs4078288		
	rs7504990		
	rs714		
Wnt signalling pathway			
SFRP4	rs1802073	North Indian	(197)
DKK2	rs17037102		
DKK3	rs3206824		
APC	rs4595552, rs11954856		
AXIN-2	rs4791171		
β-CATENIN	rs4135385		
GLI-1	rs222826) C>G		
Other genes			
KRAS	codon 25 Gln25His	Eastern India	(198)
ACE I/D	rs4646994	North Indian	(199)
DNMT3B	rs1569686	North Indian	(200)
TLR2	196-174del	North Indian	(201)
TLR4	rs4986791		
Adrenergic receptors (ADRA)	ADRA2A (C-1291G), ADR3 (T190C or Trp64Arg), ADR1 (C1165G or Arg389Gly)	North Indian	(202)

DR4 FASL	rs20575, rs20576, rs6557634, rs2234767, rs763110	North Indian	
PICE1	rs2274223, rs7922612	North Indian	(203)
Vitamin D receptor (VDR)	FokI C>T	China	(204)

*[Adapted with permission from reference (1)].

2.5 Major pathways associated with gallbladder cancer

2.5.1 PI3K/AKT/mTOR Signalling Pathway

Important cellular functions such as cell growth, motility, differentiation, metabolic activity, and apoptosis are known to be regulated by the PI3K/AKT/mTOR signalling pathway. Substantial evidences show that the PI3K/AKT/mTOR pathway is one of the most important signalling pathways that is strikingly increased in a variety of malignancies, including ‘breast cancer, gastric cancer, hepatocellular carcinoma, colorectal cancer, pancreatic cancer, cholangiocarcinoma, and gallbladder cancer’ (4, 205-207). Cell surface ligands initiate a signalling cascade by interacting with membrane receptor kinases such as G protein-coupled receptors (GPCR), epidermal growth factor receptors (EGFR), and vascular endothelial growth factor receptors (VEGFR). Phosphorylating PIP2 (phosphatidylinositol bisphosphate) to PIP3 activates the PI3K pathway (phosphatidylinositol triphosphate). PTEN, a tumoursuppressor protein that dephosphorylates PIP3 to PIP2, influences how AKT and PDK1 are activated later on (208, 209). Phosphorylated AKT either directly activates mTOR or indirectly activates mTOR by inactivating the TSC1/2 (tuberous sclerosis complex 1/2) complex, which is an inhibitor of mTOR. mTORC1 (mTOR, mLST8, PRAS40, and raptor) and mTORC2 (mTOR, Sin1 mLST8 and rictor) are the two mTOR complexes. In which p70SK6 (S6 kinase 1) is activated by mTORC1 and helps the dissociation of the inactive bound form of eIF4E (4E-BP1) to produce active eIF4E, which in turn drives cell growth and proliferation (210-212).

Recent studies have revealed the importance of this pathway in the growth and proliferation of gallbladder cancer and the therapeutic potential of its potential inhibitors (213, 214). In patients with dysplasia (66.7%), early gallbladder cancer (84.6%), and advanced gallbladder cancer (88.3%), the mTOR substrate p70S6K is phosphorylated. Additionally, 24% of individuals with chronic cholelithiasis and 64.1% with gallbladder cancer expressed phosphorylated mTOR. Advanced gallbladder cancer's prognosis was well predicted by the presence of phospho-mTOR (214, 215). There have been reports of PI3K pathway mutations in 22% of early-stage and 14.6% of advanced-stage gallbladder cancer (216). Activating mutations in PIK3CA were found in 12.5% of patients with gallbladder cancer (217). Gallbladder cancer exhibits epigenetic modification, mutation, and abnormal activation of the PI3K/mTOR pathway (218).

2.5.2 MAPK/ERK Signalling Pathway

The MAPK signalling pathway, which is essential for intracellular signal transduction and controls a variety of cellular functions, is often disrupted in cancer. Mitogen-activated protein kinase (MAPK) signalling sends growth and stress signals from the extracellular environment to intracellular machinery to carry out numerous cellular tasks like proliferation, migration, and apoptosis. The MAPKKK, MAPKK, and MAPK are the three main kinases that make up the MAPK signalling cascade. Specific ligands interact to the membrane receptor kinases/Ras complex to activate MAPKKK (Raf, ASK1, MLK, and MEKK1/4). As a result of MAPKK phosphorylation (MEK1/2, MKK4/7, and MKK3/6), which in turn phosphorylates MAPK (ERK1/2, JNK, and p38), leading to transcription of their target genes (219, 220). In order to keep the balance between apoptosis and proliferation in check, the MAPK/ERK pathway must be controlled. Due to common mutations in the RAS or RAF, this pathway is always active in the majority of malignancies (135).

The ERK/MAPK pathway was frequently activated in gallbladder cancer and chronic cholecystitis patients in a high-risk Chilean community (221, 222). By activating the

MAPK/ERK pathway, the mitochondrial glutamate transporter Solute Carrier Family 25 Member 22 (SLC25A22) encourages the growth and spread of gallbladder cancer (223). Prohibitin (PHB), a scaffold protein, was discovered to be a strong prognostic marker for gallbladder cancer, as it increases cell proliferation and invasion via the ERK pathway (224). Similarly, CCR7 (CC-chemokine receptor 7)-mediated expression of tumour necrosis factor (TNF)- increases lymph node metastasis in gallbladder cancer via activation of the ERK1/2/AP-1 and JNK/AP-1 pathways (225, 226). A systemic examination of common mutation and pathway activation found mutation and abnormal expression in MAP kinase, Wnt/ β -catenin and NF- κ B. Among the three pathways, the MAPK pathway had the highest mutation burden, with 50% of mutation in key signalling molecules including ADAM12, MAP2K1/MEK1, MAPKBP1, NF1, and PDGFR (227, 228).

lncRNAs and microRNAs are potential therapeutic targets since they control the ERK/MAPK pathway (229). MALAT1 (metastasis-associated lung adenocarcinoma transcript 1) is a lncRNA that increases gallbladder cancer proliferation and metastasis by activating the ERK/MAPK pathway (230). MALAT1 has also been proven to be an excellent predictor of gallbladder cancer recurrence and prognosis (231-233). The miR-101 microRNA is a direct target of ZFX (zinc finger protein X-linked), which decreases TGF-mediated EMT in gallbladder cancer by inhibiting MAPK/ERK/SMAD signalling (234). The MAPK pathway was inhibited by microRNA-29c-5p, a direct target of CPEB4 (Cytoplasmic polyadenylation element binding protein 4), to diminish the tumorigenic properties of cells (235).

2.5.3 EGFR pathway

The EGFR (ERBB1/HER1), also known as the ErbB receptor, is a transmembrane glycoprotein that is a member of the receptor tyrosine kinase family (RTK). In addition to ERBB1, this family also includes ERBB2 (HER2), ERBB3 (HER3), and ERBB4 (HER4) (236). ERBB1 is the most often mutant and aberrantly expressed RTK in a variety of malignancies. The

activation of receptor tyrosine kinases induces ERK/MAPK and PI3K/AKT signalling cascade, which are important for proliferation and invasion of cancer cells (237, 238). It is believed that frequent mutation and abnormal expression of ERBB1 serve as both a useful prognostic indicator and a therapeutic target for the treatment of cancer (239, 240).

Gallbladder cancer ERBB1 mutation rates have been estimated to be between 3.9%-4% (241, 242), whereas over-expression of EGFR has been found in between 44%-74% of gallbladder cancer tissues (242-244). Recently, Shen et al. found that PLEK2 (Pleckstrin 2), an oncogene known to interact with the kinase domain of EGFR, triggers EGFR signalling, which then encourages gallbladder cancer invasion and metastasis (245). Targeting ERBB1 in addition to conventional anticancer therapy has been a particular focus of recent clinical trials for gallbladder cancer. By using a dual inhibition strategy that included an anti-EGFR antibody and a tyrosine kinase inhibitor, a greater response in the reduction of EGFR signalling was observed (246, 247).

2.6 Available and potential therapies for gallbladder cancer

Available therapies

In early-stage GBC, the recommended treatment is cholecystectomy, which involves the surgical removal of the gallbladder and surrounding tissues. In some cases, nearby lymph nodes may also be removed (248). Laparoscopy is sometimes used as a guiding tool during gallbladder surgery (249). Unfortunately, surgery alone is not a viable option for curing advanced-stage cancer. However, palliative surgery may be recommended to alleviate symptoms caused by a tumour blocking the biliary system. In cases of locally advanced tumors, postoperative chemotherapy is a reasonable approach to prevent further spread before surgery (250).

A study involving 140 patients compared the outcomes of surgery alone versus surgery followed by postoperative chemotherapy. The chemotherapy regimen included two courses of

mitomycin C and infusional 5-FU, followed by prolonged oral administration of 5-FU until tumour progression. It was found that 90 percent of patients with positive lymph nodes benefited from adjuvant chemotherapy (251). Currently, the most commonly used chemotherapeutic agents for GBC are 5-FU or Gemcitabine, either as monotherapy or in combination with a platinum drug (cisplatin, carboplatin, or oxaliplatin). The combination therapies have shown better efficacy, particularly in advanced stages of the disease (99, 252). However, the clinical application of platinum drugs is limited by their side effects, lack of selectivity, high systemic toxicity, and the development of drug resistance (253).

Potential therapies

Gallbladder cancer has a low overall survival rate due to its poor prognosis and late clinical intervention, driving the scientific community to investigate new therapeutic targets (254). While there are a few treatments available such as cholecystectomy, radical gallbladder resection, palliative surgery, radiotherapy, and chemotherapy, the risk of recurrence and remission remains high (255). Additionally, these treatments often cause significant side effects and drug resistance. Therefore, alternative treatment approaches are required.

Recent studies has revealed that high expression levels of long noncoding RNAs (lncRNA), including AFAP1-AS1, MALAT1, and ROR, are associated with decreased overall survival of patients (256). Conversely, certain microRNAs (miR-20a, miR-182, and miR-155) have been found to be overexpressed in gallbladder cancer and associated with increased tumorigenic properties, making them promising targets for treatment (257).

Clinical study at Johns Hopkins has shown that adjuvant chemotherapy can improve the overall survival rate of patients with tumour recurrence, with Gemcitabine with Cisplatin, and additional chemotherapy regimens being the most effective treatments (258). However, drug resistance and harmful side effects limit the efficacy of traditional therapeutic agents, such as

cisplatin and gemcitabine, and specialized inhibitors and combination therapies are needed (258, 259). To increase the efficacy of existing chemotherapeutic drugs, researchers are exploring small molecule inhibitors of critical signalling molecules and pathways involved in gallbladder cancer.

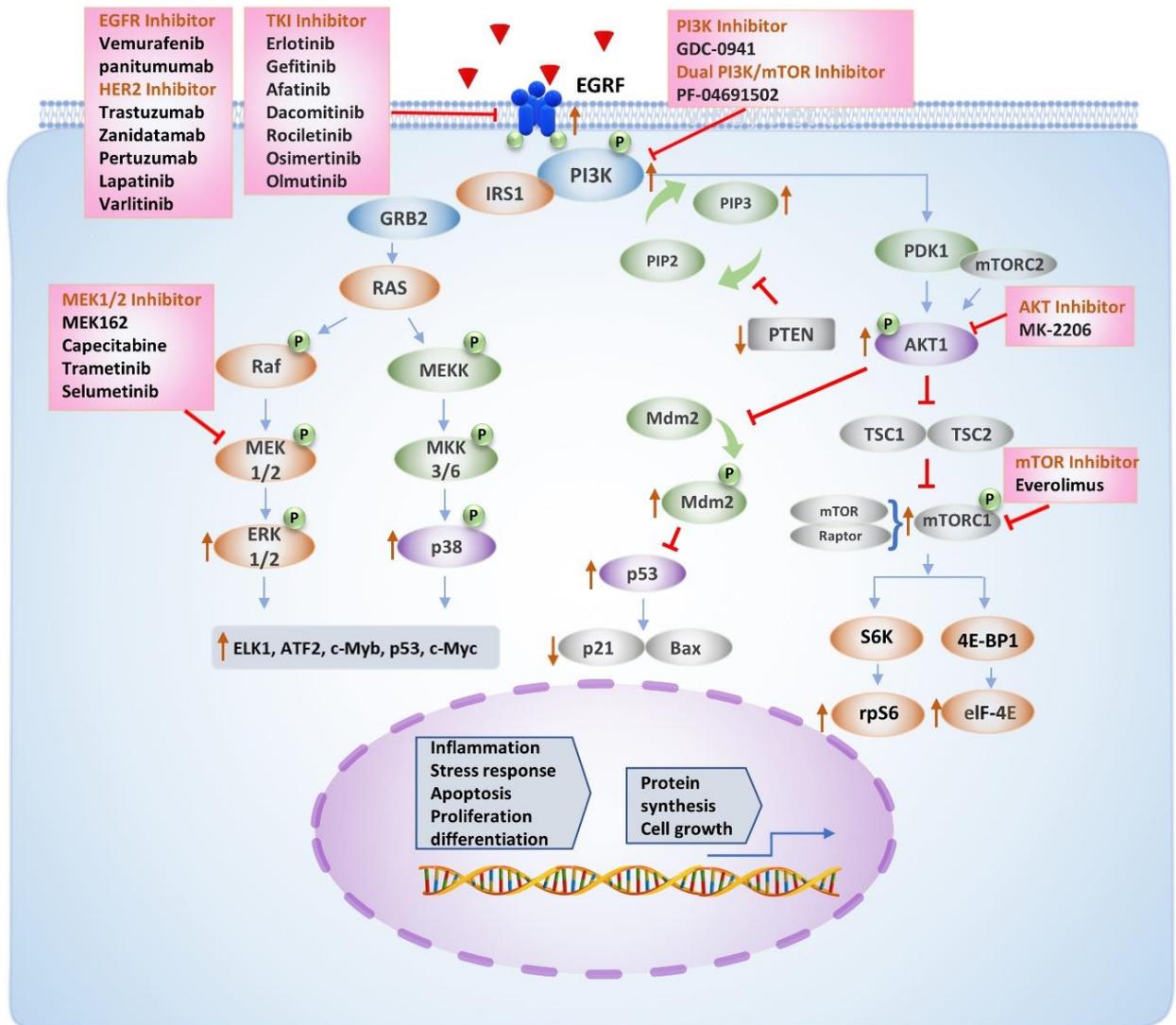


Figure 2.6.1 The Major pathways associated with gallbladder cancer. *The summary of the molecular inhibitors and therapeutic targets in the major pathways associated with gallbladder cancer. The names of the small molecule inhibitors are included in a box with a rectangle to indicate their specific signalling targets. The increased and decreased level of expression are represented by the bold red upward and descending arrows. Target molecule inhibition is indicated by bold, red T-shaped arrows. The figure is adapted from reference (260)*

2.7 Matrix metalloproteinases

MMPs belong to the matrixins protease superfamily of metzincins and are a family of zinc-containing endoproteases. They were first identified for their ability to degrade ECM components, particularly collagen, during tadpole tail resorption in the early 1960s (261). Based on their substrate specificities, MMPs can be classified into various categories, including collagenases, gelatinases, stromelysins, membrane type, and miscellaneous.

A numbering system is widely used for easy identification, as the same MMP can have multiple substrates (262). In humans, 23 MMPs have been identified. The structure of MMPs consists of a pro-peptide domain with a cysteine-switch motif, a catalytic metalloproteinase domain, a linker peptide, and a hemopexin domain (263), these domains structures are shown in

Figure 2.7.1.

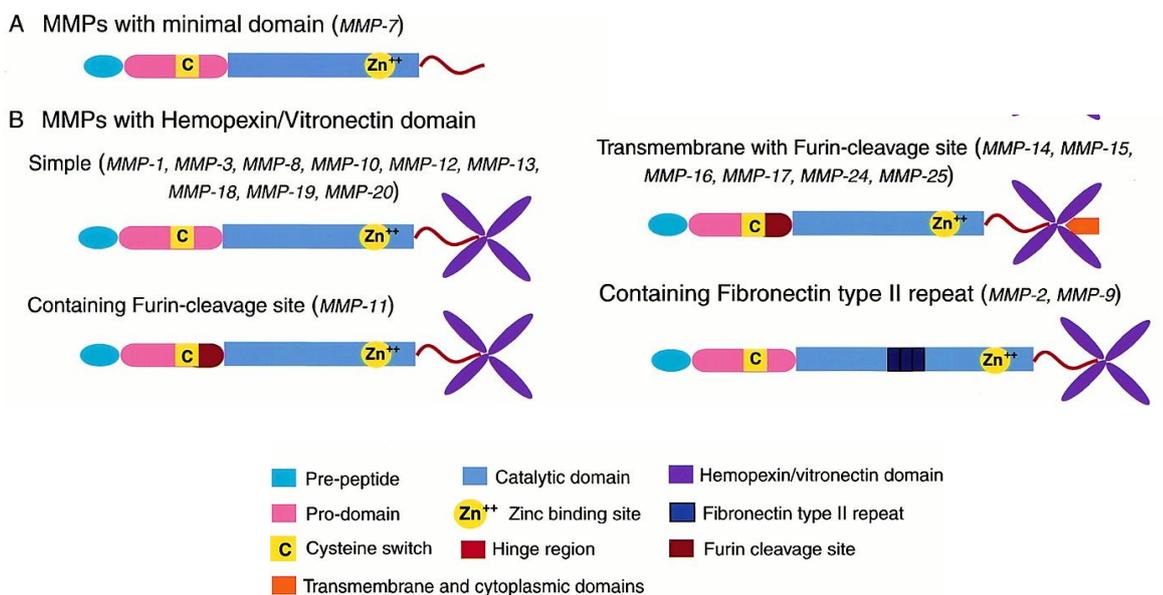


Figure 2.7.1 The Matrix metalloproteinases domain structure. *Pictorial representation of the domain organization of the various matrix metalloproteinase subgroups. A highly conserved pro-domain region has the cysteine switch. Some MMPs have a furin cleavage motif in their pro-domain that attracts furin-like enzymes. Three gelatin-binding fibronectin type II repeats can be found in the catalytic domain of the gelatinases MMP2 and MMP9. The folding of the hemopexin/vitronectin domain forms a four-bladed propeller structure. Adapted from reference (264).*

MMPs are initially expressed as inactive pro-MMP zymogens due to the interaction between the prodomain cysteine-switch motif PRCGXPD and catalytic domain zinc ions, which blocks water molecules from entering for active catalysis (265). The activation of MMPs is mediated by proteolytic removal of the prodomain interaction or chemical alteration of cysteine residues (262, 266).

2.7.1 Matrix metalloproteinase's role in tumorigenesis

MMPs in tumour initiation: Studies have shown that stromal cues increase the risk of cancer development (267), particularly in fibrotic and inflammatory settings, and some familial cancer syndromes are caused by genetic flaws that result in stromal changes prior to epithelial changes (268). In such cases, MMPs (MMP-1, -2, -7, -9 and -14) are elevated in the stroma, and this increase in MMP activity can contribute to cancer risk (269, 270). Studies using genetically altered mice suggest that MMPs play a role in early cancer formation. For instance, MMP (MMP-2, -7 and -14) overexpressing transgenic mice tend to develop spontaneous hyperproliferative lesions, and MMP overexpression generates more tumors than usual in response to various oncogenic stimuli. Conversely, mice lacking certain MMPs (33, 35, 271) or overproducing TIMP1 are less likely to develop cancer (272).

MMPs in tumour progression: Overexpression of MMP7 has been shown to promote the development of breast cancer in mice with a mammary-targeted Her2-transgene (28), while its absence in animals with the Apc^{min} mutant has been shown to slow the development of intestinal

adenomas (271). When MMP9-expressing bone marrow cells are transplanted into MMP9-deficient transgenic mice, normal carcinogenesis is restored, demonstrating the significance of inflammatory cell-derived MMP9 in skin carcinogenesis. In transgenic mice, the absence of MMP9 inhibits the development but increases the progression of human papilloma virus-16-induced squamous cell carcinomas (273). Similarly, mice with a mammary-targeting MMP14 mutation also get spontaneous mammary tumors (35). Overall, these studies suggest that MMPs function as natural tumourpromoters and can alter cancer susceptibility. The genetically altered mice and MMPs involved in tumorigenesis are listed in **Table 2.7.1**.

Table 2.7.1 MMPs induce carcinogenesis in mice with genetic modifications

<i>MMP Genotype</i>	<i>Oncogenic stimulus</i>	<i>Phenotype</i>	<i>Reference</i>
<i>MMP2</i> ^{-/-}	RIP-TAg	↓ pancreatic tumourgrowth	(33)
	Injected cells	↓ angiogenesis and tumourgrowth	(13)
<i>MMP7</i> ^{-/-}	<i>Apc</i> ^{min}	↓ intestinal adenoma formation	(271)
<i>MMTV-MMP7</i>	none	↑ mammary hyperplasias	(28)
	<i>MMTV-neu</i>	↑ mammary carcinogenesis	(28)
<i>MMTV-MMP14</i>	none	↑ mammary hyperplasias and cancers	(35)
<i>MMP9</i> ^{-/-}	K14-HPV16	↓ skin carcinogenesis	(273)
	RIP-TAg	↓ pancreatic carcinogenesis	(33)

Abbreviations: *Apc*^{min}, adenomatous polyposis coli gene with multiple intestinal neoplasia mutation; MMTV, mouse mammary tumourvirus promoter; HPV16, human papilloma virus 16 early region; K14, keratin-14 promoter; RIP, rat insulin promoter; TAg, SV40 T antigen.

2.7.2 Matrix metalloproteinase mode of action

Extracellular proteinases play a crucial role in several physiological and pathological processes (274). MMPs, a group of related enzymes, since their initial discovery, are highly effective in breaking down structural proteins in the ECM. Additionally, they can degrade various pericellular non-matrix proteins that affect how cells behave in multiple ways (269). Consequently, MMPs have an impact on various physiological and pathological processes, including embryonic development, tissue morphogenesis, wound healing, inflammatory disorders, and cancer (269, 275). These modes of action are illustrated in **Figure 2.7.2** and are further discussed in the following sections.

Cell proliferation: ECM molecules and cell surface receptors interact to control how cells behave. The capacity of cells to multiply, survive, or develop also varies in a changing environment. Since MMPs affect the ECM, it is not surprising that they can change cell behaviour. The few studies that have been done on the effects of MMPs on cell behaviour in vitro have indicated that they have an impact on cell organisation, differentiation, proliferation, survival, and apoptosis (276). For instance, inhibition of MMP2 reduces the mitogenic response of cultured vascular smooth muscle cells to PDGF (277). Furthermore, the inhibition of MMPs (MMP1, MMP2, and MMP3) in cultured human dermal microvascular endothelial cells promotes ECM deposition and restricts cell proliferation (278).

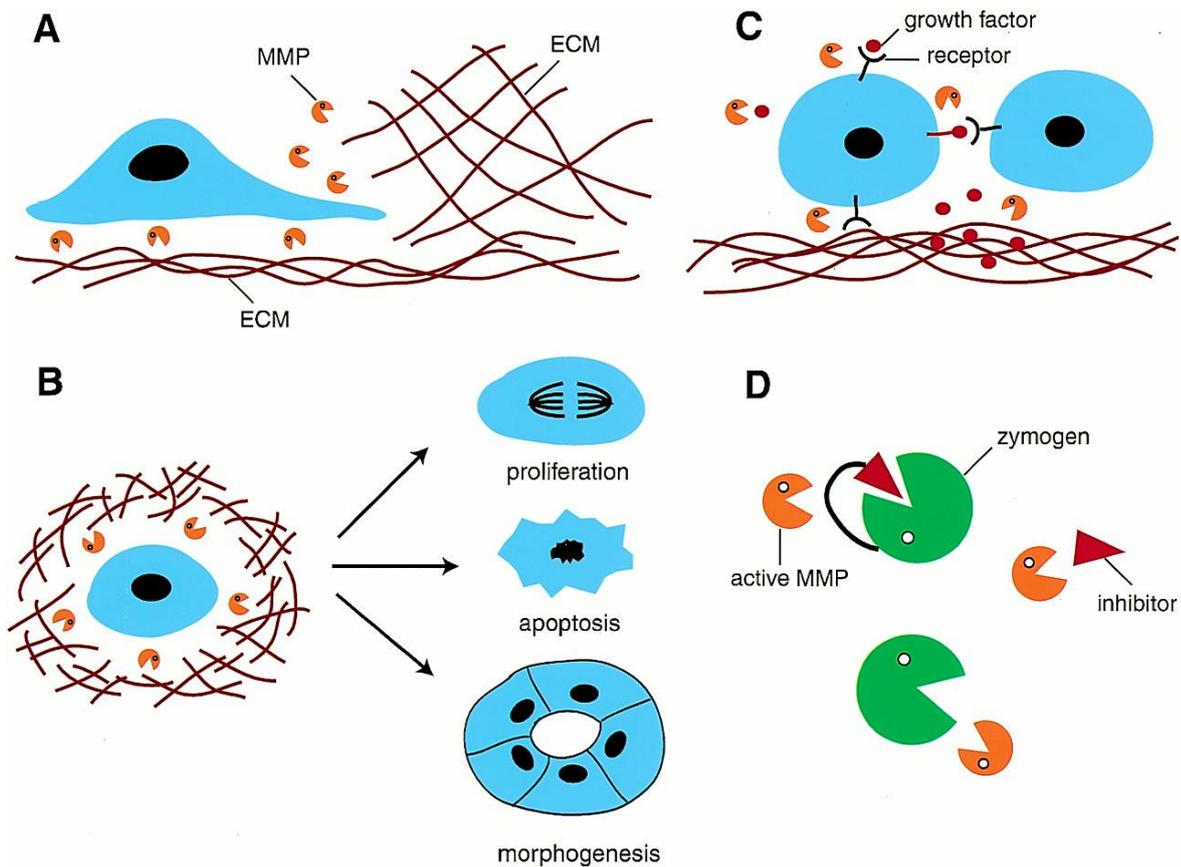


Figure 2.7.2 Matrix metalloproteinase modes of action. (A) MMPs may alter cell migration by altering the behavior of the cells from one of adhesion to one of non-adhesion type and by disrupting the ECM. (B) MMPs may change the microenvironment of the ECM, causing changes in cell growth, death, or morphogenesis. (C) By cleaving or releasing physiologically active molecules from the ECM, such as growth factors and their receptors, MMPs can alter their biological activity. (D) By cleaving the enzymes or their inhibitors, MMPs can change the equilibrium of protease activity. The figure is adapted from reference (264).

Apoptosis: MMPs have also been found to cause cell death. For example, mammary epithelial cells undergo apoptosis due to proteinases or aberrant ECM molecules, which could be due to impaired integrin signalling (279). Inhibition of MMPs (MMP2 and MMP9) activity prevents cell death and prevents the induction of ECM-degrading proteinases in mammary cells in response to a fibronectin fragment (FN120) (280). MMP2 inhibitors also reduce smooth muscle cell growth and trigger cell death in hypertrophied rat pulmonary arteries in organ culture (281). Therefore, MMPs regulate survival signals produced by specific ECM changes, affecting both positive and negative aspects of cell survival and proliferation.

Tissue morphogenesis: MMPs' ability to modify cell organization is one of their distinct actions that is relevant to tissue morphogenesis. For instance, inhibition of MMP activity reduces cell migration and three-dimensional structure formation of adipocytes when grown on a basement membrane (282). Additionally, embryonic pancreatic epithelial cells with increased MMP2 activity differentiate and group into clusters resembling islets of Langerhans in an in vitro model of pancreatic islet development, while MMP2 suppression eliminates islet morphogenesis without affecting the differentiation of endocrine cells (283).

Angiogenesis: The tubular branching, and the formation of several epithelial structures involve ECM-dependent cellular architecture, which requires precise interactions between cells and ECM molecules (276). Endothelial and epithelial cells in three-dimensional culture exhibit varying dependence on MMPs when forming tubular structures. By reducing MMP activity, HUVECs cultivated on Matrigel or collagen gels are unable to produce tubules (284, 285).

Cell migration: The cells must transition from an adhesive to a migratory phenotype before moving, and the ECM must be degraded to eliminate any physical barriers (10). MMPs have been shown to break down a wide range of ECM substrates, including collagens and non-collagenous compounds, in numerous in vitro experiments, indicating that they may act as ECM-clearing enzymes during cell migration (269) and developmental process (286).

Regulation of biologically active molecules: The activity of MMPs can affect the availability and function of biologically active molecules by either releasing them from bound proteins or ECM storage, or proteolytically activating or inactivating them. This can cause either a gain or loss of function in these molecules (10). For instance, MMP-1 and MMP3 have been shown to break down perlecan and release bound FGF (287) , while MMP2, MMP3, and MMP7 can cleave decorin, freeing TGF that was bound to decorin (288). MMPs may also regulate the bioavailability of growth factors by cleaving their binding proteins (289). For example, MMP1,

MMP2, and MMP3 can cleave IGFBP-3, and MMP1 and MMP2 can cleave IGFBP-5, thereby modulating the availability of these binding proteins and influencing IGF activity (290).

2.7.3 Genetic association studies of MMP2, 7 and 14 in cancers

MMP14: Genetic association studies have linked MMP14 variations to several diseases, including hepatocellular carcinoma (HCC) (38), ovarian cancer (39), oral cancer (291), Winchester syndrome (WS) (292), Osteonecrosis of the femoral head (ONFH) (293) and intervertebral disc degeneration (IVDD) (294). MMP14 promoter variants rs1003349 and rs1004030, have been shown to regulate MMP14 expression by facilitating the binding of Sp1 and RR1(295). Previous studies in different populations have identified certain risk alleles of these variants to be associated with other diseases such as osteoporosis (296), focal and segmental glomerulosclerosis, and mesangiocapillary glomerulonephritis (295). In the Taiwanese population, the allele "T" of rs1004030 was found to be a risk allele for both HCC and uterine cervix neoplasia (297).

MMP7: In Chinese women with breast cancer, MMP7 polymorphisms rs11568818 and rs11225297 with rare alleles 'G' and 'A', respectively, are associated with poor survival (298). Studies have shown that various genetic variations in MMP7 increase the risk of developing different types of cancers such as squamous cell carcinoma, gastric cardiac adenocarcinoma (299, 300), and non-small cell lung carcinoma (301). Two functional promoter variations, A181G (rs11568818) and C153T (rs11568819), have been reported to interact with numerous nuclear proteins and exhibit allele-specific regulation of gene expression (302-305). Studies have found that patients with gastric cancer (299) and hypertension (304) have allele-specific binding of Cyclic AMP Response Element-Binding Protein (CREB) at the promoter variant rs11568818, leading to increased MMP7 expression. However, the impact of these MMP7 promoter variations on GBC has not been extensively studied, with only one study by Sharma

et al. indicating that rs11568818 may be a risk factor for GBC in the north Indian population (52).

MMP2: MMP2, like most MMPs, is polymorphic and modulates enzymatic activity and gene expression (306). Studies have identified two promoter variants of MMP2, rs243865 C>T (-1306) and rs242866 G>A (-1575), as risk factors for metabolic syndrome (307), end-stage kidney disease (308), coronary triple-vessel disease (309). Additionally, variant rs243865 is also considered risk factors for intracranial aneurysms (310) and Alzheimer's disease (311). Furthermore, rs243865 C>T (-1306) and rs2285053 C>T (-735) with alleles "C" and "C," respectively, increase the risk of lung (312), nasopharyngeal (313) and esophageal carcinoma (314, 315). Interestingly, the functional variants rs243865 C>T (-1306) and rs2285053 C>T (-735) show strong linkage disequilibrium (LD), which influences the binding of the Sp1 binding site in an allele-dependent manner. The allele C>T transition destroys the Sp1 binding element at both loci, resulting in a drastic decrease in the activity of the MMP2 promoter (314, 316). Moreover, the presence of haplotype T-T (rs243865 and rs2285053, respectively) leads to a 3.7-fold decrease in MMP2 transcription compared to the C-C haplotype, indicating allele-dependent control of MMP2 expression (316). Despite several studies highlighting the significance of these matrix metalloproteinases and their genetic correlation with various cancers and disorders (**Table 2.7.2**), none provide an in-depth mechanistic analysis of these genetic variants in the Indian population.

Genetic association studies of MMP2, 7 and 14 with GBC

An isolated study by Sharma *et al.* from Indian population showed association of SNPs present in MMP2 (rs2285053, rs243865), MMP7 (rs11568818), MMP9 (rs17577) and TIMP2 (rs8179090) with gallbladder cancer (52). However, the underlying mechanism of these variants in GBC has yet to be explored.

Table 2.7.2 Genetic association studies of MMPs in various cancer

SNP of MMP	Disease	OD	CI-95%	P value	PMID	Population
rs1799750 (MMP-1)	bladder cancer	0.62	0.39-0.98	0.042	PMID: 23819551	Polish
rs11644561 (MMP-2)	breast cancer	0.6	0.3-1	0.098	PMID: 19454611	Chinese
rs243865 (MMP-2)	prostate cancer	1.32	1.05– 1.66	0.016	PMID: 29069837	Chinese
rs243865 (MMP-2)	lung cancer	0.63	0.45– 0.89	0.009	PMID: 28827732	Chinese
rs243865 (MMP-2)	oral cancer	0.73	0.58– 0.91	0.006	PMID: 28595731	Indian
rs243865 (MMP-2)	cervical cancer	1.46	1.18-3.55	0.032	PMID: 26526578	Chinese Han
rs243865 (MMP-2)	breast cancer	1.27	1.10-1.47	0.001	PMID: 21161369	Chinese
rs243865 (MMP-2)	GBC	1.59	1.04-2.42	0.03	PMID: 22621753	North India
rs2285053 (MMP-2)	gastric adenocarcinoma	2.87	1.42-5.83	0.003	PMID: 29285307	Chinese
rs2285053 (MMP-2)	GBC	1.81	1.1-2.8	0.008	PMID: 22621753	North India
rs11643630 (MMP-2)	breast cancer	0.8	0.7-1	0.046	PMID: 19454611	Chinese
rs3025058 (MMP-3)	gastric cancer	1.75	1.21– 2.53	0.0027	PMID: 22121090	Indian
rs522616 (MMP-3)	gastric cancer	2.19	1.58– 3.03	0.0001	PMID: 22121090	Indian
rs617819 (MMP-3)	gastric cancer	1.08	0.79– 1.48	0.634	PMID: 22121090	Indian
rs679620 (MMP-3)	gastric adenocarcinoma	2.15	1.04-4.45	0.04	PMID: 29285307	Chinese
rs7935378 (MMP-7)	breast cancer	1.3	0.9-1.8	0.048	PMID: 18648013	Chinese
rs10895304 (MMP-7)	prostate cancer	3.37	1.56-7.26	0.001	PMID: 20605361	USA
rs11568818 (MMP-7)	endometrial cancer	2.03	1.21-3.39	0.017	PMID: 20113256	Taiwan
rs11568818 (MMP-7)	GBC	1.3	0.9-1.81	0.051	PMID: 22621753	North Indian
rs17098236 (MMP-7)	Ovarian Cancer	1.19	1.01– 1.42	0.04	PMID: 20628624	Australia
rs2292730 (MMP-20)	ovarian cancer	2.03	1.39– 2.96	0.0002	PMID: 25867973	Caucasian
rs2236307 (MMP-14)	cervical neoplasia	2.13	1.03-4.37	0.037	PMID: 22527983	Taiwan
rs1003349 (MMP-14)	COPD	1.13	0.88-1.45	0.3456	PMID: 23267696	Chinese Han



CHAPTER 3

Materials and Methods



Vinay J
NISER, Bhubaneswar

3.1 Materials

3.1.1 Study participants

The blood samples of study subjects were collected from SCB Medical College & Hospital and tissue samples were collected from Acharya Harihar Regional Cancer Centre. The study was approved by Institutional Ethics Committee, NISER, Bhubaneswar (protocol number-NISER/IEC/2017/02). Also, the study was approved by ethical committee of collaborative hospitals, Acharya Harihar Regional Cancer Centre with protocol number (062/IEC/AHRCC), and SCB Medical College & Hospital with protocol number (123/22.11.2014). The prospective study included subjects of all age groups and gender between 2014 and 2020. Informed written consent was taken from all recruited study subjects. The study adheres to the principles of the Helsinki Declaration.

The inclusion and exclusion criteria for GBC involve all patients with diagnostically confirmed GBC and natives of Odisha state. We excluded patients with incomplete or missing consent, diagnosed with some other cancers, and non-natives of Odisha state. Similarly, we included control samples from a random Odisha population with no history of chronic diseases or cancer. We excluded all control subjects who were not residents of Odisha and had a history of chronic illness. Demographics, and clinical and histopathological features of the study subjects are shown in

Table 3.1.2 and **Table 3.1.1**, respectively.

Table 3.1.1 Demographic profile of the gallbladder cancer patients and controls recruited in the study

Study subjects with blood samples			
Characteristics	Cases (n = 314)	Control (n = 323)	P value
Age^a	53.39 ± 11.12	42.94 ± 11.12	0.0001
Gender^b			
Female	181	127	0.0001
Male	133	196	
Gallbladder cancer patients with tissue samples			
Age		58 ± 8.46	58.3 ± 9.2
Gender		Male	Female
		6	20
Gallstone		Present	Absent
		9	17

^a Student's T-test was used to compare mean values of age

^b Chi-Square test was used to compare the difference in frequency of male and female

Table 3.1.2 Gallbladder cancer patients clinical and histopathological characteristics

TumourHistological type	Number of GBC patients
Papillary Adenocarcinoma	1
Signet ring cell type Adenocarcinoma	1
Invasive Adenocarcinoma	24
Histological grade	
G1	12
G2	10
G3	4
Local invasion	
Adjacent liver	2
Lymphovascular invasion	3
Perineural invasion	13
Periductal invasion	2
Cystic duct invasion	8
Periportal lymph node	4

3.1.2 Cell lines

Gallbladder cancer cell lines TGBC1TKB and G415 was procured from the Riken BRC Cell Bank repository (Ibaraki, Japan). Human embryonic kidney cells (HEK293T) were procured from NCCS (Pune, India). The cellular origin and source of procurement are listed in the following **Table 3.1.3**.

Table 3.1.3 List of cell lines used in the study and their cellular origin

Name of cell lines	Details	Source
HEK293T	Tissue: Embryonic kidney Morphology: Epithelial Culture type: Adherent cells	NCCS, Pune
TGBC1TKB	Tissue: Gallbladder Morphology: Epithelial Culture type: Adherent cells	Riken, Japan
G415	Tissue: Gallbladder Morphology: Epithelial Culture type: Adherent cells	Riken, Japan

3.1.3 Vector constructs, shRNAs, antibodies and primer sequences

A list of all vector constructs, sh-RNAs, antibodies and primers used in this study are given in Appendix I, II, III and IV, respectively.

3.2 Chemicals and reagents

DNA extraction and PCR clean-up or Gel purification kit: HiPurA Blood DNA Isolation Kit (HiMedia), Pinpoint Slide DNA Isolation System (Zymo Research), and NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel).

DNA extraction by manual method: RBC Lysis buffer (HiMedia) (0.32M Sucrose, 1mM MgCl₂, 1% Triton X-100, 12mM Tris-HCl, pH adjusted to 7.6), sodium dodecyl sulfate (SDS), Proteinase K buffer (20 mM Tris-HCl, 4 mM Na₂EDTA, 100 mM NaCl, pH 7.4), Proteinase

K (HiMedia), NaCl (MP Biomedicals), Ethanol (MP Biomedicals), Nuclease Free Water (GeNei), 10X TE (GeNei) and Phenol-Chloroform-Isoamyl alcohol mixture (Tris saturated Phenol: Chloroform: Isoamyl Alcohol 25:24:1) (HiMedia).

Polymerase Chain reaction:

Taq polymerase Buffer (GeNei), 1.5 mM MgCl₂, 0.5 μM of each primer (IDT), 100mM dNTP mix (GeNei), 100ng of genomic DNA, 0.5 unit of Taq DNA polymerase (GeNei) and 2.5μl of DMSO (MP Biomedicals) and Nuclease Free Water (GeNei).

Agarose gel electrophoresis: LE Agarose (Lonza), 50X TAE (50 mM EDTA sodium salt, 2M Tris, Glacial acetic acid 1M), 1Kb DNA ladder (NEB), 50 bp DNA ladder (NEB), 6X DNA loading dye (GeNei).

Sangers sequencing:

BigDye Terminator v.3.1 cycle sequencing kit (Applied Biosystems), Sequencing buffer (5X), RR100 mix (1X), Primer (2 pmol), DNA template (50-100ng), HiDi formamide and Nuclease Free Water (GeNei).

Sequencing Clean-up:

125mM EDTA, 3M NaOAc pH-4.6, Nuclease Free Water (Genei) and Ethanol (MP Biomedicals).

Plasmid isolation: Plasmid mini kit (Qiagen), Plasmid midi kit (Qiagen), Isopropanol (MP Biomedicals), Nuclease Free Water (GeNei), 10X TE (Tris-EDTA 0.5M, pH 7) (GeNei).

Western blotting:

Lysate preparation: Halt™ Protease and Phosphatase Inhibitor Single-Use Cocktail (100X), RIPA Lysis and Extraction Buffer (Thermoscientific), Pierce Protease Inhibitor Mini Tablets (Thermoscientific), Glycerol (MP Biomedicals) and β-Mercaptoethanol (MP Biomedicals).

SDS-PAGE and Buffers: Tris (MP Biomedicals), Glycine (MP Biomedicals), Acrylamide (Invitrogen), Bis-acrylamide (Sigma), APS (MP Biomedicals), SDS (MP Biomedicals), TEMED (MP Biomedicals), and Tween 20 (MP Biomedicals).

Quantification and Detection: Pierce™ BCA Protein Assay Kit (Thermoscientific), Polyvinylidenedifluoride membrane (PVDF) (Millipore), Skim milk powder (MP Biomedicals), Bovine Serum Albumin (BSA) (MP Biomedicals), 4 color prestained protein ladder (Puregene), Page-ruler prestained protein ladder (Thermoscientific), Bromophenol blue (Sigma), Ponceau-S stain (HiMedia), Restore plus stripping buffer (Thermo Fischer Scientific), and Super-Signal west Femto-substrate (Thermoscientific).

Bacterial culture: LB Broth- Miller (HiMedia), LB Agar- Miller (Lonza), Ampicillin (100 mg/ml) (Sigma) in autoclaved milli-Q water.

Competent Cell preparation:

Inoue Buffer: 250 mM KCl (MP Biomedicals), 55mM MnCl₂.4H₂O (MP Biomedicals), 15 mM CaCl₂.2 H₂O (MP Biomedicals), 10 mM PIPES in autoclaved milli-Q water.

Cloning: KpnI (New England Biolabs), XhoI (New England Biolabs), AgeI (New England Biolabs), EcoRI (New England Biolabs), BbsI (New England Biolabs), NcoI-HF (New England Biolabs), T4DNA Ligase (New England Biolabs), and Quick ligase (New England Biolabs).

Immunohistochemistry (IHC):

10% Buffered formalin (10% formalin, 0.025M NaH₂PO₄, 0.046 M Na₂HPO₄ in Milli Q water), Formalin solution (HiMedia), Paraffin (Thermoscientific), Ethanol (MP Biomedicals), Acetone (Merck), Poly-L-Lysine (Merck, USA), Xylene (Merck, USA), Haematoxylin (Himedia), Eosin (Himedia), Tris-buffered saline (0.05M Tris, 0.8% NaCl), Tris-EDTA buffer pH-9 (10 mM

Tris, 1 mM EDTA), EnVision Flex Wash Buffer (Dako, Denmark), EnVision Flex Substrate Buffer (Dako), EnVision Flex DAB+ Chromogen (Dako), EnVision™ Flex HRP secondary antibody (Dako), EnVision Flex peroxidase blocking reagent (Dako), DPX mountant (Fisher Scientific), Sodium Citrate target antigen retrieval buffer (10mM Sodium Citrate, 0.05% Tween 20, pH 6.0), and EnVision™ Flex target Retrieval solution, high pH (Dako).

Cell culture: DMEM (Hi-media), RPMI1640 (HiMedia), DPBS pH 7.4 (HiMedia), Trypsin-EDTA solution 0.25% (HiMedia), PSA antibiotic cocktail [Penicillin (10000 IU/mL), Streptomycin (10 mg/mL), and Amphotericin B (25 µg/mL)] (MP Biomedicals), Trypan blue (0.4% in PBS) (HiMedia), Monensin sodium salt (Sigma), Fetal Bovine Serum (HiMedia), DMSO (MP Biomedicals, India), Ethanol (Merck).

Transfection and stable cell-line generation: Opti-MEM (Thermoscientific), Lipofectamine 3000 (Thermoscientific), Puromycin (1mg/ml) (Sigma).

Cell proliferation assay: CellTiter 96 Aqueous One solution reagent (Promega).

Luciferase assay: Dual-Luciferase® Reporter Assay System (Promega), Nunc Micro Well 96-Well, Nunclon Delta-Treated, Flat-Bottom Microplate (Thermoscientific).

Chromatin immunoprecipitation (ChIP) assays: ChIP Kit (ab500) (abcam).

Electrophoretic mobility shift assay (EMSA): LightShift Chemiluminescent EMSA Kit (Thermoscientific), Chemiluminescent Nucleic Acid Detection Module Kit (Thermoscientific), NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermoscientific), and Biodyne B Nylon Membrane, 0.45 µm, 8 cm x 12 cm (Thermoscientific).

Transwell migration and invasion assay: Phosphate Buffered Saline (pH- 7.4), Matrigel Growth factor reduced (Corning), Millicell Hanging Cell Culture Insert, 8 µm pore-size, 12-

well (Millipore), Paraformaldehyde (HiMedia), Methanol (Merck), Crystal violet (HiMedia), and Giemsa (Fisher scientific).

3.3 Methodology

3.3.1 DNA extraction

To obtain genomic DNA from both case and control subjects, 4 mL of peripheral blood was drawn into vials coated with EDTA, a chelating agent that prevents blood from clotting. The phenol-chloroform extraction method was employed for DNA isolation. Initially, 0.5 mL of blood was centrifuged at 11,000 rotations per minute (rpm) for 10 minutes at 25°C, and the resulting supernatant was discarded. Next, 1 mL of a red blood cell (RBC) lysis buffer was added to the cell pellet, which was then broken down using rapid pipetting. The sample was again centrifuged at 11,000 rpm for 5 minutes, and any remaining RBCs were removed by repeating the previous step with 200µl of RBC lysis buffer. The resulting pellet was then washed with 200µl of Milli-Q water using centrifugation. To proceed with DNA extraction, 80 µl of proteinase K buffer and 10 µl of 10% sodium dodecyl sulphate (SDS) were added and mixed until foaming was observed. Subsequently, 100µl of chilled sodium chloride (NaCl) was added and mixed by tapping, followed by adding another 200µl of MQ and mixing. Then, 400µl of phenol: chloroform (Tris-saturated Phenol: Chloroform: Isoamyl Alcohol 25:24:1) was added and vigorously mixed by inverting the microcentrifuge tubes. The sample was then centrifuged at 12,000 rpm for 10 minutes, and the resulting upper transparent aqueous phase was collected. To precipitate the DNA, 1 ml of chilled absolute ethanol was added and mixed, followed by centrifugation at 13,000 rpm for 10 minutes. The pellet was washed with 200µl of 70% ethanol using centrifugation at 13,000 rpm for 5 minutes. The supernatant was discarded, and the pellet was air-dried overnight at room temperature. Finally, the dried DNA pellet was

dissolved in 30µl of Tris-EDTA buffer and quantitated using microvolume UV spectrophotometer NanoDrop-One (Thermoscientific).

4.1.1 Polymerase chain reaction

To genotype the target region containing single nucleotide polymorphisms (SNPs), primers sets were designed using the Primer-BLAST tool and used to amplify the region by Polymerase Chain Reaction (PCR). The primer sequences and annealing temperatures are provided in **Table 3.3.1**. The PCR reactions were carried out in the 25µl of reaction mix containing Taq Buffer (GeNei, 1.5 mM MgCl₂, 0.5 µM of each primer (from IDT, USA), 100 mM dNTP mixture (GeNei), 0.5 unit of Taq DNA polymerase (GeNei), and 100 ng of genomic DNA (

Table 3.3.2). The reaction mixture was initially incubated at 94°C for 5 minutes, followed by 30 cycles of denaturation at 95°C for 30 seconds, annealing at specific temperatures for 30 seconds, and extension at 72°C for 45 seconds, with a final extension at 72°C for 4 minutes (

Table 3.3.3). The resulting amplicons were stored at -20°C until further use. The amplified DNA was sequenced to detect the presence of SNPs. The genotyping data obtained from sequencing the PCR products were used to identify and compare the genetic variations between the case and control groups.

Table 3.3.1 List of primers used in the study

SI No	Oligo Name	Primer Sequence 5' to 3'	Annealing temperature
1	MMP2_F_SET1	GGAGTTCCCCATCACAGCTTA	54°C
2	MMP2_R_SET1	GCCTCGTATAGTGCGAGATG	
3	MMP2_F_SET2	CCCAAGCCGCAGAGACTTTT	54°C
4	MMP2_R_SET2	GCCTGACTTCAGCCCCTAAAC	
5	MMP7_F_SET1	CCTGAATGATACCTATGAGAGCAGT	54°C
6	MMP7_R_SET1	CATAGCTGCCGTCCAGAGAC	
7	MMP7_F_SET2	CACCCAATTTGTGGCTTGTGTG	53°C
8	MMP7_R_SET2	CATGGTAATTGAGCACTGTGAGC	
9	MMP14_F_SET1	CAGAGGAATCAAGCCACTCAGA	54°C

10	MMP14_F_SET1	TCCTCTCCGAATAGAGGCTGT	53°C
11	MMP14_F_SET2	GCTGACTGGCTTTGTGCTTAAAT	
12	MMP14_R_SET2	CAAAGTTCCCGTCACAGATGTTG	

Table 3.3.2 PCR reaction mix

Components	Volume
10X Taq buffer	2.5 µl
dNTP's (10 mM)	2 µl
Forward Primer (10 µM)	0.5 µl
Reverse Primer (10 µM)	0.5 µl
Taq polymerase (3U/µl)	0.33 µl
Nuclease Free Water	to 25 µl

Table 3.3.3 Thermal cycling condition for PCR reactions

Steps	Temperature	Time in Seconds	Cycle(s)
Initial Denaturation	95°C	60	1
Denaturation	95°C	30	30
Annealing	53-54°C	30	
Extension	72°C	45	
Final Extension	72°C	240	1
	4°C	Hold	

3.3.2 PCR gel purification

After the completion of PCR, the resulting product was subjected to agarose gel electrophoresis to visualize and confirm the amplification of the target region. Specifically, a 1% Agarose gel was prepared in 1X TAE buffer, and the PCR product was loaded into the wells. The gel was then run at 100V for 25 minutes using a horizontal electrophoresis unit. The resulting gel was then visualized using a UV transilluminator and compared to a DNA ladder to determine the expected molecular size of the target band. A clean scalpel was used to carefully excise the single amplified band from the gel.

To purify the PCR product, the NucleoSpin Gel and PCR Clean-up kit from Macherey-Nagel was used, following the manufacturer's instructions as detailed in Appendix V.

3.3.3 Sanger's sequencing

The purified PCR templates were quantified before proceeding to the sequencing PCR reaction. For sequencing, the BigDye Terminator v.3.1 cycle sequencing kit from Applied Biosystems was used. The PCR reaction mix consisted of 4 μ l of Nuclease Free Water (GeNei), 0.3 μ l of RR100 mix, 2 μ l of Sequencing buffer (5X), 50 ng of template DNA, and 2 μ l of primer (2 pmol) in a total reaction volume of 10 μ l (**Table 3.3.4**). The thermal cycling sequencing PCR reaction was performed for 30 cycles, with each cycle consisting of denaturation at 96°C for 10 seconds, annealing at 50°C for 5 seconds, and extension at 60°C for 4 minutes (**Table 3.3.5**). The reaction mix was then stored at -20°C until further use.

Table 3.3.4 Sequencing PCR reaction mix

Components	Volume
5X Sequencing buffer	2 μ l
Primer (1 pmol)	2 μ l
Template DNA (50-80 ng)	1 μ l
Ready Reaction mix (RR100)	0.33 μ l
Nuclease Free Water	to 10 μ l

Table 3.3.5 Thermal cycling condition for sequencing PCR reactions

Steps	Temperature	Time in Seconds	Cycle(s)
Initial Denaturation	96°C	60	1
Denaturation	96°C	10	30
Annealing	50°C	05	
Extension	60°C	240	
Hold	4°C	Until ready to purify	

3.3.4 Sequencing reaction cleanup

After the completion of sequencing PCR, the next step was to purify the sequencing reaction product to remove any unused di-deoxy nucleotide triphosphate, salts, and primers. The purification process began by preparing two master mixes. The first one, Master mix I, contained 10 μ l of Nuclease Free Water (NFW) and 2 μ l of 125 mM Ethylenediaminetetraacetic acid (EDTA) per reaction, while the second one, Master mix II, consisted of 2 μ l of 3M Sodium Acetate (NaOAc) pH-4.6 and 50 μ l of absolute ethanol per reaction. Then, 10 μ l of each sequencing reaction was mixed with 12 μ l of Master mix I, and the mixture was thoroughly mixed before adding 52 μ l of Master mix II to each reaction. The reaction mix was then incubated for 20 minutes, and after that, it was centrifuged at 12500g for 15 minutes at room temperature. The supernatant was discarded, and the pellets were washed with 250 μ l of freshly prepared 70% ethanol and centrifuged at 12500g for 10 minutes. The supernatant was then removed, and the pellet was left to air dry in a dark place at room temperature. The air-dry pellet was suspended in 12 μ l of HiDi-Formamide polymer and loaded into 96 well sequencing plates (Applied Biosystems) and covered with 96 well plate septa. The plate was subjected to denaturation at 95°C for 5 minutes and snap-chilled on ice for 2 minutes to linearize the templates before capillary electrophoresis in 3130xl Genetic Analyzer (Applied Biosystems). Finally, the sequencing reads were analyzed using Sequencing Analysis Software v5.3 (Applied Biosystems) and BioEdit v7.1. 1 (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>).

3.4 Cloning

3.4.1 DH5 α Ultra-competent cell preparation

The Inoue Method was followed to prepare DH5 α Ultra-competent cells (317). The process began by preparing a primary culture, in which a single colony streak was cultured in 5 ml of LB broth and incubated at 37°C for 16 hours. Next, 1 ml of this starter culture was transferred to a 250 ml conical flask containing LB broth media and incubated at 18°C with shaking at 225

rpm until the OD reached 0.55. Once the OD reached 0.55, the culture was transferred to ice and incubated for an additional 10 minutes. The cells were then harvested by centrifugation at 2500g for 10 minutes at 4°C. After centrifugation, the supernatant LB media was carefully removed using filter papers, and the pellet was gently dispersed in 20 ml of Inoue buffer by shaking the falcon tube.

Following the dispersion of cells, 1.5 ml of molecular grade DMSO was added to the mixture. To prepare for long-term storage, 50 µl aliquots were made into 1.5 ml microcentrifuge tubes while maintaining the tubes at 4°C during the entire process. The aliquots were then snap-chilled in liquid nitrogen and stored at -80°C for future use.

3.4.2 Cloning of MYB and SOX10 shRNA

Designing of shRNA probes:

Our general guideline to design shRNA included determining the optimal 21-mer targets in our gene of interest. We also considered factors such as low G-C content, low stability of the sense 3' end, avoiding targeting introns, and avoiding stretches of 4 or more nucleotide repeats. To minimize the degradation of off-target mRNAs, we used NCBI's BLAST program to select sequences with at least three nucleotide mismatches to all unrelated genes.

Once we identified the target sequences, the shRNA inserts were designed, as shown in **Figure 3.4.1**, by inserting the sense and antisense sequences in respective positions without changing the loop sequence and sticky end of EcoRI and AgeI. The designed shRNA oligos were procured from Integrated DNA Technologies (IDT). A detailed list of vectors and primers used are given in Appendix I and IV, respectively.

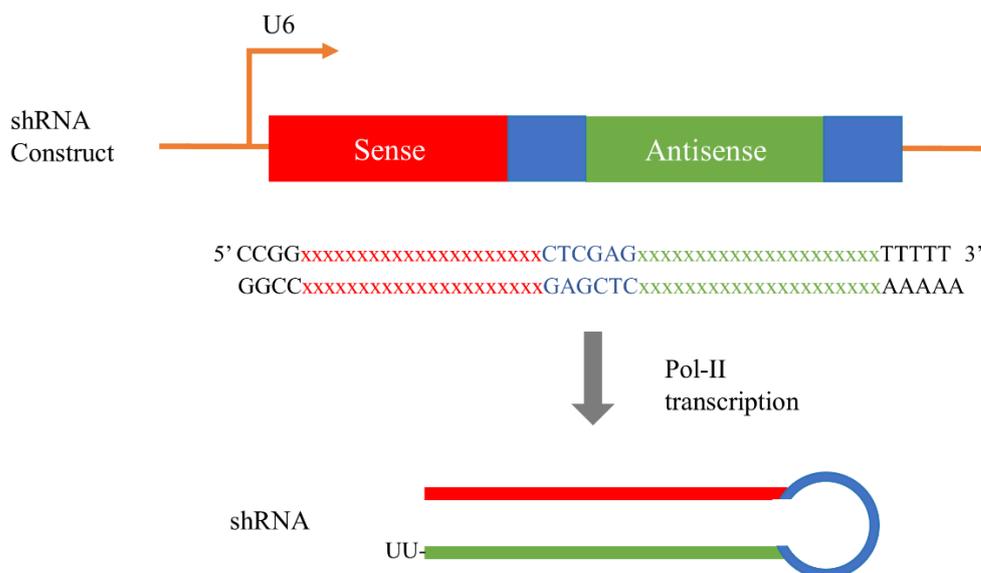


Figure 3.4.1 Designing of shRNA for pLKO.1 TRC Cloning Vector. The shRNA map showing position of sense (red coloured sequence) and antisense (green coloured sequence) of target sequence along with loop (blue coloured sequence) and sticky ends for EcoRI and AgeI (black coloured sequence).

Annealing of oligos: The forward and reverse oligos were annealed by resuspending them in nuclease-free water to a concentration of 20 μM and then incubating the mixture for 4 minutes at 95°C. This was followed by slow cooling to room temperature over a period of several hours (12 Hours). The components of the annealing reaction are given in **Table 3.4.1**.

Table 3.4.1 Annealing of Oligos

Components	Quantity
Forward oligo (10 μM)	5 μl
Reverse oligo (10 μM)	5 μl
NEB buffer 2 (10X)	5 μl
Nuclease-free Water	to 50 μl

pLKO.1 vector restriction digestion: The pLKO.1 TRC cloning vector was digested with AgeI and EcoRI restriction enzymes (

Table 3.4.2) for 2 hour at 37°C, and the digested DNA was purified using a Qiaquick gel extraction kit. The manufacturer's detailed instructions, have been described in Appendix V.

Table 3.4.2 Digesting pLKO.1 TRC Cloning Vector

Component	Quantity
pLKO.1 Vector	5 μ g
10X rCutSmart Buffer	5 μ l (1X)
EcoRI-HF	1.0 μ l (20 units)
AgeI-HF	1.0 μ l (20 units)
Nuclease-free Water	to 50 μ l

Ligation of shRNA into pLKO.1 vector: To ligate the target shRNA into the pLKO.1 vector (**Figure 3.4.2**), the annealed oligos were mixed with the purified digested vector in a T4 ligation reaction. Before the ligation reaction, the concentration and purity of DNA were determined using microvolume UV spectrophotometer NanoDrop-One (Thermoscientific). The ratio of insert to vector used for the ligation was 5:1. The amount of insert and vector DNA to be used in the ligation reaction mixture was calculated using the ligation calculator tool provided by New England Biolabs (NEB). The ligation reaction was carried out overnight at 16°C in a thermal cycler. The required components of the ligation reaction mixture are given in **Table 3.4.3**.

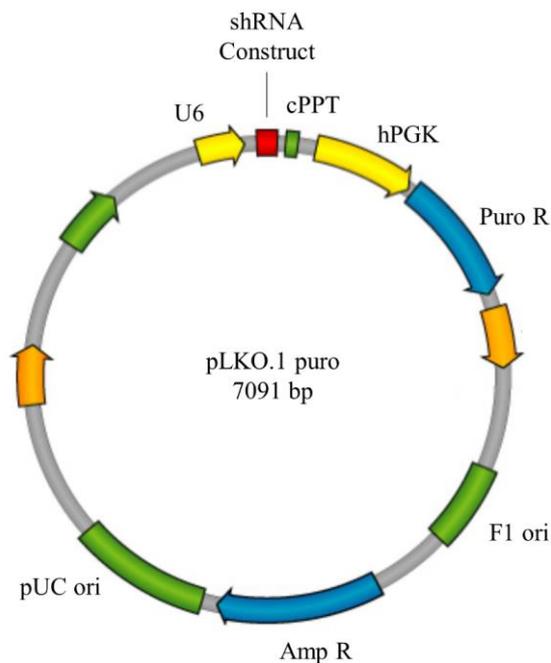


Figure 3.4.2 pLKO.1 vector map. *The pLKO.1 vector map showing the sequence position of shRNA insert under U6 promoter. Also, the map depicts positions of various selection markers such as puromycin and ampicillin.*

Table 3.4.3 T4 DNA ligase reaction mix

Components	Volume
Annealed oligo (58bp) (2 μ M)	1.65 ng (~2 μ l)
Digested pLKO.1 Vector (7026bp) (40 ng)	1 μ l
T4 DNA ligase buffer (10X)	2 μ l
T4 DNA ligase	1.0 μ l (20 units)
Nuclease-free Water	to 20 μ l

Transformation of ligation product: The process of transforming ultra-competent DH5 α cells with ligation product was conducted through a heat-shock method. Initially, the ultra-competent DH5 α cells were thawed on ice for 10 minutes. Following this, 10 μ l of ligation mix was added to the cells in the tube, flicked gently 3-4 times, and left to incubate on ice for another 20 minutes. Heat shock was then applied by incubating the tube in a water bath at 42°C for 45 seconds, after which the tube was immediately placed back on ice for 3 minutes. The cells were then treated with 500 μ l of SOC media, and the tubes were further incubated for 1 hour at 37°C in shaking at 220 rpm. Finally, the bacterial suspension was spread evenly on LB agar plates containing 100 μ g/ml ampicillin using an L-shaped spreader and incubated overnight at 37°C.

Screening of positive clones: One colony was selected from the plate and used to inoculate a master plate with colony number and 5 ml LB broth with 100 μ g/ml ampicillin, which was incubated at 37°C for 16 hours with shaking at 220 rpm. Plasmid DNA was then isolated from the bacterial culture using a miniprep kit (Qiagen) (The manufacturer's detailed instructions are in Appendix V). The purified plasmid was used for restriction digestion and the master plate with colony streak was kept aside to be used for colony PCR. Finally, positive clones were confirmed through colony PCR and restriction digestion.

Colony PCR: The master plate containing colony streak was used as template for each colony PCR. The components and cycling conditions used for each PCR reaction mixture are given in **Table 3.4.4** and

Table 3.4.5, respectively. The molecular size of amplified PCR product with insert cloned is expected to be 497bp and the size without insert is 439bp. The amplified PCR product was then analysed using agarose gel electrophoresis to determine the expected band position of the amplified shRNA insert by comparing it with corresponding molecular weight markers in the gel under an ultraviolet transilluminator.

Table 3.4.4 PCR reaction mix

Components	Volume
10X Taq buffer	2.5 μ l
dNTP's (10 mM)	2 μ l
hU6 Forward Primer (10 μ M)	0.5 μ l
LKO.1 Reverse Primer (10 μ M)	0.5 μ l
Taq polymerase (3U/ μ l)	0.33 μ l
Nuclease Free Water	to 25 μ l

Table 3.4.5 Thermal cycling condition for PCR reaction

Steps	Temperature	Time in Seconds	Cycle(s)
Initial Denaturation	95°C	60	1
Denaturation	95°C	30	30
Annealing	55°C	30	
Extension	72°C	45	
Final Extension	72°C	240	1
	4°C	Hold	

Colony screening through restriction digestion: Further confirmation of positive clones was done through the release of insert by the restriction digestion of the purified recombinant vector with EcoRI and NcoI. The reaction mixture used is given in **Table 3.4.6**. The double-digested

plasmid DNA was analyzed using agarose gel electrophoresis, and the expected band position for a positive clone was found by two fragments consisting 1973bp and 5053bp. The band positions were visualized by comparing them with corresponding molecular weight markers in the gel under an ultraviolet transilluminator.

Table 3.4.6 Screening for insert clones by restriction digestion

Component	Volume
pLKO.1 Vector (7026bp)	1 μ g
10X rCutSmart Buffer	5 μ l (1X)
EcoRI-HF	1.0 μ l (20 units)
NcoI-HF	1.0 μ l (20 units)
Nuclease-free Water	to 50 μ l

Confirmation of positive clones by sequencing: Finally, colonies found to be positive, based on the results from colony PCR and restriction digestion, were sequenced to confirm and rule out any mutations in the target site. The sequencing process was carried out using hU6 F primer. Sequencing PCR reaction mix for each primer and cycling conditions are given in **Table 3.4.7** and **Table 3.4.8**, respectively.

Table 3.4.7 Sequencing PCR reaction mix

Components	Volume
5X Sequencing buffer	2 μ l
hU6 F Primer (1 pmol)	2 μ l
Plasmid DNA (100 ng)	2 μ l
Ready Reaction mix (RR100)	0.33 μ l
Nuclease Free Water	to 10 μ l

Table 3.4.8 Thermal cycling condition for sequencing PCR reactions

Steps	Temperature	Time in Seconds	Cycle(s)
Initial Denaturation	96°C	60	1
Denaturation	96°C	10	30
Annealing	50°C	05	
Extension	60°C	240	
Hold	4°C	Until ready to purify	

Plasmid isolation by Midiprep or Maxiprep: After confirming the positive clones, we performed midiprep/maxiprep to obtain a higher yield of the selected positive clone. To carry out midiprep, we utilized the plasmid midi kit. The protocol for isolating plasmid DNA using the aforementioned kit has been provided in Appendix V.

3.4.3 Cloning of Luciferase vectors

Designing of luciferase assay inserts: The 50 to 55 bp long oligoes consisted of 30 bp promoter sequence having SNPs positioned at the centre as insert flanking KpnI and XhoI restriction sites along with 4 bp spacers to facilitate binding of restriction enzymes. These oligoes were commercially procured from Integrated DNA Technologies (IDT). The detailed pictorial representation of luciferase inserts design were explained in following **Figure 3.4.3**.

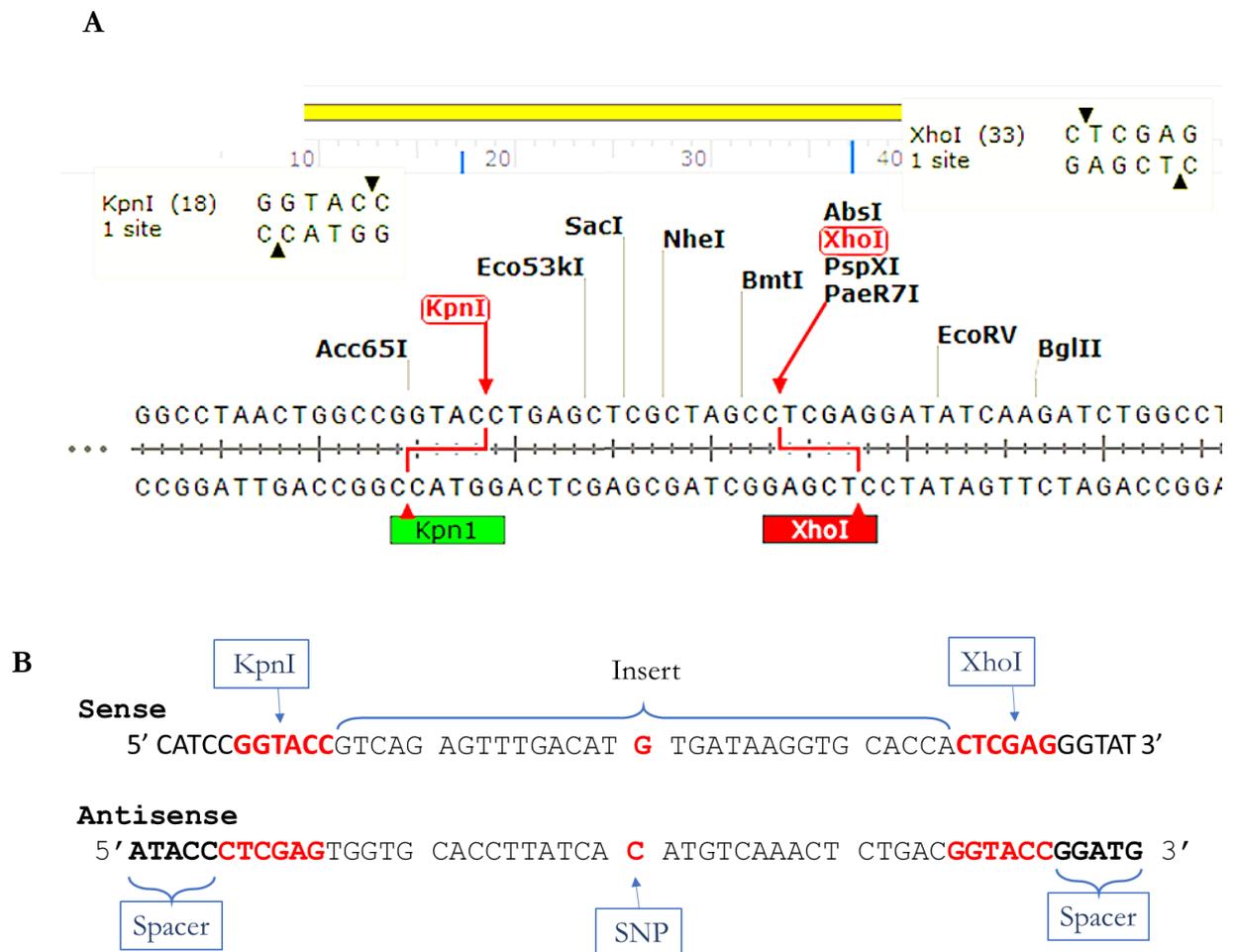


Figure 3.4.3 Designing of luciferase assay inserts. (A) The PGL 4.23 vector sequence showing position of *KpnI* and *XhoI* restriction enzyme and formation of sticky ends. (B) The luciferase inserts with both sense and antisense sequences showing *KpnI* and *XhoI* restriction site along with spacer sequence.

Annealing of oligos: The forward and reverse oligos were annealed by resuspending them in nuclease-free water to a concentration of 20 μM and then incubating the mixture for 4 minutes at 95°C. This was followed by slow cooling to room temperature over a period of several hours (12 Hours). The components of the annealing reaction are given in following

Table 3.4.9.

Table 3.4.9 Annealing of Oligos

Components	Volume
Forward oligo (10 μM)	5 μl
Reverse oligo (10 μM)	5 μl
NEB buffer 2 (10X)	5 μl
Nuclease-free Water	to 50 μl

pGL4.23 luciferase vector restriction digestion: The pGL4.23 luciferase reporter vector was digested with *KpnI* and *XhoI* restriction enzymes (**Table 3.4.10**) for 2 hour at 37°C, and the digested DNA was purified using a Qiaquick gel extraction kit. The manufacturer's detailed instructions, have been described in Appendix V.

Table 3.4.10 Digestion of pGL4.23 luciferase vector

Components	Volume
pGL4.23 Vector	5 μg
10X rCutSmart Buffer	5 μl (1X)
<i>KpnI</i> -HF	1.0 μl (20 units)
<i>XhoI</i> -HF	1.0 μl (20 units)
Nuclease-free Water	to 50 μl

Ligation of luciferase insert into pGL4.23: The ligation of luciferase inserts was carried out by using T4 DNA ligase enzyme (New England Biolabs). Prior to the ligation reaction, the purity and concentration of insert DNA was assessed using microvolume UV spectrophotometer NanoDrop-One (Thermoscientific). The ratio of insert to vector for ligation was maintained at

5:1. The ligation calculator tool provided by New England Biolabs (NEB) was used to calculate the amount of insert and vector DNA needed for the ligation reaction mixture. The ligation reaction was performed overnight at 16°C in a thermal cycler. The required components for the ligation reaction mixture were calculated and are listed in **Table 3.4.11**.

Table 3.4.11 T4 DNA ligase reaction mix

Components	Quantity
Annealed oligo (54bp)	2.47 ng
Digested pGL4.23 vector (4268bp) (39 ng)	1 µl
T4 DNA ligase buffer (10X)	2 µl
T4 DNA ligase	1.0 µl (20 units)
Nuclease-free Water	to 20 µl

Transformation of ligated product: The process of transforming ultra-competent DH5α cells with ligation product was conducted through a heat-shock method. The detailed procedure has been previously explained in the **Section 3.4.2**.

Screening of positive clones: The screening of positive clones was performed by first selecting a single colony from the agar plate and using it to inoculate a master plate labeled with the corresponding colony number, as well as a 5 ml LB broth containing 100 µg/ml ampicillin. This culture was then incubated for 16 hours at 37°C with shaking at 220 rpm. Plasmid DNA was extracted from the bacterial culture using a miniprep kit (Qiagen), following the detailed instructions provided by the manufacturer (see Appendix V). The purified plasmid DNA was subjected to sequencing confirmation, and the master plate containing the colony streak was utilized for colony PCR. Ultimately, positive clones were confirmed through both colony PCR and sequencing.

Colony PCR: In order to screen the positive clones, each colony on the master plate was subjected to colony PCR using specific components and cycling conditions (**Table 3.4.12** and **Table 3.4.13**, respectively). The expected molecular size of the PCR product with the cloned insert was 199bp, while the size without insert was 169bp. After amplification, the PCR product was analysed by agarose gel electrophoresis to determine the band position of the amplified

shRNA insert, which was compared with corresponding molecular weight markers under an ultraviolet transilluminator.

Table 3.4.12 Colony PCR reaction mix

Components	Volume
10X Taq buffer	2.5 μ l
dNTP's (10 mM)	2 μ l
RV Primer (10 μ M)	0.5 μ l
RV Reverse Primer (10 μ M)	0.5 μ l
Taq polymerase (3U/ μ l)	0.33 μ l
Nuclease Free Water	to 25 μ l

Table 3.4.13 Thermal cycling condition for colony PCR reactions

Steps	Temperature	Time in Seconds	Cycle(s)
Initial Denaturation	95°C	60	1
Denaturation	95°C	30	30
Annealing	55°C	30	
Extension	72°C	45	
Final Extension	72°C	240	1
	4°C	Hold	

Sequencing confirmation of positive clones: Sequencing confirmation of positive clones was performed to validate and eliminate the possibility of any mutations in the target site. Only colonies that were confirmed as positive based on colony PCR results were sequenced. RV forward primer was used in the sequencing process. The sequencing PCR reaction mix and cycling conditions were prepared as described in **Table 3.4.14** and **Table 3.4.15**, respectively.

Table 3.4.14 Sequencing PCR reaction mix

Components	Volume	
5X Sequencing buffer	2 μ l	
hU6 F Primer (1 pmol)	2 μ l	
Plasmid DNA (100 ng)	2 μ l	
Ready Reaction mix (RR100)	0.33 μ l	
Nuclease Free Water	to 10 μ l	

Table 3.4.15 Thermal cycling condition for sequencing PCR reactions

Steps	Temperature	Time in Seconds	Cycle(s)
Initial Denaturation	96°C	1 min	1
Denaturation	96°C	10	30
Annealing	50°C	05	
Extension	60°C	4 min	
Hold	4°C	Until ready to purify	

Plasmid isolation by Midiprep: Following the confirmation of positive clones, midiprep procedures were conducted to obtain a greater yield of the selected clones. For midiprep, the plasmid midi kit was employed, and the detailed protocol for plasmid DNA isolation utilizing this kit is provided in Appendix V.

3.4.4 sgRNA design and CRISPR-Cas9 cloning

Identification of target site: The target sequence was scanned for the presence of PAM sequence nearby MMP14 promoter variants rs1004030 and rs1003349 for genomic editing. Identifying the PAM (protospacer adjacent motif) site is essential, as it is required for Cas9 binding. The PAM is typically NGG for the commonly used Cas9 from *Streptococcus pyogenes* (Figure 3.4.4). Our study used the sgRNA design tool CHOPCHOP (<https://chopchop.cbu.uib.no/>; last accessed on June 12, 2019). The tool helps to identify potential targets with increased specificity and minimum off-target effects. The list of sgRNAs used is given in Table 3.4.16 and Figure 3.4.4.

Table 3.4.16 List of oligonucleotides used to generate sgRNAs for the generation of CRISPR-Cas9 based MMP14 promotor deletion construct

S. No.	Oligonucleotide name	Primer Sequence 5' to 3'
1	M14sgRNA set1 sense	CACCGCAGCCCCCTGCTGTCCATCG
2	M14sgRNA set1 antisense	AAACCGATGGACAGCAGGGGGCTGC
3	M14sgRNA set2 sense	CACCGTCCTTTCCTGGTTGGGGACG
4	M14sgRNA set2 antisense	AAACCGTCCCCAACCAGGAAAGGAC

sgRNA design: The sgRNA1 and sgRNA2 was designed as showed in **Figure 3.4.4 A and B**. The sgRNA was designed consists of 20 bases complementary sequence to target genomic region following the 5'-NGG motifs in target genomic DNA, which is required for Cas9 cleavage. The BbsI site overhangs were added to sgRNA guide sequence to facilitate ligation into pSpCas9(BB) vector. In addition, a G-C base pair (grey rectangle) was added at the 5' end of the sgRNA guide sequence, which is required for optimal transcription under U6 promoter.

pSpCas9(BB) vector restriction digestion: The restriction digestion of pSpCas9(BB) with BbsI clears the Type II restriction sites, creating a linearized vector with BbsI sticky ends. These sticky ends are then ligated with the annealed sgRNA's with BbsI sticky overhangs. The pSpCas9(BB) vector was digested with BbsI restriction enzyme (**Table 3.4.17**) for 2 hours at 37°C, and the digested DNA was purified using a Qiaquick gel extraction kit. The manufacturer's detailed instructions have been described in Appendix V.

The process of annealing, ligation, and transformation and screening of colonies were already discussed in the previous section. For a detailed procedure please refer to **Section 3.4.2**.

Table 3.4.17 Digestion of pSpCas9(BB) vector

Component	Volume
pSpCas9(BB) Vector	5 µg
10X rCutSmart Buffer	5 µl (1X)
BbsI-HF	1.0 µl (20 units)
Nuclease-free Water	to 50 µl

3.5 CRISPR Cas9 mediated genome editing

The CRISPR/Cas9 system was utilized to delete the 119 bp MMP14 promoter region flanking variants rs1004030 and rs1003349 in G415 cells. The identification of target site, sgRNA design and pSpCas9(BB) plasmid constructions detailed explanation is given in previous **Section 3.4.4**. In brief, the set of sgRNA was designed using the online tool (<https://chopchop.cbu.uib.no/>; last accessed on June 12, 2019) targeting 119 bp, spanning promoter variants, with minimal off-target effects. The sgRNA oligonucleotides (provided in Appendix II) were procured commercially from Integrated DNA Technologies [IDT] and were cloned into the BbsI restriction site of pSpCas9(BB)-2A-Puro (PX459) (Gift from Feng Zhang, Addgene #62988).

To accomplish the deletion, the G415 cells were co-transfected with sgRNA1 and sgRNA2 containing CRISPR construct at 70% confluency using lipofectamine 3000 (Invitrogen). The cells were then subjected to puromycin-based selection (1 $\mu\text{g}/\mu\text{l}$ in complete media) 24 hours post-transfection. Single-cell colonies were subsequently isolated and grown further. The detailed transfection and generation of stable line procedure is explained in **Section 3.6.4**. Genomic DNA was isolated from the single-cell derived colonies using DNeasy Blood & Tissue Kits (Qiagen). The homozygous deletion was confirmed through PCR and Western blot analysis. The confirmed clones with homozygous deletion were selected for further experiments.

3.6 Cell culture and cell-based assays in current study

3.6.1 Culturing of human gallbladder cancer and HEK293T cell lines

The TGBC1TKB and HEK293T cell lines were cultured and maintained in DMEM supplemented with 10% heat-inactivated FBS, 0.025 $\mu\text{g}/\text{ml}$ Amphotericin B, 100 units/ml of Penicillin, and 50 $\mu\text{g}/\text{ml}$ Streptomycin. The GBC cell line G415 was cultured and maintained in RPMI1640 with 15% heat-inactivated FBS, 100 units/ml of Penicillin, 0.025 $\mu\text{g}/\text{ml}$ Amphotericin B, and 50 $\mu\text{g}/\text{ml}$ Streptomycin. All the cell lines were grown in T-25/T-75 flasks at a temperature of 37°C, 5% CO₂, and 95% humidity in an incubator. The culture medium was changed every two days, and the cells were sub-cultured at 70-80% confluence using trypsin-EDTA solution. The cells were then harvested and used for further experiments.

3.6.2 Subculture, trypsinization and passage of cell lines

To subculture or passage the cells, they were allowed to grow until reaching 70-80% confluency. The old media was discarded from the flasks, and the cells were washed twice with DPBS. Then, 1 ml and 3 ml of pre-warmed trypsin-EDTA solution were added to the T-25 and T-75 flasks, respectively, and the flasks were placed back in the 37°C incubator to allow the cells to detach from the flask surface. Once the cells were detached, trypsin activity was neutralized by adding complete growth media that contained FBS. A sterile serological pipette was used to aspirate the cell suspension from the flasks, and the suspension was then centrifuged at 300g for 5 minutes. The supernatant was discarded, and the cell pellet was resuspended in complete media. The total viable cell count was determined using the trypan-blue method, and a hemocytometer under an inverted microscope. Afterward, the required number of viable cells was seeded into culture plates in fresh flasks.

3.6.3 Storage and revival of mammalian cell lines

To store the cells, they were allowed to grow up to 70-80% confluency and were then trypsinized as previously described (**Section 3.6.2**). The resulting cell pellet was resuspended in complete media, and the total viable cell count was determined using the trypan-blue method and a hemocytometer under an inverted microscope. The cell pellet was resuspended in a freezing mix consisting of 10% DMSO in FBS, at a concentration of 1×10^6 cells/ml. Next, 1 ml of the cell suspension was aliquoted into each cryovial, transferred to a -1°C cooler, kept overnight at -80°C , and finally transferred to liquid nitrogen for long-term storage.

To revive the cryopreserved cells, the frozen cell vials were removed from liquid nitrogen and immediately placed in a 37°C water bath. Next, the cell suspension from the vial was transferred to a 15 ml centrifuge tube containing pre-warmed complete growth media. The centrifuge tube containing the cell suspension was then centrifuged at 200g for 5 minutes. The resulting cell pellet was resuspended in a sufficient volume of complete growth media and transferred to a T-25 cell culture flask. The flask was then incubated inside a 37°C incubator, maintaining 5% CO_2 and 95% humidity.

3.6.4 Stable and transient transfection of cell lines

For transfecting various vector constructs into different cell lines, Lipofectamine 3000 reagent was employed. The cells were seeded at a suitable density prior to transfection, aiming for 70-80% confluency within 24-48 hours. To ensure successful transfection, complete growth media was replaced with fresh media one hour before transfection. A master mix of DNA was created by diluting DNA in Opti-MEM Medium and adding P3000 Reagent in a DNase, RNase, and pyrogen-free 1.5 ml microfuge tube. Similarly, in another 1.5 ml tube, Lipofectamine 3000 reagent was added to Opti-MEM medium and mixed well. The appropriate ratio of DNA:P3000:Lipofectamine:Cells for different cell culture plates was established and presented in **Table 3.6.1**. Both tubes were incubated at 25°C for 5 minutes, then mixed by gentle pipetting and further incubated at 25°C for 20 minutes to form the transfection complex. The reaction mixture containing the transfection complex was added drop-wise to the cells in the plate, followed by gentle swirling for uniform distribution of the complex. Cells with the complex

mix were transferred to an incubator maintained at 37°C and 5% CO₂ for 24 hours and treated according to the experimental protocol.

Table 3.6.1 Reaction setup for Lipofectamine 3000 mediated transfection[#]

Culture plate	Vol. of growth medium	DNA Mix I			Lipid Mix II	
		Opti-MEM (µl)	P3000 (µl)	DNA (µg)	Opti-MEM (µl)	L3000 (µl)
6-well	2 ml	125	5	2.5	125	7.5
12-well	1 ml	50	2	1	50	3
24-well	500 µl	25	1	0.5	25	1.5
96-well	100 µl	5	0.2	0.1	5	0.3

[#]Adapted from reagent user guide, Thermofischer scientific

After 48 hours of transfection, antibiotic selection was initiated to isolate stable clones. To select stable clones for TGBC1TKB and G415 cells transfected with MMP14 expression vector, puromycin was added at a concentration of 2 µg/ml. On the other hand, for TGBC1TKB and G415 cells transfected with shRNA for SOX10 and MYB, puromycin was used at a concentration of 1 µg/ml. G415 cells that were transfected with pSpCas9(BB)-M14^{119/-} vector construct was subjected to puromycin selection at a concentration of 2 µg/ml to obtain stable clones. The selected clone's expression were then verified by Western blot and maintained at a concentration of 1 µg/ml puromycin.

3.6.5 Cell proliferation assay

For the cell proliferation/growth assay, 2×10^3 cells were seeded in each well of a 96-well plate in complete growth media and incubated in a humidified incubator at 37°C with 5% CO₂ for 24 hours. Afterward, complete growth media was replaced with reduced serum growth media containing 2% FBS, and the cells were incubated for 96 hours to allow growth. Following this, the growth media was replaced with 100 µl of fresh complete media and 20 µl of pre-warmed CellTiter 96 Aqueous One solution reagent was added. The plate was then returned to the incubator for 1 hours. Absorbance was measured at a wavelength of 490 nm using the Varioscan multimode microplate reader (Thermoscientific). The experiment was performed with five replicates in each group.

3.6.6 Colony formation assay

For this experiment, 1000 cells (For TGBC1TKB, 1500 cells) were seeded into each well of a 6-well plate with complete media and incubated at 37°C in a humidified incubator with 5% CO₂

for 15-20 days. After the incubation period, the cells were washed once with 1X PBS buffer and fixed with methanol and acetic acid (3:1) fixation solution for 5 minutes. Subsequently, the cells were washed with 1X PBS buffer and stained with 0.5% crystal violet solution in PBS for 20 minutes. After a gentle wash under tap water, the plates were dried by keeping them inverted on paper towels. Digital photographs of the cell colonies were captured using a Nikon camera and the colonies were counted using Fiji, an ImageJ software plugin (318). The experiment was performed in triplicates.

3.6.7 Scratch wound healing assay

For this assay, approximately 0.3×10^6 cells were seeded in each well of a 12-well plate with complete growth medium. The plates were gently swirled to ensure even cell distribution and incubated at 37°C until 90-100% confluence was reached. Once the cells reached confluency, the growth medium was removed and the cells were washed twice with pre-warmed 1X DPBS. A fine scratch was made in the centre of each well using a P200 (200 µl) pipette tip, and the wells were washed again with pre-warmed 1X DPBS to remove any detached cells. Then, 2 ml of complete medium was added to each well and the plates were placed back into the incubator. Images of the wound were taken using an inverted microscope (Nikon, 10X objective lens) every 12 hours until complete wound closure. The area of wound closure was analysed using ImageJ software (318). This assay was performed in triplicate.

3.6.8 Transwell migration assay

The Millipore transwell chambers with 8 µm pore size were used to conduct the transwell migration assay. Prior to the assay, the cells were starved overnight using media with reduced serum. The viable cell count was determined, and 0.5×10^6 cells suspended in 400 µl of media were seeded in membrane filter inserts, which were then placed in a 12-well plate. The lower chamber of the well was filled with complete media (1.2 ml) to induce cell migration, and the plates were incubated for 12 hours. After incubation, the plates were removed from the incubator, and the inserts were carefully removed from the plate with forceps and washed twice with PBS. Then, 500 µl of 4% PFA was added to each insert, and the inserts were incubated for 10 minutes to facilitate fixation. The inserts were then washed twice with PBS, and 500 µl of methanol was added to each insert, followed by incubation for 20 minutes to facilitate permeabilization. After two additional washes with PBS, 500 µl of either Crystal Violet solution or Giemsa solution was added to the inserts, which were then incubated in the dark for 15 minutes. The inserts were washed twice with PBS, and non-migrating cells on the upper side

of the inserts were removed using a cotton swab. The experiment was performed in triplicates, and images were captured using an inverted microscope (BX53, Olympus). ImageJ software (318) was used to count the cells in five different random view fields.

3.6.9 Transwell Matrigel invasion assay

To conduct the transwell Matrigel invasion assay, the same steps were followed as the transwell migration assay apart from an additional step for pre-coating of Corning Matrigel (Growth Factor Reduced Basement Membrane Matrix) (Corning) in the upper chamber of Millipore transwell inserts (8 μm pore size). For pre-coating, 100 μl of 1 mg/ml growth factor reduced Matrigel was added to the upper chamber of the insert and allowed to settle down for 4-6 hours in the incubator at 37°C. The remaining steps of the method were the same as described in the previous section (**Section 3.6.8**).

3.7 Luciferase assay

For the reporter assays, we used pGL4.23 luciferase reporter vector with minimal promoter and pGL4.74 renilla vector, which were obtained from Promega. We commercially procured 50 to 55 bp long oligoes consisting of a 30 bp promoter sequence with SNPs located in the center, flanked by KpnI and XhoI restriction sites (IDT). The oligoes were annealed and cloned into pGL4.23 vector using double digestion at the KpnI-XhoI site (KpnI-HF and XhoI-HF, New England Biolabs). The detailed cloning procedure is explained in **Section 3.4.3**. For transfection, the cells were seeded in a 12-well plate and transfected with pGL4.23 vector (1 μg) and pGL4.74 vector (10ng) using Lipofectamine 3000 (Invitrogen) at 80% confluency. The detailed transfection procedure is explained in **Section 3.6.4**. Transfection efficiency was normalized by renilla reporter activity. After 24 hours of post-transfection, we prepared cell lysates using the Dual-Luciferase reporter assay kit (Promega), as per the manufacturer's instructions. The reporter activities were measured with a Varioskan Flash Multimode reader (Thermo Fisher Scientific). The values for luciferase activity were normalized with renilla reporter activity and used for further analysis. Each experiment was repeated independently with three replicates.

3.8 Electrophoretic mobility shift assay (EMSA)

In order to identify putative transcription factor binding elements around rs1004030 and rs1003349 variants in the MMP14 promoter region, we used the JASPAR database with a

default relative profile score threshold of 80% (<https://jaspar.genereg.net/>, last accessed on March 24, 2019). We procured 39-mer oligonucleotides centered at rs1004030 and rs1003349 corresponding to all allelic combinations with and without 5'-end biotin labeling from IDT. A detailed list of EMSA oligoes is provided in Appendix II. The probes were annealed by incubating at 95°C for 5 minutes and gradually cooling overnight. The detailed annealing procedure has been previously explained in the **Section 3.4.2**. We isolated nuclear protein from TGBC1TKB cells using the NE-PER kit (Invitrogen) and estimated the concentration using the BCA protein assay kit (Invitrogen). We performed EMSA using the LightShift Chemiluminescent EMSA Kit (Invitrogen) according to the manufacturer's instructions. We used 2 µg of nuclear protein extract and 20 fmol of 5'-biotinylated annealed double-strand oligonucleotides for protein-DNA binding reactions. The specificity of binding was confirmed by challenging it with 200-fold excess (4 pmol) of unlabeled double-stranded oligonucleotides. We performed a supershift assay by incubating the final binding reaction mix with 2 µg of MYB and SOX10 antibodies (Cell signalling) for 20 minutes. We ran the DNA-protein complex on an 8% native polyacrylamide gel in TBE (0.5X) buffer. After transferring the complex to a nylon membrane, and post-UV crosslinking, the membrane was detected using a chemiluminescence detection kit (Invitrogen). We repeated each experiment independently with at least three replicates.

3.9 Chromatin immunoprecipitation assay (ChIP)

The chromatin immunoprecipitation (ChIP) assay was performed using the ChIP Kit following the manufacturer's instructions (Abcam). To summarize, cells were fixed with 4% buffered formalin and lysed using buffer A and buffer B provided in the kit. The cells were then centrifuged and resuspended in buffer C and buffer D/PI mix. The resulting DNA was sheared using an ultrasonicator (Cole-Parmer) to obtain optimal DNA fragment sizes of 100-1000 bp. The sheared DNA was incubated overnight with 5 µg of ChIP-grade antibodies for SOX10 (Cell Signalling) and MYB (Cell Signalling). A negative isotype control was also included using normal rabbit IgG isotype control antibody (Abgenex). The following day, protein A sepharose beads were added for immunoprecipitation. After immunoprecipitation, DNA was purified, and the fold enrichment of the target region was determined using specific primer sets (provided in Appendix IV) through quantitative real-time PCR (qRT-PCR).

3.10 Immunohistochemistry

3.10.1 Poly-L-Lysine coating of slides

The glass slides were coated with poly-L-lysine by as per the following protocol. Initially, the slides were cleaned with soap and soaked in 1% acetic acid (acetic acid: ethanol in 1:99 ratio) for 20 minutes. The slides were rinsed twice with tap water and distilled water, and then warmed and dried in an incubator. The dried slides were dipped twice in 0.01% poly-L-lysine solution and left for 5 minutes each time. Afterward, the coated slides were left to dry overnight at 37°C.

3.10.2 Immunohistochemistry staining

The 5 µm thin sections were cut from FFPE biopsy specimens and mounted on poly-L-lysine coated glass slides. The slides were then deparaffinized by heating at 80°C for 1 hour, followed by two consecutive xylene washes. The tissue sections were rehydrated by immersing in a gradient of alcohol from 100% to 50%. Heat induced epitope retrieval was carried out using high pH buffer for MMP14, and low pH citrate buffer for MMP7 and MMP2. After this, all the slides were treated with Envision Peroxidase Blocker (Dako) for 15 minutes, followed by incubation with the respective primary antibodies for 1 hour in a humidified chamber at room temperature. Subsequently, the sections were washed and incubated with Envision Flex HRP secondary antibody (Dako) for 30 minutes, followed by addition of liquid DAB substrate (Dako). The sections were then counterstained with haematoxylin, dehydrated, and mounted with coverslip. To normalize the staining scores across multiple rounds of IHC for MMP2, MMP7, and MMP14 antibodies, HeLa cell blocks were used as positive controls. A list of the antibodies used in the entire procedure can be found in Appendix III.

3.10.3 Immunohistochemistry scoring

The immunohistochemistry scoring was performed according to the Allred Score system described by Fedchenko et al, 2014 (319). Thin sections of tumour tissue and adjacent normal tissue were used for the scoring, and the expression levels of proteins were evaluated based on the brown signals present in the cytoplasm, nucleus, or membrane of the cells. The intensity of staining was scored from 0 to 3 (0 = negative, 1 = weak, 2 = moderate, and 3 = strong), while cell positivity was scored from 0 to 5 (0 = 0% cell positivity, 1 = 1% cell positivity, 2 = 1%-10% cell positivity, 3 = 11%-33% cell positivity, 4 = 34%-66% cell positivity, 5 = 67%-100% cell positivity). The scores for intensity and cell positivity were added to obtain the final Allred score for each tissue section. Tissue sections with Allred scores of 0-2 were considered

negative/weak, those with scores of 3-6 were considered moderate, and those with scores of 7-8 were considered intense/strong. Imaging was performed using an upright light microscope (CX31, Olympus) with 10X and 40X objective lenses.

3.11 Western blotting

3.11.1 Preparation of mammalian cell lysate

Cells were cultured in cell culture plates and dishes in a CO₂ incubator with 37°C temperature and 95% humidity until they reached 70-80% confluency. After washing the cells with ice-cold 1X PBS, ice-cold RIPA buffer (Thermoscientific) was added in a suitable volume to the cells for lysis. Ice-cold scrapers were used to ensure maximum lysis efficiency. The lysate was collected in cold-microcentrifuge tubes and centrifuged at approximately 14,000g at 4°C for 15 minutes to collect the supernatant after pelleting down the cell debris. An aliquot of the sample was stored at -80°C for protein estimation by BCA method. To the remaining lysate, 4X Laemmli Sample Buffer (BioRad) was added to a final concentration of 1X and boiled at 95°C for 5 minutes. The protein samples were stored at -80°C until further use.

3.11.2 Protein estimation

To estimate the protein concentration of the whole cell lysate, BCA protein assay kit (Pierce) was used. Initially, BSA standards were prepared by making serial dilutions of Bovine Serum Albumin (BSA) at 1.5-fold from 2 mg/ml (stock solution) to 20 µg/ml. Then, BCA reagent A and BCA reagent B were mixed at a ratio of 50:1 to prepare a working reagent. Triplicates of 10 µl each of the whole cell lysate samples and BSA standards were dispensed in a 96 well plate. Next, 200 µl of the working reagent was added to each well followed by incubation at 37°C for 30 minutes. After incubation, the plate was cooled to room temperature for 10 minutes and then the absorbance was measured at 562 nm using the Varioskan LUX multimode microplate reader (Thermoscientific).

3.11.3 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used to separate and visualize proteins based on their molecular mass. A Mini-PROTEAN tetra system (Bio-Rad) was used for gel electrophoresis. Stacking gels at 5% and resolving gels at 10% or 12% were prepared for SDS-PAGE. The volume and concentration of components required to cast a gel of predetermined volume are specified in **Table 3.11.1**. Electrophoresis was performed using 1X TGS electrophoresis buffer in the electrophoresis

chamber, which was connected to a power pack. The protein samples stored in a -80°C freezer were thawed on ice and boiled at 95°C for 5 minutes. $20\ \mu\text{g}$ of protein lysates were loaded into each well along with the prestained protein ladder as the size standard and were run at a constant voltage of 100V until the bromophenol blue dye front exited from the base of the gel. A detailed description of various stock solutions and buffers required for the entire procedure can be found in Appendix VI.

Table 3.11.1 The volumes and concentrations of components to prepare SDS-PAGE gels*

Components	Resolving gel (10%) for 10ml	Resolving gel (12%) for 10ml	Components	Stacking gel (5%) for 3 ml
	Volume in ml	Volume in ml		Volume in ml
Milli Q water	4	3.3	Milli Q water	2.1
1.5 M Tris (pH 8.8)	2.5	2.5	1 M Tris (pH 6.8)	0.38
30% acrylamide	3.3	4	30% acrylamide	0.5
10% APS	0.1	0.1	10% APS	0.03
10% SDS	0.1	0.1	10% SDS	0.03
TEMED	0.004	0.004	TEMED	0.003

*Adapted from (320).

3.11.4 Immunoblotting

In this study, immunoblotting was carried out after performing SDS-PAGE to transfer and visualize the proteins. After SDS-PAGE, the PVDF membrane was activated with methanol and then placed on the gel. The transfer cassette was assembled and transferred to the Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad) filled with chilled transfer buffer, and connected to a power-pack. The assembly was then transferred to a 4°C room for 16 hours at a constant voltage of 30V. Following the transfer, the membrane was stained with Ponceau stain and then washed with 1X TBST to remove the stain. Non-specific binding of IgG was blocked by incubating the membrane in a blocking solution of 5% BSA or 5% skim milk powder in 0.01% TBST for one hour in a shaker. The membrane was then incubated with primary antibody solution at an appropriate dilution (Appendix III) for 1 hour at room temperature followed by overnight incubation at 4°C . On the following day, the membrane was washed thrice with 1X TBST for 5 minutes each, and then transferred to the corresponding HRP-conjugated secondary

antibody solution. The membrane was kept at constant shaking at room temperature for 1 hour during incubation with the secondary antibody, followed by washing thrice with TBST for 5 minutes each. SuperSignal West Femto Maximum Sensitivity Substrate (Thermoscientific) was used for immunoblot detection, and Chemidoc XRS⁺ equipped with Image Lab analysis software version 6 (Bio-Rad) was used for imaging.

3.11.5 Stripping and reprobing of blots

To probe the PVDF membrane with a different antibody, stripping of the previous antibody was carried out using restore plus stripping buffer (Thermoscientific). The membrane was washed thrice with 0.1% TBST for 5 minutes each before incubating with the stripping buffer at 37°C with constant shaking for 15-20 minutes. After stripping, the membrane was washed thrice with TBST for 10 minutes each at room temperature. Blocking of the membrane, incubation with the new primary antibody, and detection were carried out as described in **Section 3.11.4**.

3.12 *In-silico* analysis

3.12.1 Differential Gene Expression Analysis of RNA-Seq data from GEO database

RNA-Seq Data Extraction from GEO database: RNA-Seq data in the fastq format for the two distinct gallbladder cancer study were downloaded from the NCBI Gene Expression Omnibus (GEO) database (accession number: GSE139682 and GSE132223) using the SRA toolkit.

Quality Control and Processing of Raw Reads: Quality control and processing of raw reads were performed using the fastp program (<https://github.com/OpenGene/fastp>). The raw reads were filtered for low-quality reads, trimmed for adapter sequences, and filtered for reads shorter than 50 nucleotides.

Alignment of Clean Reads to the Reference Genome:

Clean reads were aligned to the reference genome using the HISAT2 program (<https://ccb.jhu.edu/software/hisat2/index.shtml>). HISAT2, a software that maps RNA-Seq reads to a genome with a high degree of accuracy and sensitivity. The reference genome and annotation files were obtained from the Ensembl database for homosapiens. HISAT2 was run with default parameters to align the reads to the reference genome.

Transcript Assembly and Quantification: Transcript assembly and quantification were performed using the StringTie program (<https://ccb.jhu.edu/software/stringtie/>) to generate

transcript-level expression values. The output file from HISAT2 was used as input for StringTie, and transcript assembly was performed with default parameters.

Differential Gene Expression Analysis:

Differential gene expression analysis was performed using the Deseq2 R package (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>). Gene expression counts were extracted from the output file generated by StringTie, and imported into R for further analysis. Deseq2 was used to calculate the normalized read counts, and to identify differentially expressed genes (DEGs) between the GBC and control groups. The normalized counts and differential expressed gene sets representing upper quartile (top 25% of total differentially expressed genes) were used for pathways and network analysis by DAVID, GSEA and Cytoscape tool.

Functional enrichment and pathway analysis: The normalized counts and differential expressed gene sets representing upper quartile (top 25% of total differentially expressed genes) were further used for pathways and network analysis. Functional enrichment analysis was performed using the GSEA package. Gene ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) (DAVID, <http://david.abcc.ncifcrf.gov/>), pathway analysis were performed to identify enriched biological processes and pathways. Pathway enrichment analysis results are further visualized and interpreted in Cytoscape using its EnrichmentMap.

3.12.2 Gene Set Enrichment Analysis (GSEA) for differentially expression genes in two gallbladder cancer expression data set from GEO

We carried out GSEA analysis by pre-processing the gene expression data using R software. The data with differential expressed gene sets representing upper quartile (top 25% of total differentially expressed genes) were then put into the GSEA software along with the chosen gene set collection. The gene set collection was chosen based on the biological question being addressed. We used Molecular Signatures Database (MSigDB) collection of predefined gene sets, including canonical pathways, gene ontology (GO) terms, and transcription factor targets. The GSEA analysis result shows enrichment score (ES) for each gene set. The output of the analysis shows a ranked list of genes based on their correlation with the disease phenotype. The enrichment plot provides a graphical view of the enrichment score for a gene set.

3.12.3 Visualization and interpretation of pathway enrichment analysis using Cytoscape EnrichmentMap tool

The pathway enrichment analysis using Cytoscape EnrichmentMap consists of four main steps. First, we carried out pre-processing of gene expression data using Deseq2 R package. This involves normalizing the data, filtering out low-expressing genes, and performing log₂ transformation. Second, we selected Molecular Signatures Database (MSigDB) collection of predefined gene sets, including canonical pathways, gene ontology (GO) terms, and transcription factor targets for further analysis.

Third, we carried out Cytoscape analysis to identify enriched gene sets and biological processes. We created co-expression network based on the Pearson correlation coefficient, with the threshold for edge creation set to a correlation coefficient of 0.4. Further, we carried out EnrichmentMap analysis to identify enriched gene sets and organize them into a network. The analysis involves identifying gene sets that are significantly enriched based on a p-value cut-off of 0.05 and a false discovery rate (FDR) cut-off of 0.25 and an overlap coefficient cut-off of 0.5. The enriched gene sets are then visualized as a network using the EnrichmentMap plugin.

Finally, we interpreted the result by identifying the most significantly enriched gene sets and biological processes based on a predefined FDR threshold of less than 0.05. The top enriched gene sets and biological processes are then further analysed for biological relevance using gene annotations and functional information provided by the gene set collection.

3.11.5. DAVID Bioinformatics online tool for pathway analysis and visualisation

DAVID pathway analysis and visualisation involves several steps. First, we pre-processed expression data and gene list (Entrez gene IDs) in TXT format (top 25% of total differentially expressed gens) are uploaded to the DAVID website. Second, we select the functional annotation clustering tool with specific biological functions, pathways, or cellular components. Third, we analysed pathway annotations, based on a modified Fisher exact test with a significance threshold of 0.05. Then the enriched gene ontology terms, and pathways, are clustered based on their similarities. Finally, the results are visualized as a scatterplot or bubble plot, which displays the enrichment score and the p-value for each cluster.

3.12.4 Somatic mutation and Copy number alteration analysis using cBioportal

We carried out somatic mutations and copy number analysis in MMP2, MMP7, and MMP14 using cBioPortal which involves following steps. First, we accessed the cBioportal website

(<http://www.cbioportal.org/>) and selected the TCGA dataset for gallbladder cancer. Second, we selected MMP2, MMP7, and MMP14, genes and retrieved information on their mutations and copy number alterations in the dataset. Finally, we used the cBioportal's visualization tools to analyze the mutations and copy number alterations of these genes and generate various types of plots such as bar plots, and scatterplots.

3.12.5 Study of methylation status in TCGA data

We used TCGA datasets (TCGA-CHOL) to carried out methylation analysis of MMP2, MMP7, and MMP14. First, we accessed the SMART App website (<http://www.bioinformatics.com/smartapp>) and uploaded the TCGA-CHOL dataset. Second, we selected, MMP2, MMP7, and MMP14 genes using the SMART App search function. This will give us a list of DNA methylation probes that are associated with these genes. Finally, we generated DNA methylation plots for each methylation probe, which show the level of DNA methylation between cases and control samples in the dataset.

3.12.6 Analysis of TCGA data using UCSC Xena browser

TCGA data analysis using UCSC Xena browser typically involves following steps. First, we selected the TCGA datasets (TCGA-CHOL) from the UCSC Xena browser website. Second, we select the desired molecular data type, such as gene expression. Third, we selected the desired patient cohort and clinical data, such as tumourtype and stage. Fourth, we downloaded the TCGA data from the UCSC Xena browser website. This can be done by using the UCSC Xena browser's download feature, which allows us to download the selected TCGA dataset as a TSV or TXT file. Finally, we visualized the results of the analysis using the UCSC Xena browser's visualization tools, which includes tools for generating heatmaps, scatterplots, boxplots, and other types of plots.



CHAPTER 4

Results and Discussion



Vinay J
NISER, Bhubaneswar

Objective 1

To find out the association of MMP14 promoter variants with GBC and its molecular mechanism

Findings of this section have been published.

- **Vinay J**, Mishra D, Meher D, Dash S, Besra K, Pattnaik N, Singh SP, Dixit M. Genetic association of MMP14 promoter variants and their functional significance in gallbladder cancer pathogenesis. *J Hum. Genet.* 66, 947–956 (2021)
- **Vinay J**, Ananya Palo, Kusumbati Besra, Manjusha Dixit. Gallbladder cancer associated genetic variants rs1003349 and rs1004030 regulate MMP14 expression by altering SOX10 and MYB binding sites. *Human Molecular Genetics*, 2023

4.1 Introduction

The MMPs are the principal moderators of the tumourmicroenvironment during initiation, progression, and metastasis (11, 12, 321, 322). The presence of MMPs in the tumourmicroenvironment is associated with aggressive tumourbehaviour and poor prognosis (13, 14). Among MMPs, MMP14 has been found to be upregulated in several types of cancer, including breast (17, 36), oral (323), prostate (324), cervical (325), gallbladder (100), lung (19, 37), liver (38), ovary (39), colon, bladder, and gastric cancers (40). MMP14 is primarily localized to invadopodia or the leading edge of the cell membrane, and its expression is essential for invasion through the extracellular collagen basement (326, 327). While the role of MMP14 in tumorigenesis and metastasis is well established in many cancers, its molecular role in GBC remains largely unknown, highlighting the study lacunae.

The MMP14 gene has been found to induce tumorigenesis in transgenic mice, leading to mammary gland abnormalities and adenocarcinoma (35). In addition, MMP14 is known to activate latent TGF β 1 (10) and RANLK (41), which modulate the tumourmicroenvironment through cancer-associated fibroblasts and tumor-associated macrophages, thus promoting tumour initiation (42). Moreover, MMP14 plays a dual role in ECM remodelling by activating MMP2 and MMP13, and by directly cleaving ECM molecules such as gelatin, collagen type I-III, and fibronectin. Also, MMP14 and MMP2 activate TGF- β in a CD44-dependent manner by cleaving LTBP-1 protein, promoting invasion and metastasis (10, 286). While a recent study in Indian population has showed association of SNPs present in MMP2 (rs2285053, rs243865), MMP7 (rs11568818), MMP9 (rs17577), and TIMP2 (rs8179090) with gallbladder cancer (52), but no studies have yet explored the potential involvement of MMP14 in GBC development.

Genetic association studies have revealed that SNPs in MMP14 are associated with various cancers, such as hepatocellular (38), ovarian (39), and oral cancer (291). MMP14 expression is

regulated by two promoter variants, rs1003349 and rs1004030, which aid in binding Sp1 and RR1 (295). In the Taiwanese population, the rs1004030 allele "T" has been identified as a risk allele for both HCC and uterine cervix neoplasia (297). While most of these studies have focused on genetic associations, limited showed the underlying molecular mechanisms. These genetic variants can predispose individuals to altered tumour initiation and metastatic potential. Based on the aforementioned observations, our hypothesis is that the overexpression of MMP14 may contribute to GBC tumorigenesis. This overexpression could be influenced by specific genetic variants within the MMP14 gene. However, no previous study has investigated the role of MMP14 genetic variants in the development of GBC.

Therefore, in the current chapter we conducted a case-control study focusing on this candidate gene to identify variants that may impact expression levels. Specifically, we examined promoter SNPs, namely rs1004030 and rs1003349, which have been shown to affect the binding of transcription factors (27).

In this chapter, we conducted a case-control study to show the genetic association between MMP14 promoter variants and gallbladder cancer, as well as investigate their potential functional role in regulating MMP14 expression. Along with the genetic association study of these MMPs, detailed functional studies were carried out by using various genetic, molecular, and functional assays to understand their role in GBC pathogenesis. However, the study not only helps to understand the transcriptional regulation, but also helps to identify the molecular basis of genetic predisposition.

Results

4.2 Screening of MMP14 promoter SNPs in GBC

We sequenced 50 GBC and 50 control samples as screening sets to identify SNPs in the MMP14 promoter region. We have screened all variants in the 1500 bases upstream of the MMP14 transcription start site. We found four SNPs in the region, out of which variants g. 22836440T>C (rs1004030), g. 22836454G>T (rs1003349), and g. 22836839T>C (rs1042703) showed allelic variations, while g. 22836867 C>T (rs587777039) was found non-polymorphic.

4.3 Discovery set allele and genotype frequency analysis

We analysed all three MMP14 promoter variants, rs1004030, rs1003349 and rs1042703, to understand population specific effect size or odds of disease exposure and minor allele frequencies. As shown in **Table 4.3.1**, all the three polymorphisms were significantly associated with GBC in the discovery set with p value <0.05. These variants were tested for genetic association with GBC in the subsequent validation set. The analysis of allelic and genotype frequencies for each SNP in discovery set is discussed in following sections.

The variant rs1004030 T>C shows significant association with GBC

The allelic and genotypic distribution of the variant between the case and control groups showed 'T' as a major allele and 'C' as a minor allele in our study population (**Table 4.3.1**). rs1004030 T>C showed significant association with GBC (p = 0.0176, OR = 2.15, 95% CI, 1.13-4.08). rs1004030 was further analysed in the validation set.

rs1003349 G>T shows significant association with GBC

The genotype and allele frequencies of gallbladder cancer and the control population were compared, the promoter variant rs1003349 was found to be significantly associated with GBC in discovery set. The minor allele 'T' confers protective effect (p = 0.0439) on GBC with odds ratio of 0.53 (95% CI, 0.29-0.99). The variant rs17098318 was further tested for genetic association in the validation set (**Table 4.3.1**).

The variant rs1042703 T>C shows significant association with GBC

The allelic and genotypic distribution of the variants between the case and control shows 'T' as a major allele and 'C' as a minor allele in our study population. MMP14 promoter variant rs1042703 showed significant association with GBC in discovery set ($p = 0.0407$, OR = 2.19, 95% CI, 1.02-4.69). The variants rs1042703 were tested for genetic association with GBC in the subsequent validation set (Table 4.3.1).

Table 4.3.1 Allelic and genotypic frequency distribution of MMP14 promoter variants in GBC and control subjects (Discovery set)

1	rs1004030	Allele or genotype count (frequency)		P value	OR (95% CI)
		Patients (N = 50)	Control (N = 50)		
Genotypic model	TT ^{ref}	24 (0.48)	33 (0.66)		
	TC	17 (0.34)	14 (0.28)	0.2524	1.67 (0.69-4.03)
	CC	9 (0.18)	3 (0.06)	-	-
Allelic model	T ^{ref}	65 (0.65)	80 (0.80)	0.0176	2.15 (1.13-4.08)
	C	35 (0.35)	20 (0.20)		
2	rs1003349				
Genotypic model	GG ^{ref}	30 (0.60)	20 (0.40)		
	GT	17 (0.34)	24 (0.48)	0.0783	0.47 (0.20-1.09)
	TT	3 (0.06)	6 (0.12)	-	-
Allelic model	G ^{ref}	77 (0.77)	64 (0.64)	0.0439	0.53 (0.29-0.99)
	T	23 (0.23)	36 (0.36)		
3	rs1042703				
Genotypic model	TT ^{ref}	31 (0.62)	39 (0.78)		
	TC	15 (0.30)	10 (0.20)	0.1773	1.89 (0.74-4.78)
	CC	04 (0.08)	1 (0.08)	-	-
Allelic model	T ^{ref}	77 (0.77)	88 (0.88)	0.0407	2.19 (1.02-4.69)
	C	23 (0.23)	12 (0.12)		

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference
P value of Chi-square test with Yates continuity correction

4.4 Analysis of Hardy-Weinberg equilibrium (HWE) and study power

Variants rs1004030, rs1003349, and rs1042703 allelic frequencies were tested for HWE using a significance cut-off of $p < 0.05$. The HWE test was passed by the promoter variations

To find out the association of MMP14 promoter variants with GBC and its molecular mechanism

rs1004030 and rs1003349 with $p = 0.2352$ and 0.0637 , respectively. On the other hand, rs1042703 ($p = 0.0000$) did not satisfy HWE, hence we excluded the SNP from further analysis (Table 4.4.1). The study power was calculated for rs113823671 and rs17098318 using PS statistical power analysis programme. The set parameter to calculate power of the study were, 80% beta power and 0.05 alpha significance level.

Table 4.4.1 Sample size and HWE calculations for MMP14 promoter SNPs

Sl. No.	SNP	Calculated Sample size*		HWE p-value [#]
		Case	control	
1	rs1004030	293	293	0.2352
2	rs1003349	305	305	0.0637
3	rs1042703	313	313	0.0000

* PS statistical power program-based analysis at 80% beta power and 0.05 alpha significance level

[#] HWE was calculated based on actual number of controls samples included in study.

4.5 Demographic profile of study subjects

This study included 314 GBC patients and 323 healthy controls. In Table 4.5.1, the demographic characteristics of the study subjects are listed. The mean age of GBC incidence in our study population was found to be 53.39 years ($SD \pm 11.12$). The gender wise analysis of GBC incidence shows high GBC incidence among females compared to male patients ($P = 0.0001$).

Table 4.5.1 Demographic profile of the gallbladder cancer patients and controls recruited in the study

Characteristics	Cases (n = 314)	Control (n = 323)	P value
Age ^a	53.39 ± 11.12	42.94 ± 11.12	0.0001
Gender^b			
Female	181	127	0.0001
Male	133	196	

^a Student's T-test was used to compare mean values of age

^b Chi-Square test was used to compare the difference in frequency of male and female

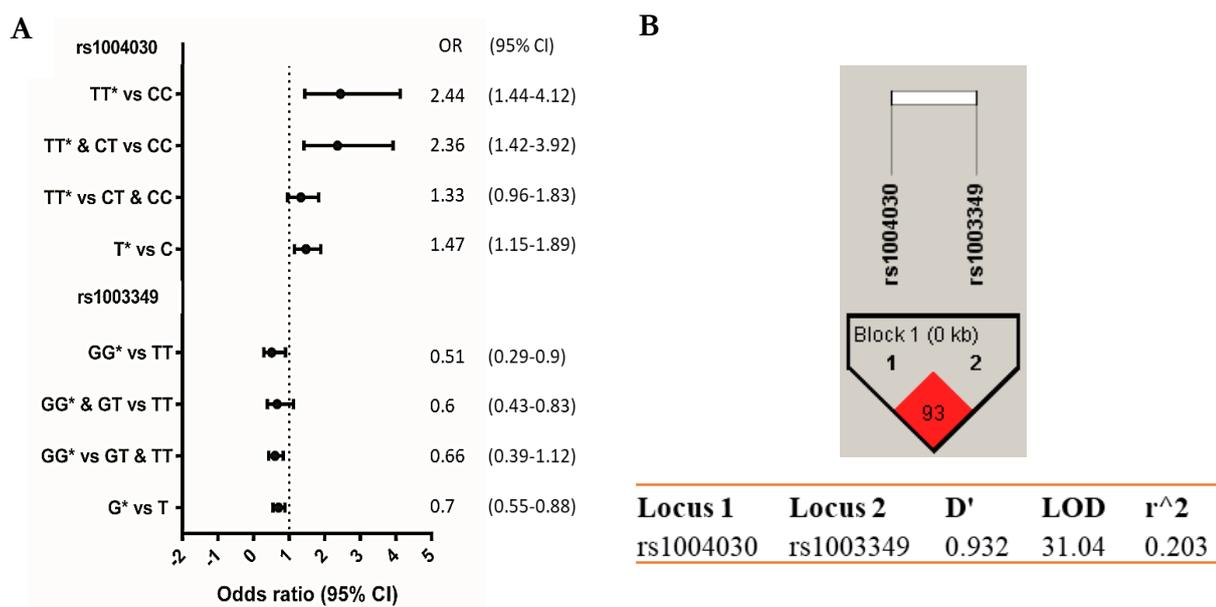


Figure 4.5.1 Gallbladder cancer genetic risk forest plot and linkage disequilibrium plot for MMP14 promoter variants rs1004030 and rs1003349. (A) The odds ratios and 95% confidence intervals of each variants genotype and its individual allelic effect on gallbladder cancer are shown as dark spots and matching horizontal lines. * Reference genotype or allele, odds ratio (OR), and confidence interval (CI). (B) A linkage disequilibrium plot created with the software programme Haploview. Pairwise D' values are used to illustrate linkage disequilibrium (LD). The amplitude and relevance of pairwise LD are shown by shading, with a red-to-white gradient showing higher-to-lower LD values.

4.6 The rs1004030 allele C is associated with an increased risk of gallbladder cancer

Allelic and genotypic distribution of rs1004030 between GBC patients and the control population found allele 'T' as major and 'C' as minor allele in our study population. The frequency of allele 'C' found more in GBC patients than the control population. The association at the allelic level showed that patients with minor allele 'C' increase the risk for gallbladder cancer ($p = 0.0019$, OR = 1.47, 95% CI 1.15-1.89) (Table 4.7.1 and Figure 4.5.1). Allelic interaction analysis revealed that rs1004030 is more likely to follow a recessive model that imparts risk for gallbladder cancer ($p = 0.0006$, OR = 2.36, 95% CI 1.42-3.92). After adjusting for potential confounders such age and gender, rs1004030 exhibited increased association with GBC in allelic ($p = 0.0001$, OR = 1.79, 95% CI 1.35-2.38), recessive ($p = 0.0002$, OR = 3.08, 95% CI 1.72-5.51), and dominant models ($p = 0.0084$, OR = 1.64, 95% CI 1.13–2.37). This suggests that age and gender significantly increased association of variant rs1004030 with GBC.

4.7 The variant rs1003349 is associated with gallbladder cancer

Allelic distribution of rs1003349 in our study population found allele 'T' as minor and 'G' as major allele. A risk analysis at the allelic level showed that the minor allele 'T' at rs1003349 confers a protective effect ($p = 0.0031$), with an odds ratio of 0.7 (95% CI 0.55-0.88) (**Table 4.7.1** and **Figure 4.5.1**). Various genetic models were tested, and it was discovered that the minor allele 'T' at rs1003349 is associated with a lower risk of gallbladder cancer under the dominant model, with a p value of 0.0021 (OR = 0.6, 95% CI 0.43-0.83). The confounding factors were adjusted using binomial logistic regression analysis and found that the allelic ($p = 0.0008$, OR 0.63, 95% CI 0.48-0.82), dominant ($p = 0.0003$, OR 0.5, 95% CI, 0.34-0.73) and recessive models ($p = 0.0266$, OR = 0.502, 95% CI 0.27–0.92) retained its association with a decreased risk for GBC. Adjusting for confounding factors significantly increased variant rs1003349 association with GBC.

Table 4.7.1 Allelic and genotypic frequency distribution of rs1004030 and rs1003349 in GBC and control subjects

	Allele or Genotype	Allele or Genotype count (frequency)		<i>p</i> -value	<i>p</i> -value ^b
		Case (<i>n</i> = 300)	Control (<i>n</i> = 301)		
rs1004030					
Genotypic	TT ^{ref}	145 (0.48)	167 (0.55)		
	TC	102 (0.34)	109 (0.36)	0.6713	0.0001
	CC	53 (0.18)	25 (0.08)	0.0007	0.0022
Allelic	T ^{ref}	392 (0.65)	443 (0.74)	0.0019	0.0001
	C	208 (0.35)	159 (0.26)		
Dominant	TT ^{ref}	145 (0.48)	167 (0.55)	0.0794	0.0084
	CT + CC	155 (0.52)	134 (0.45)		
Recessive	TT + CT ^{ref}	247 (0.82)	276 (0.92)	0.0006	0.0002
	CC	53 (0.18)	25 (0.08)		
Additive	TT vs TC vs CC			0.0026	
rs1003349					
Genotypic	GG ^{ref}	151(0.48)	111 (0.36)		
	GT	137 (0.44)	161 (0.52)	0.0058	0.0023
	TT	26 (0.08)	37 (0.12)	0.0192	0.2450
Allelic	G ^{ref}	439 (0.70)	383 (0.62)	0.0031	0.0008
	T	189 (0.30)	235 (0.38)		
Dominant	GG ^{ref}	151(0.48)	111(0.36)	0.0021	0.0003
	GT + TT	163 (0.52)	198 (0.64)		

Recessive	GG + GT ^{ref}	288 (0.92)	272 (0.88)	0.1262	0.0266
	TT	26 (0.08)	37 (0.12)		
Additive	GG vs GT vs TT			0.007	

n Sample size

^{ref} Reference allele

^b Adjusted *p* values for confounding factors like age and gender

* *p* values indicate significance after Bonferroni correction.

4.8 Linkage disequilibrium and haplotype analysis of MMP14 promoter variants

Maximum likelihood estimates were used to analyse the degree of LD and haplotype association (based on polymorphisms present in the MMP14 promoter sequence). The frequency distribution of the haplotype revealed haplotype T-G (rs1004030 and rs1003349, respectively) as the most frequent haplotype (**Table 4.8.1**). We found no significant association of GBC with any of the haplotypes. Also, the correlation coefficient r^2 between rs1004030 and rs1003349 was 0.203 which suggests weak linkage disequilibrium, suggesting independent effects on the alteration of risk for GBC.

Table 4.8.1 Haplotype association of the variants rs1004030 and rs1003349 with gallbladder cancer

Haplotype	Frequencies		OR (95% CI)	<i>p</i>-value
	Case (314)	Control (323)		
T-G ^{ref}	0.36 (113)	0.36 (116)	1	-
T-T	0.29 (92)	0.38 (121)	0.78 (0.54-1.14)	0.1949
C-G	0.34 (107)	0.26 (85)	1.30 (0.88-1.90)	0.1910

^{ref} Reference haplotype

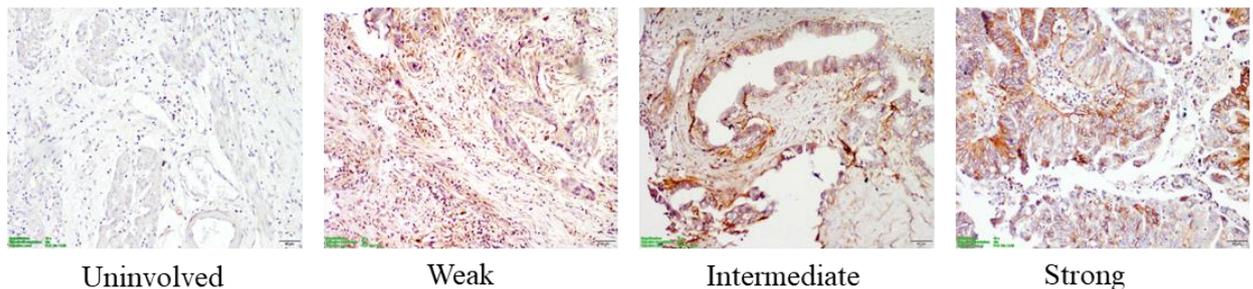
4.9 MMP14 expression in gallbladder cancer tissue is between the genotypes of variant rs1004030 and rs1003349

MMP14 expression was studied using immunohistochemistry in 26 gallbladder cancer tissues and adjacent normal tissue. MMP14 was predominantly expressed in glandular epithelial cells of tumour tissue (**Figure 4.9.1A**). The expression of MMP14 was significantly higher in the gallbladder tumour tissue (Median Allred score 6) compared to normal tissue (Median Allred score 4, $p = 0.0001$) (**Figure 4.9.1B**). This result was also supported by TCGA data (TCGA-CHOL) (tcga-data.nci.nih.gov/tcga, accessed on 08-04-2020), showing a significant increase in MMP14 expression in tumour tissue ($p = 0.0001$) (**Figure 4.9.1C**).

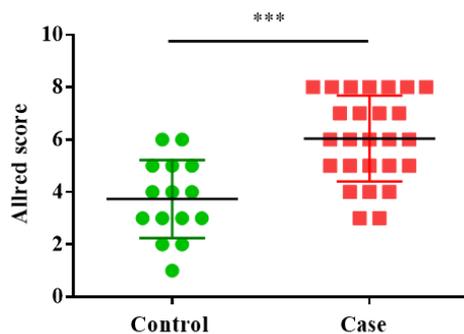
Next, we analysed the MMP14 expression levels in GBC tissues samples having different genotypes. At rs1004030, patients with genotype 'CC' (mean Allred score 8.0 ± 0.0) had significantly ($p = 0.0001$) higher MMP14 expression compared to patients with genotype 'TT' (mean Allred score 3.6 ± 0.25). Similar to this, MMP14 expression was significantly ($p = 0.0001$) higher among those with the 'CT' genotype (mean Allred score 6.29 ± 0.27) than with the 'CC' genotype (**Figure 4.9.2A**). Altogether, we found a clear effect of 'C' allele copy number on MMP14 expression, with each copy of allele 'C' increasing its expression.

Additionally, patients with the genotype 'GG' (mean Allred score 6.47 ± 0.47) at locus rs1003349 showed significantly ($p = 0.0178$) higher MMP14 expression than carriers of the 'TT' genotype (mean Allred score 5.17 ± 0.17). However, there was no substantial difference between the expression levels of carriers of the 'GT' (mean Allred score 5.80 ± 0.80) and 'TT' genotypes ($p = 0.4783$). (**Figure 4.9.2B**). Overall, these findings support the involvement of MMP14 promoter variations in the regulation of MMP14 gene expression.

A



B



C

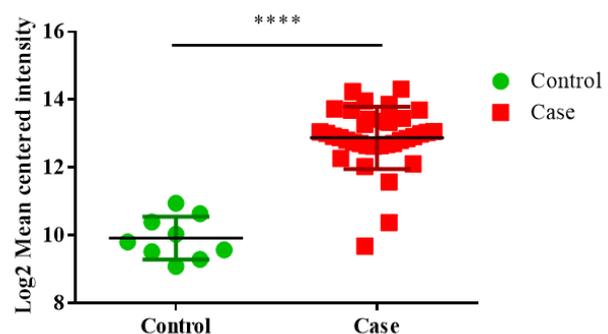


Figure 4.9.1 The MMP14 protein expression in gallbladder cancer and surrounding normal tissue. (A) Representative images of immunohistochemical staining in gallbladder cancer tissue, ranging from weak or negative to intermediate and strong. The scale bar indicates 50 m with a 10X objective and 10X magnification. (B) Immunohistochemistry scoring of MMP14 in samples of gallbladder tumour tissue and nearby normal tissue; the X axis

indicates the sample type, case, and control. Y axis displays Allred score. The unpaired non-parametric two-sided Mann Whitney test was used to analyse expression difference. (C) MMP14 RNA-seq expression data from the TCGA database (TCGA-CHOL), with the X axis representing sample type, GBC samples, and normal samples, and the Y axis representing Log2 mean intensity value. Unpaired non-parametric two-sided Mann Whitney test was used for the statistical test.

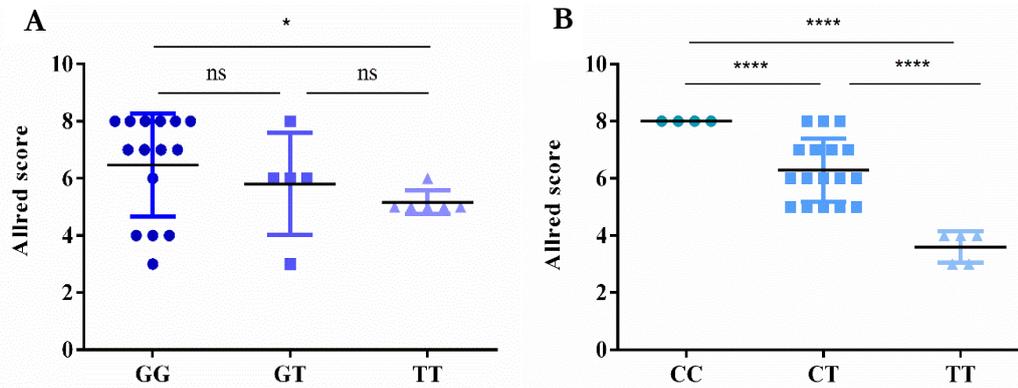


Figure 4.9.2 MMP14 protein expression levels in gallbladder cancer patients with different genotypes of variants rs1004030 and rs1003349. (A) Analysis of MMP14 expression between the rs1004030 genotypes groups; rs1004030 genotypes shown on the X axis and corresponding Allred scores on the Y axis. (B) Similarly, for rs1003349; genotype is shown on the X axis and Allred score is shown on the Y axis. Two-sided unpaired student's t-test with Welch's correction was used for the statistical test.

4.10 Risk variants rs1004030 and rs1003349 may elevate MMP14 expression

We analysed the genomic region around these variants in the UCSC genome browser containing Encyclopaedia of DNA Elements (ENCODE) data to determine the regulatory role of loci consisting of the SNPs rs1004030 and rs1003349. Both variations are present in the cis-Regulatory Element (CRE), which has significant DNase and H3K4me3 signals, indicating the presence of an active promoter in the region (**Figure 4.10.1A**). In GBC, we hypothesized that mutant allele would change how the gene is regulated by forming or destroying binding sites for repressor or enhancer elements.

We hypothesized that these SNPs might modify the binding sites of transcription factors, thereby impacting transcript levels. To confirm this in GBC cell lines, we cloned the promoter

region surrounding rs1004030 and rs1003349 into the luciferase minimal promoter vector pGL4.23 and transfected it into GBC lines G415 and TGBC1TKB. In both cell lines, the pGL4.23_MMP14_promoter fragment (containing the wild type allele) demonstrated a significant increase in luciferase activity compared to the minimal promoter vector pGL4.23, highlighting the regulatory role of this region (**Figure 4.10.1B** and **C**). To further validate the impact of individual SNPs on transcription, we conducted luciferase assays with combinations of risk alleles. In the TGBC1TKB cell line, the combination of 'C and G' risk alleles showed a significant change in luciferase activity compared to the combination of single risk alleles 'T and G' at rs1004030 and rs1003349, respectively (**Figure 4.10.1B**). However, in the G415 cell line, although the combinations of alleles 'C and T' and 'T and G' showed a significant difference in luciferase activity compared to 'T and T', there was no additive effect of the two risk alleles (C and G) on luciferase activity (**Figure 4.10.1C**). These findings suggest that the risk allele at rs1004030 may have a greater impact on MMP14 expression compared to rs1003349, but this effect is specific to the cell line. We further explored the potential role of putative transcription factors that may affect MMP14 expression in an allele-dependent manner.

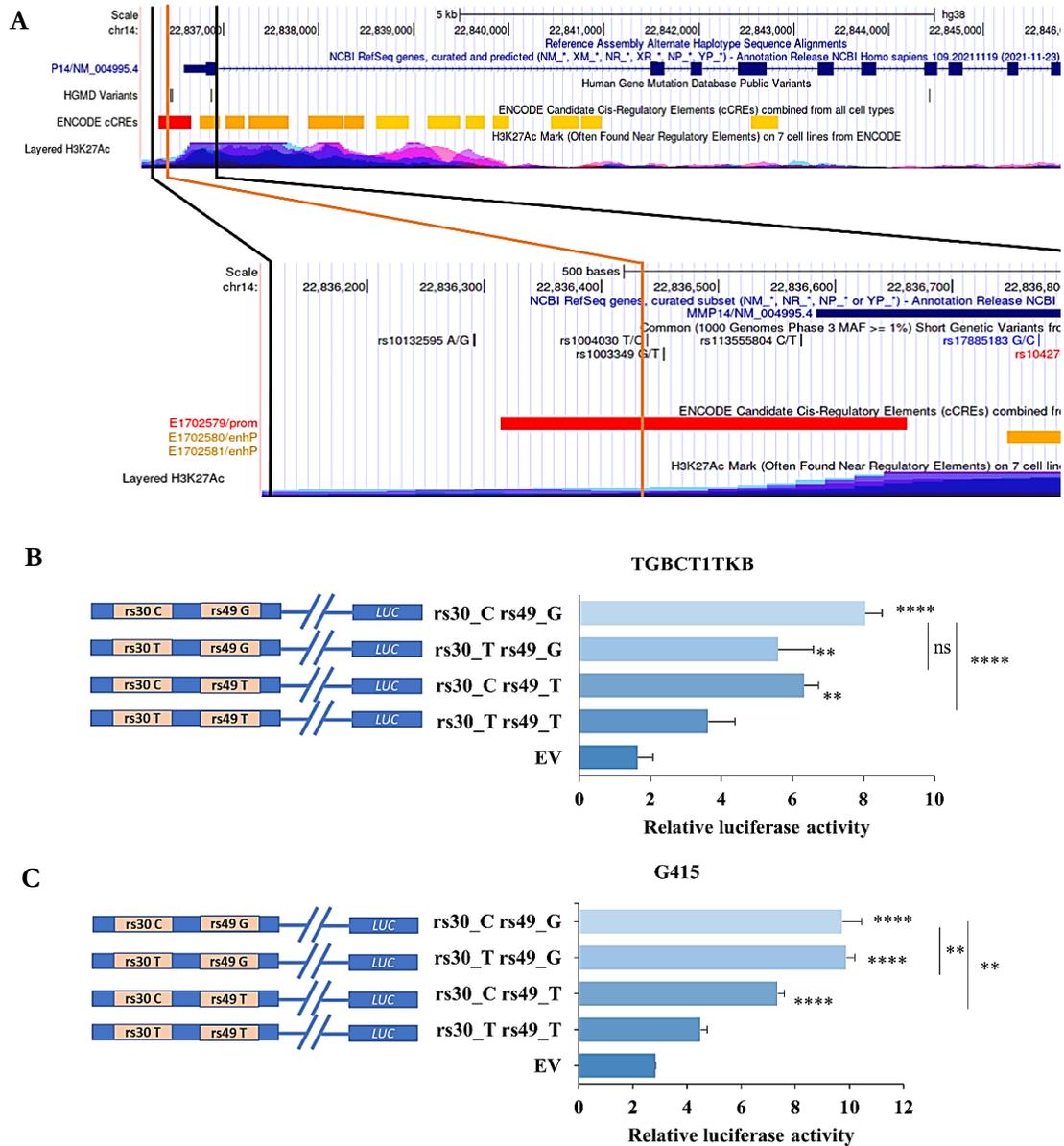


Figure 4.10.1 Evaluation of the functional potential of promoter variants rs1004030 and rs1003349. The allelic effect of promoter variants rs1004030 and rs1003349 on reporter luciferase activity (firefly/renilla) was assessed by cloning the 63 bp genomic region around the variations into the pGL4.23 vector. (A) ENCODE data indicating the presence of variants rs1004030 and rs1003349 inside cis-Regulatory Element (CRE) with high DNase and H3K4me3 area (indicated with red band). (B) TGBC1TKB cells exhibit a statistically significant increase in luciferase activity when the risk allele, 'C' / 'G', is present (8.36 ± 0.43 compared to the wild-type allele, 'T' / 'T', (3.51 ± 0.76 ; $P = 4.3E-06$)). (C) The G415 cells exhibit a significantly higher level of luciferase activity when the risk allele combination 'C'/'G' (9.73 ± 0.72) is compared to the wild type alleles 'T' / 'T' (4.49 ± 0.25) ($P = 3.55E-12$). The experiments were carried out in triplicate, and the results are shown as mean \pm SD. The student's *t*-test was used to determine the statistical significance of the difference between groups: ns $P > 0.05$, ** $P < 0.01$, **** $P < 0.0001$; ns, not significant.

4.11 The risk alleles 'G' (rs1004030) and 'C' (rs1003349) provide binding sites for transcription factors SOX10 and MYB

We conducted in-silico analysis of the promoter region surrounding rs1004030 and rs1003349 SNPs using the JASPAR transcription factor database to identify potential transcription factors. The results revealed that the alleles 'G' and 'C' at rs1003349 and rs1004030, respectively, are potential binding sites for the transcription factors MYB and SOX10, but not with the allele 'T'. To validate these findings, we performed EMSAs using nuclear extracts from TGBC1TK cells and 39 bp oligonucleotides representing the genomic region encompassing rs1003349 and rs1004030 with alleles 'G' and 'C', respectively. The results showed that a specific protein complex bound to the labelled oligonucleotide with allele 'C' at rs1004030, as was evident by the band shift observed in Lane 2, **Figure 4.11.1A**, whereas the labelled oligos with allele 'T' did not show any binding (Lane 5, **Figure 4.11.1A**). This binding was specific to allele 'C' as the addition of unlabelled oligonucleotides containing allele 'C' abolished the shift (Lane 3, **Figure 4.11.1A**). Furthermore, the presence of a super shift in the protein-DNA complex mobility in the presence of a SOX10-specific antibody confirmed the specificity of the protein binding to allele 'C' at rs1004030 (Lane 4, **Figure 4.11.1A**).

Similarly, the EMSA results for rs1003349 showed a shift in the electrophoretic mobility of the labeled oligonucleotide with allele 'G' (Lane 2, **Figure 4.11.1C**), indicating binding of a protein to this allele. The specific binding to allele 'G' was confirmed by the disappearance of the shift upon addition of unlabelled oligonucleotides containing allele 'G' (Lane 3, **Figure 4.11.1C**). Moreover, the presence of a supershift in the protein-DNA complex mobility in the presence of a MYB-specific antibody provided further evidence that the bound protein complex at rs1003349 includes the MYB transcription factor (Lane 4, **Figure 4.11.1C**). These findings confirm that the alleles 'C' and 'G' at loci rs1004030 and rs1003349 facilitate the binding of transcription factors SOX10 and MYB, respectively.

To further validate the allele-specific binding of SOX10 and MYB to rs1004030 and rs1003349, respectively, chromatin immunoprecipitation (ChIP) assays were performed in G415 cells. The results showed a significant enrichment of the genomic region surrounding rs1004030 and rs1003349 following immunoprecipitation with SOX10-specific and MYB-specific antibodies, respectively, compared to the IgG control (**Figure 4.11.1B and D**). This enrichment was further increased with the overexpression of SOX10 and MYB, confirming the specificity of the

binding. Additionally, the enrichment of the MYB binding site was more prominent compared to the SOX10 binding site. Overall, these findings provide *in vivo* evidence of the direct binding of SOX10 and MYB to loci rs1004030 and rs1003349, respectively, and suggest their putative role in the expression of MMP14.

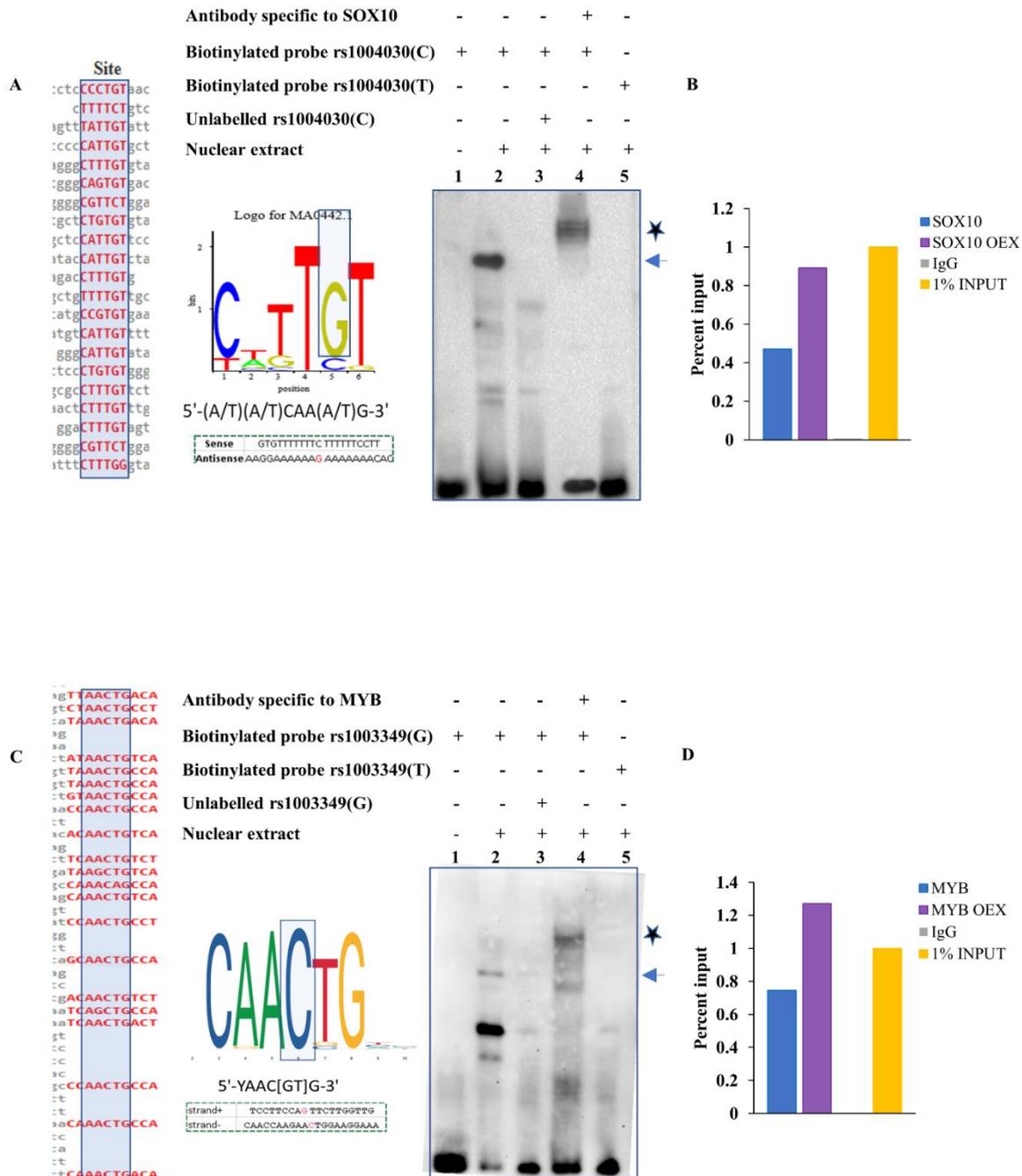


Figure 4.11.1 The risk alleles 'C' (rs1004030) and 'G' (rs1003349) provide binding sites for transcription factors SOX10 and MYB. EMSA and ChIP assays showing the risk alleles 'C' (rs1004030) and 'G' (rs1003349) provide binding sites for transcription factors SOX10 and MYB. (A) The biotinylated "C" allele leads to a shift in mobility, as shown by the arrowhead

(protein-DNA complex) (lane 2). There is no shift at rs1004030 with the 'T' allele (lane 5). In Lane 3, labeled and unlabelled probes (allele "C") compete with one another for protein binding, which results in a shift disappearance that indicates an allele-specific protein-DNA interaction. Supershift (marked with a star, lane 4) is produced when this complex is incubated with SOX10-specific antibodies, demonstrating the SOX10 protein's allele-specific binding. (B) A bar graph for the ChIP assay analysis of the genomic region encompassing rs1004030 indicates chromatin enrichment for the SOX10 binding site in wild-type G415 cells, G415 cells pulled with IgG as a negative control, G415 cells with SOX10 overexpression, and G415 cells pulled with 1% input. (C) A biotinylated probe with the allele "G" at rs1003049 exhibits a mobility shift (a protein-DNA complex), which is indicated by an arrowhead (lane 2), whereas the labeled probe with the allele "T" exhibits no shift (lane 5), showing a specific interaction of the protein-DNA complex with allele "G." When competing with unlabeled probes carrying the allele "G," the shift was abolished (lane 3). After incubating this protein-DNA complex with an anti-MYB antibody, the supershift was seen at lane 4. Indicating allele-specific binding of transcription factor MYB. (D) A bar graph for the ChIP assay analysis of the genomic region at rs1003349 reveals chromatin enrichment for the MYB binding site in wild-type G415 cells, G415 cells with MYB overexpression, G415 cells pulled with negative IgG control, and in 1% input.

4.12 Perturbation of SOX10 and MYB expression alters MMP14 expression

Based on our knowledge of the active promoter signature and transcription factor binding in the regulatory region, we investigated the direct regulation of MMP14 expression by SOX10 and MYB. To examine this, we perturbed the expression of SOX10 and MYB and examined the resulting changes in MMP14 expression levels. We conducted these experiments using G415 and HEK293T cell lines, as they possess binding sites for SOX10 and MYB, unlike the TGBC1TKB cell line (**Figure 4.12.2**).

We observed that stable knockdown of SOX10 and MYB significantly decreased MMP14 expression in both HEK293T and G415 cells (**Figure 4.12.1A and B**). Conversely, ectopic expression of SOX10 and MYB in HEK293T and G415 cells led to a significant increase in MMP14 expression (**Figure 4.12.1C and D**). These findings provide further evidence of the regulatory role of SOX10 and MYB expression levels in controlling MMP14 expression.

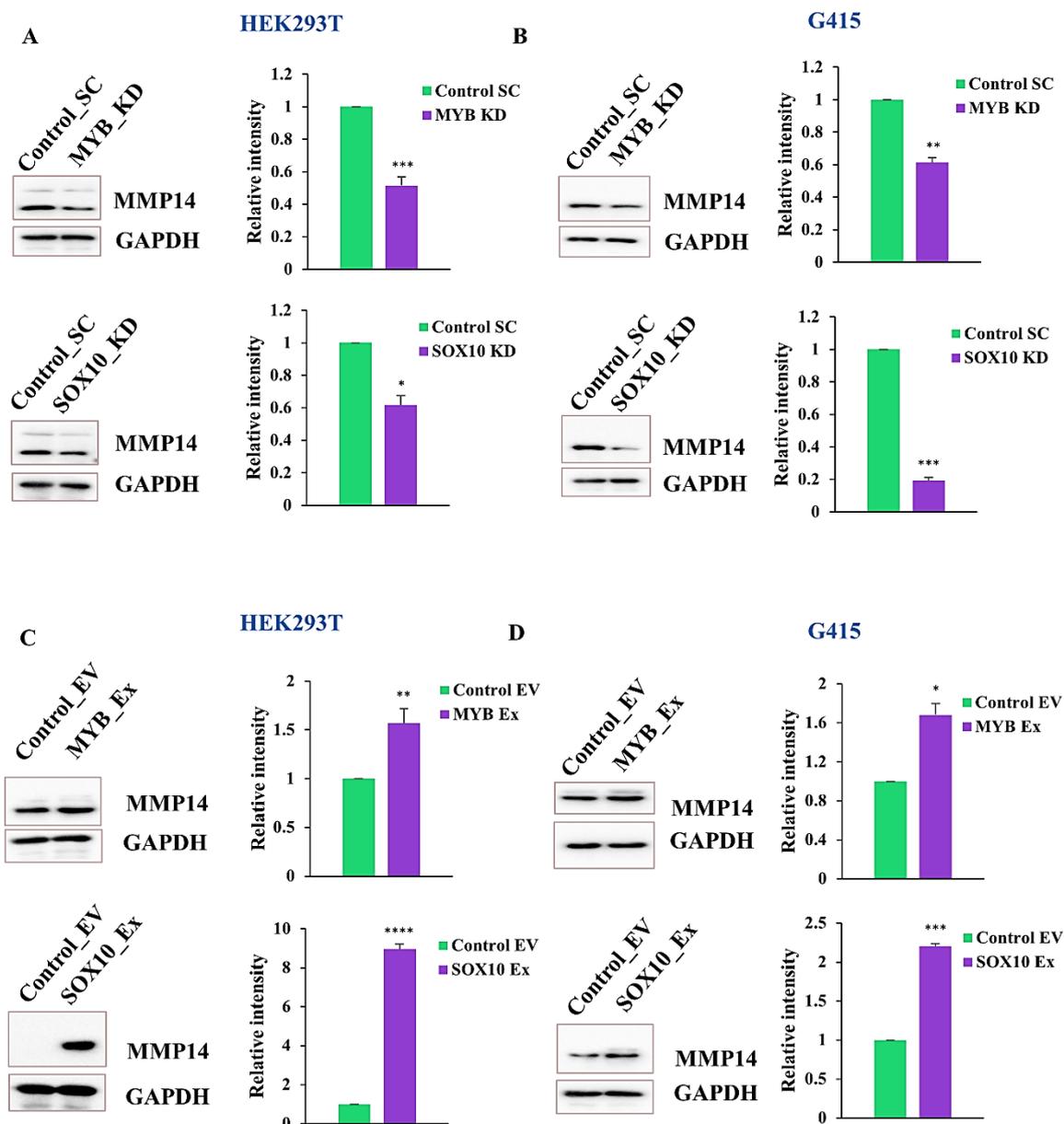


Figure 4.12.1 MYB and SOX10 expression levels regulate MMP14 expression. *In HEK293T and G415 cells, MYB and SOX10 were knocked down and ectopically expressed, and the effect on MMP14 expression was determined using a Western blot. (A and B) MMP14 expression is significantly reduced in the upper panel when MYB (MYB KD) is knocked down in (A) HEK293T ($P = 0.0002$) and (B) G415 ($P = 0.0054$) cells compared to control (SC). The expression of MMP14 is significantly reduced in the lower panel's HEK293T ($P = 0.0219$) and G415 ($P = 0.0006$) cells with SOX10 knockdown (SOX10 KD) compared to control (SC). (C and D) In the upper panel, the MMP14 level is significantly higher in the ectopic MYB (MYB Ex)-expressed cells of (C) HEK293T ($P = 0.0052$) and (D) G415 ($P = 0.0283$) than in control (EV) cells. In the lower panel, cells with elevated SOX10 expression (SOX10 Ex) ($P = 3.23E-05$) and G415 ($P = 0.0007$) show significantly higher MMP14 expression than controls (EV). The experiments were carried out in triplicate, and the results are shown as mean SD. The statistical significance of the difference between the groups was determined using the student's *t*-test, with * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$.*

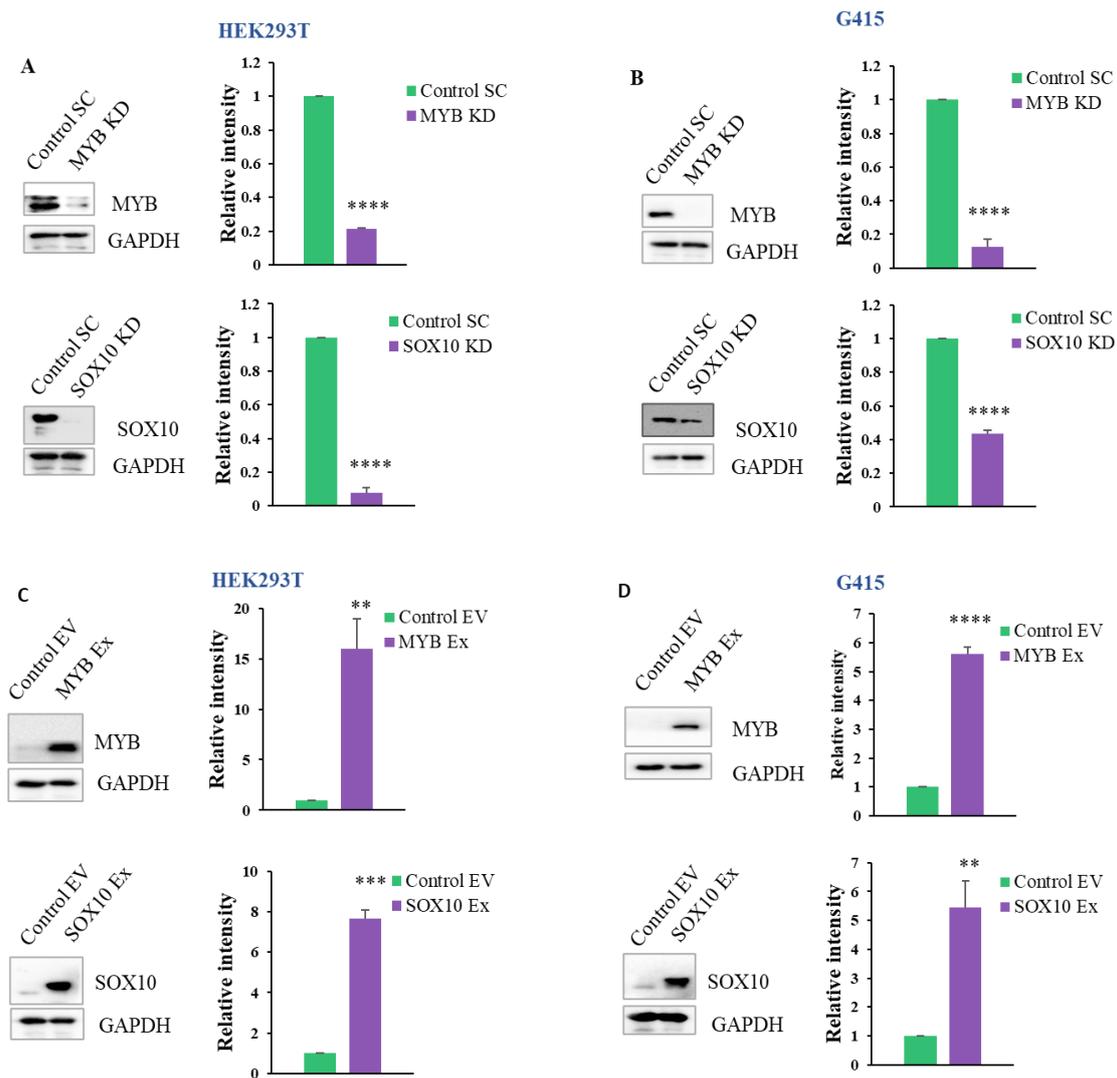


Figure 4.12.2 MYB and SOX10 expression perturbation in the HEK293T and G415 cell lines. Western blot analysis of two GBC cells reveals MYB and SOX10 perturbation. (A and B) In HEK293T cells, stable knockdown results in lower MYB and SOX10 expression levels in the MYB KD ($p = 1.32E-06$) and SOX10 KD ($p = 2.16E-06$) groups compared to the control SC group. Similarly, when compared to the control SC group, G415 cells exhibit decreased expression in the MYB KD ($p = 1.43E-05$) and SOX10 KD ($p = 2.49E-06$) groups. (C and D) In HEK293T cells, transient expression reveals higher levels of MYB ($p = 0.0061$) and SOX10 ($p = 0.0003$) expression in the MYB Ex and SOX10 Ex groups compared to the control EV group. Similarly, G415 cells exhibit increased expression in the MYB Ex ($p = 1.24E-05$) and SOX10 Ex ($p = 0.0078$) groups when compared to the control EV group. The values are shown as mean standard deviation. Statistics were calculated between the groups using the student's *t*-test. **, $P \leq 0.01$, ***, $P \leq 0.001$, ****, $P \leq 0.0001$.

4.13 Knockout of rs1004030 and rs1003349 containing promoter region reduces MMP14 expression and abolishes the effect of SOX10 and MYB on MMP14 expression

We tested the hypothesis that if SOX10 and MYB bind to the loci at rs1004030 and rs1003349, the elimination of this region will not show any effect of SOX10 and MYB level perturbation on the MMP14 expression. Using the CRISPR/Cas9 system, we generated a knockout cell line (G415119^{-/-}) with a 119 bp deletion encompassing the loci rs1004030 and rs1003349 (**Figure 4.13.1A and B**). The colonies derived from single cells with homozygous deletion exhibited minimal expression of MMP14 protein (**Figure 4.13.1C**). When we transiently overexpressed the transcription factor SOX10 in the G415119^{-/-} cells and the corresponding control cells, we observed retaining of reduced MMP14 expression levels in the knockout cells (G415119^{-/-}), but not in the control cells (**Figure 4.13.1D**). Likewise, the transient overexpression of MYB did not lead to an increase in MMP14 protein expression in G415119^{-/-} cells compared to the G415 non-deleted cells (**Figure 4.13.1E**). These findings further support the idea that variants within the cis-regulatory active promoter region facilitate the binding of transcription factors SOX10 and MYB to regulate the expression of MMP14.

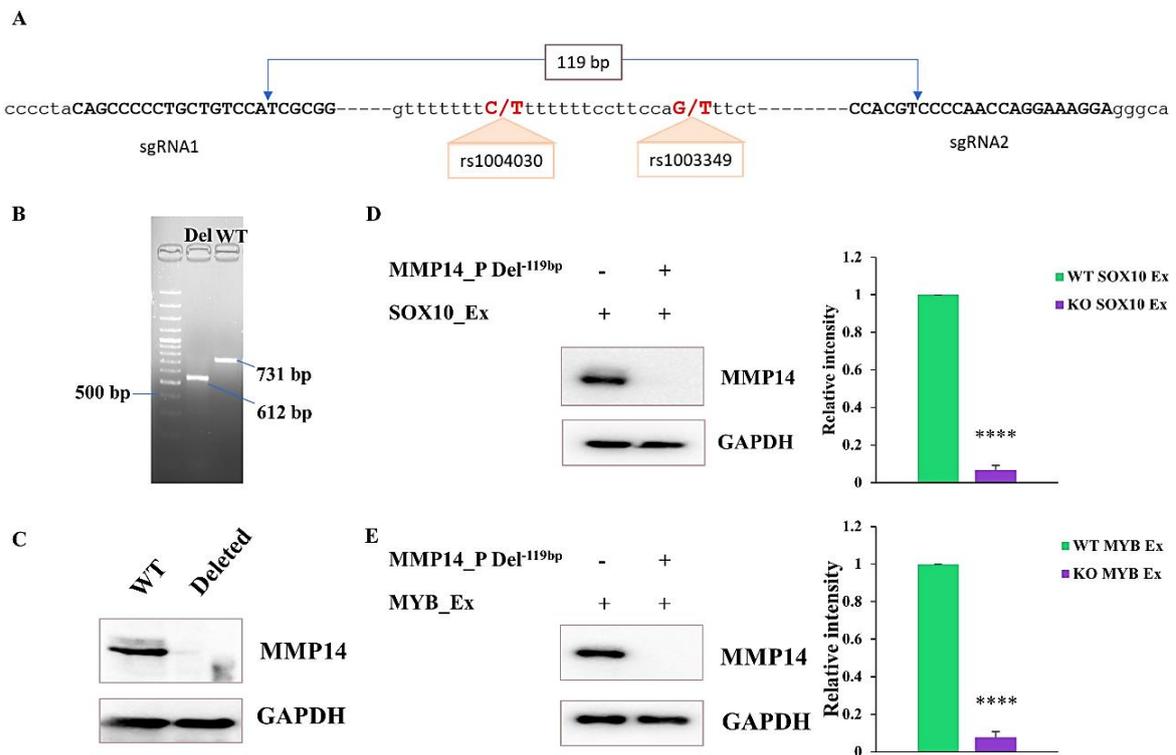


Figure 4.13.1 The effect of SOX10 and MYB on MMP14 expression in promoter knockout cells. CRISPR/Cas9-mediated deletion of 119 bp of genomic sequence around the variants rs1004030 and rs1003349 was performed to validate SOX10 and MYB binding sites. (A) A graphic representation of the genomic region displays the locations of SNPs (allelic variation is in bold), the binding positions of sgRNA1 and sgRNA2 (sequence in bold), and the arrowheads pointing to the CRISPR/Cas9 cleavage site. (B) The agarose gel picture depicts the target region being amplified by PCR in CRISPR/Cas9 deleted (del) and wild-type (WT) cells. (C) Western blot data showing MMP14 expression in wild-type (WT) and homozygous promoter fragment-deleted cells (MMP14 P Del-119bp). (D and E) In cells with homozygous deletion of the promoter segment (MMP14 P Del-119bp) and wild-type cells, the Western blots and accompanying densitometric graphs demonstrate MMP14 expression after perturbing the expression of SOX10 and MYB. The experiments were carried out in triplicate, and the results are shown as mean SD. The statistical significance of the difference between the groups was determined using the student's t-test, **** P 0.0001.

4.14 SOX10 and MYB expression levels have an allele-specific effect on reporter gene expression

To further confirm the specificity of the binding of SOX10 and MYB in an allele-specific manner and its impact on transcription regulation, we conducted a rescue experiment by

restoring MYB/SOX10 expression and evaluating the effect on reporter gene expression. HEK293T cells were transfected with modified MYB/SOX10 expression and luciferase constructs containing either the risk or wild-type alleles at SOX10 and MYB binding sites. In stable MYB knockdown HEK293T cells, there was no significant difference in luciferase activity between the 'G' and 'T' alleles of rs1003349. However, when MYB was transiently overexpressed in the HEK293T MYB knockdown cells with the 'G' allele, there was a significant increase in luciferase activity compared to the 'C' allele of rs1003349 (**Figure 4.14.1A**).

Similarly, we validated the binding of SOX10 by transfecting luciferase constructs into HEK293T cells with SOX10 knockdown. There was no significant change in luciferase activity between the risk allele 'C' and wild-type allele 'T' of rs1004030. However, when SOX10 was overexpressed, there was a significant increase in luciferase activity between the two groups (**Figure 4.14.1B**). These findings indicate that SOX10 and MYB bind to the MMP14 promoter in an allele-specific manner and regulate its expression.

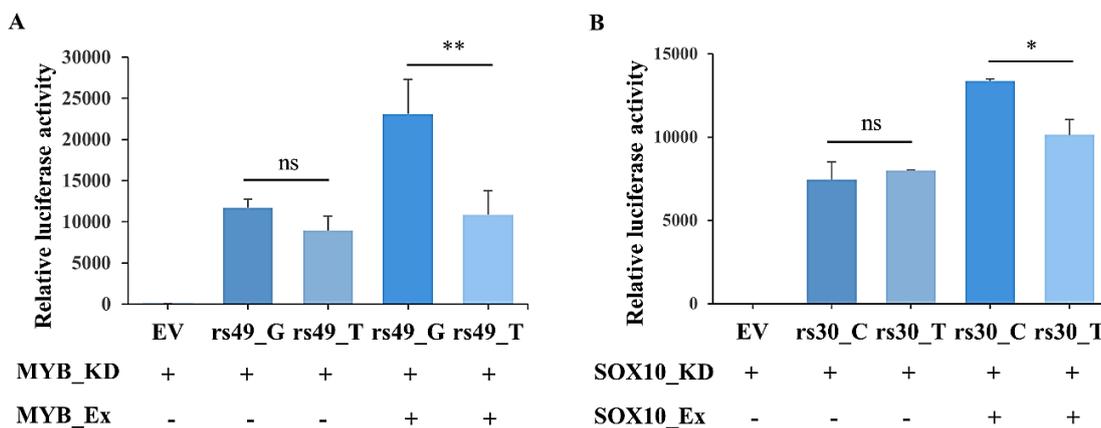


Figure 4.14.1 SOX10 and MYB expression levels have an allele-specific effect on reporter gene expression. Luciferase assay was used to further confirm SOX10 and MYB's allele-specific binding to the MMP14 promoter. Using allele-specific pGL4.23 constructs bearing a 63 bp genomic area surrounding the variants rs1004030 and rs1003349, luciferase activity was measured in HEK293T cells with SOX10 and MYB depletion (KD) and with the expression of SOX10 and MYB rescued. (A) In HEK293T cells with MYB depletion (MYB KD), there was no

visible difference in the relative luciferase activity (firefly/renilla) between the G allele (rs49 G) and T allele (rs49 T) groups. The allele 'G' group (231144189) of the MYB expression rescue (MYB Ex) exhibits significantly higher relative luciferase activity as compared to the allele 'T' group (108592921) ($P = 0.0023$). (B) In HEK293T cells with SOX10 deletion (SOX10 KD), there was no significant difference in the relative luciferase activity (firefly/renilla) between the G allele (rs30 G) and T allele (rs30 T) groups. In HEK293T SOX10 KD cells, SOX10 expression rescue (SOX10 Ex) revealed a significant increase in relative luciferase activity (firefly/renilla) with risk allele 'G' (13368129) compared to allele 'T' (10136928) ($P = 0.0517$). The experiments were carried out in triplicate, and the results are shown as mean SD. The statistical significance of the difference between the groups was determined using the Student's *t*-test, with $nsP > 0.05$, $*P 0.05$, $**P 0.01$; ns, not significant.

4.15 MMP14 increases tumorigenic properties of GBC cell lines

To assess the impact of increased MMP14 expression on tumorigenic properties, we generated GBC cell lines with stable overexpression of MMP14 and conducted cell-based assays (**Figure 4.15.2**). We investigated the effect on cell proliferation using MTS and colony formation assays. The results showed that elevated MMP14 expression significantly enhanced the proliferation ability of G415 and TGBC1TKB cells (**Figure 4.15.1A** and **B**). Likewise, the colony-forming ability of G415 and TGBC1TKB cell lines increased with MMP14 overexpression (**Figure 4.15.1C** and **D**). To evaluate the effect on the metastatic potential of cells in terms of migratory property, we performed wound healing and transwell invasion assays using Boyden chambers. The findings revealed a significant increase in wound healing for both the cell lines with ectopic MMP14 expression compared to the control groups (**Figure 4.15.1E** and **F**). Similarly, MMP14 overexpression led to a significant increase in transwell invasion for both the cell lines (**Figure 4.15.1G** and **H**). Overall, these findings suggest that MMP14 expression plays a role in promoting tumorigenic properties in G415 and TGBC1TKB cells.

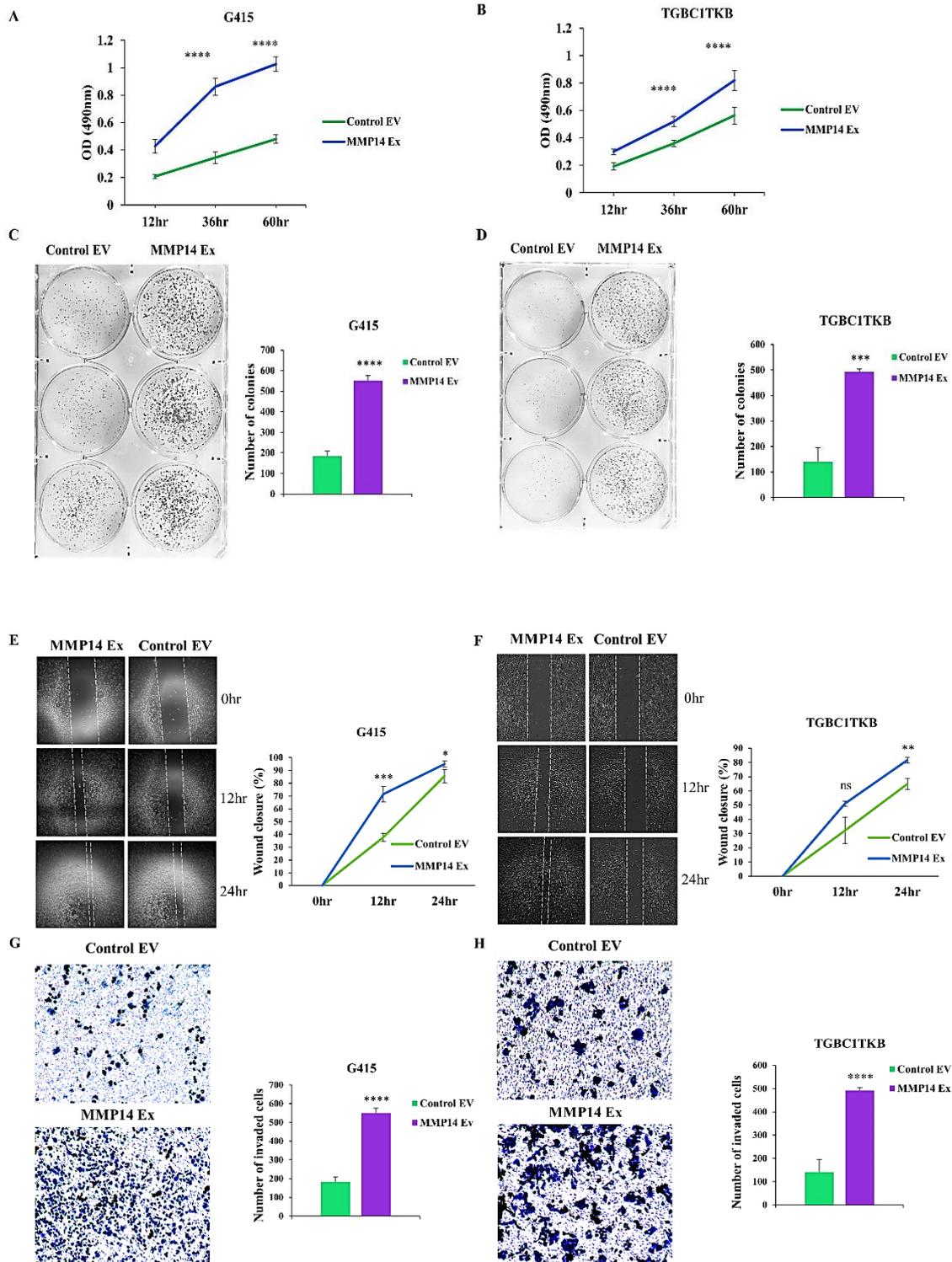


Figure 4.15.1 MMP14 expression affects the tumorigenic characteristics of GBC cells. In two GBC cell lines, G415 and TGBC1TKB, cell-based assays were carried out to study the effect of MMP14 on cellular properties. (A and B) The MTS proliferation assay in G415 ($P = 8.05E-10$) and TGBC1TKB ($P = 6.15E-05$) cells reveals that the MMP14 ectopic expression (Ex) group proliferated more than the control (EV) group. (C and D) The colony formation assay using the cells G415 ($P = 0.0001$) and TGBC1TKB ($P = 0.0009$) showed that the ability

to form colonies was considerably higher in the MMP14 ectopic expression (Ex) group of cells than in the control (EV) group. (E and F) The wound healing assay in G415 ($P = 0.0290$) and TGBC1TKB ($P = 0.0093$) cells shows fast healing at 24 hr when compared to the control EV group, while MMP14 Ex cells show a significant increase in wound healing when compared to the control EV group. (G and H) In the transwell invasion assay, MMP14 ectopic expression (Ex) cells invade more readily than control (EV) cells in G415 ($P = 3.33E-05$) and TGBC1TKB ($P = 1.93E-05$) cells. The experiments were carried out in triplicate, and the results are shown as mean SD. The statistical significance of the difference between the groups was determined using the student's *t*-test, with * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$.

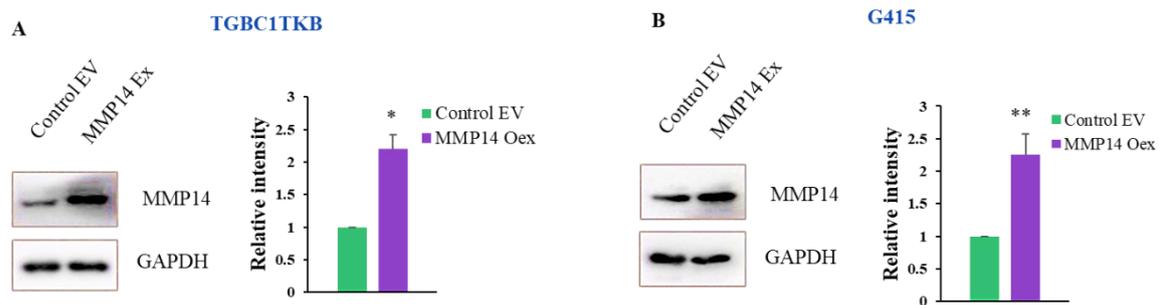


Figure 4.15.2 Stable expression of MMP14 in TGBC1TKB and G415 cell lines. Western blot analysis showing expression levels of MMP14 in two GBC cell lines. (A and B) MMP14 Ex (expression vector for MMP14) group exhibits higher MMP14 expression compared to Control EV (Empty vector) group in TGBC1TKB ($p = 0.0264$) and G415 ($p = 0.0077$) cells. The values are shown as mean standard deviation. Statistics were calculated between the groups using the student's *t*-test. *, $P \leq 0.05$, **, $P \leq 0.01$.

4.16 MMP14 expression in gallbladder cancer tissues samples

Immunohistochemistry was conducted to examine MMP14 expression in tumour tissue and adjacent normal tissue from 29 GBC patients. The results showed high MMP14 expression in tumour tissue compared to adjacent normal tissue (Figure 4.16.1A). Notably, MMP14 expression gradually increased with higher tumour grades, highlighting the significance of MMP14 in advanced gallbladder cancer (Figure 4.16.1B and D). Tumour or tumour stroma tissue exhibited higher or moderate MMP14 expression compared to uninvolved tissue (Figure 4.16.1C).

Additionally, we analyzed the expression data of SOX10, MYB, and MMP14 from three GEO datasets (GSE138772, GSE132223, and GSE139682) to examine the correlation between their

expression levels. The results demonstrated a positive correlation between MYB and SOX10 expression levels with MMP14 levels in the patients (**Figure 4.16.1E and F**). These findings further support the role of MMP14 as an oncogene.

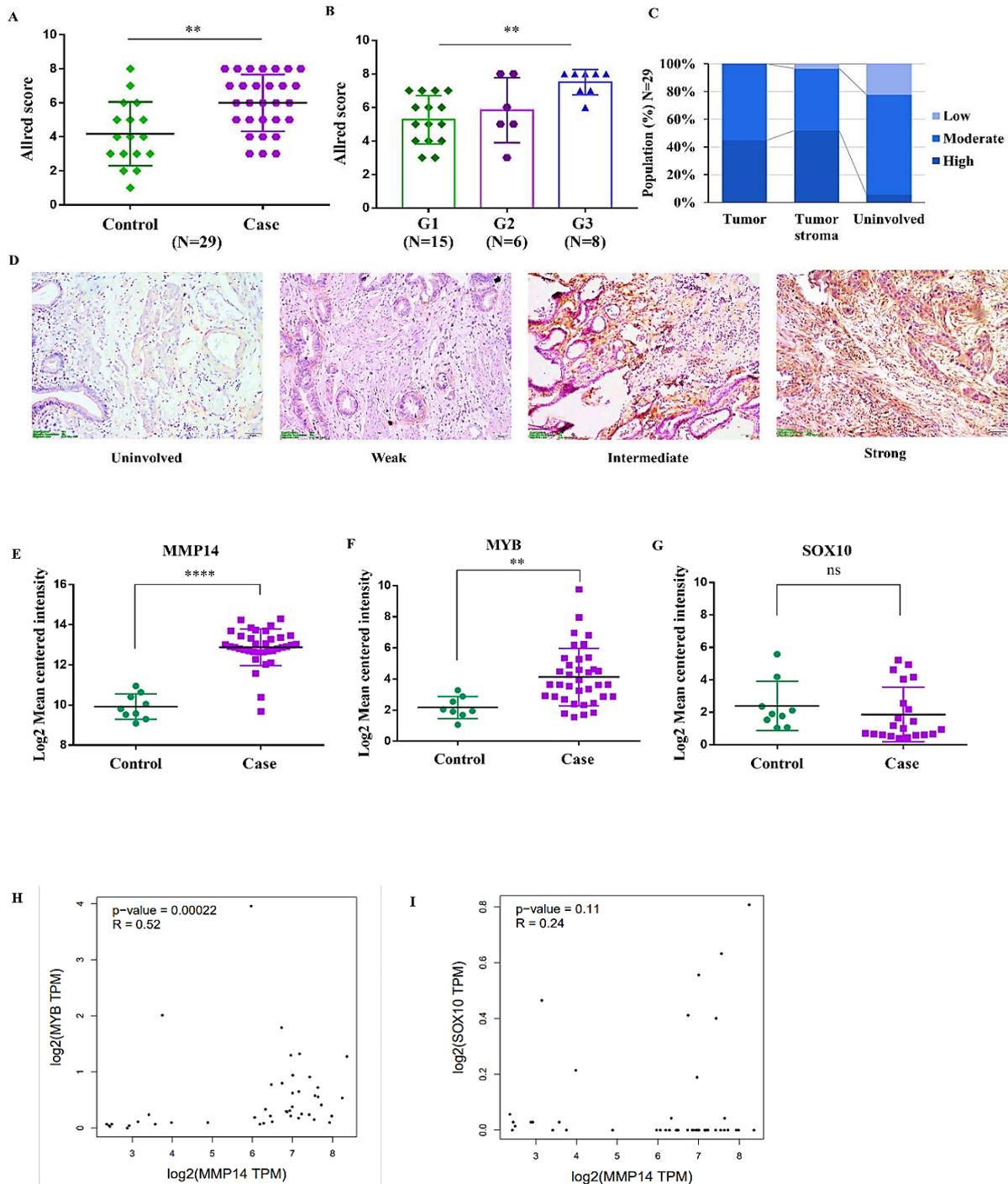


Figure 4.16.1 The expression and distribution of MMP14 protein in gallbladder tumour tissue and surrounding normal tissue. *The expression and distribution of MMP14 protein in*

*gallbladder tumour tissue and surrounding normal tissue. Analysis of the gallbladder and adjacent normal tissue using immunohistochemistry. (A) Patient samples showed increased MMP14 expression in gallbladder tumour tissue as compared to adjacent normal tissue samples ($P = 0.0022$). (B) The grade-wise distribution of MMP14 expression reveals a steady rise with advanced tumour grade ($P = 0.0054$). (C and D) Representative images of gallbladder cancer tissue immunohistochemical staining with varying degrees of staining from weak or negative, intermediate, and strong. Also shown are percentage distributions of Allred scores in tumour, tumour stroma, and normal tissue. The scale bar is 50 micrometres (magnification 10x, objective 10x). (E and F) Graphs from the GEO datasets illustrate the expression correlation (Pearson) of MMP14 with MYB (E) and SOX10 (F) in gallbladder cancer (GSE138772, GSE132223, and GSE139682). The statistical tests employed were the unpaired non-parametric two-sided Mann-Whitney test, the unpaired non-parametric one-way Kruskal-Wallis test, and the Pearson correlation analysis. ns $P > 0.05$, ** $P \leq 0.01$, **** $P \leq 0.0001$.*

4.17 MYB inhibitor treatment reduces the tumorigenic properties in GBC cells

MYB inhibitor treatment was found to have a significant effect on reducing the tumorigenic properties of GBC cells. In order to assess potential therapeutic strategies for individuals carrying risk alleles for genetic variants rs1003349 and rs1004030, cell-based assays were conducted using G415 cells treated with the MYB inhibitor Monensin sodium (1 μ M, PubChem SID: 24896992) (328, 329). Treatment with the MYB inhibitor resulted in a dose-dependent decrease in the expression of MMP14, as shown in **Figure 4.17.2**. Furthermore, inhibition of the MYB transcription factor significantly inhibited cell proliferation and colony formation in G415 cells, as depicted in **Figure 4.17.1A** and **B**. Additionally, treated cells demonstrated a notable reduction in migration and invasion efficiency compared to untreated cells, indicating a decrease in the metastatic potential of G415 cells, as illustrated in **Figure 4.17.1C** and **D**.

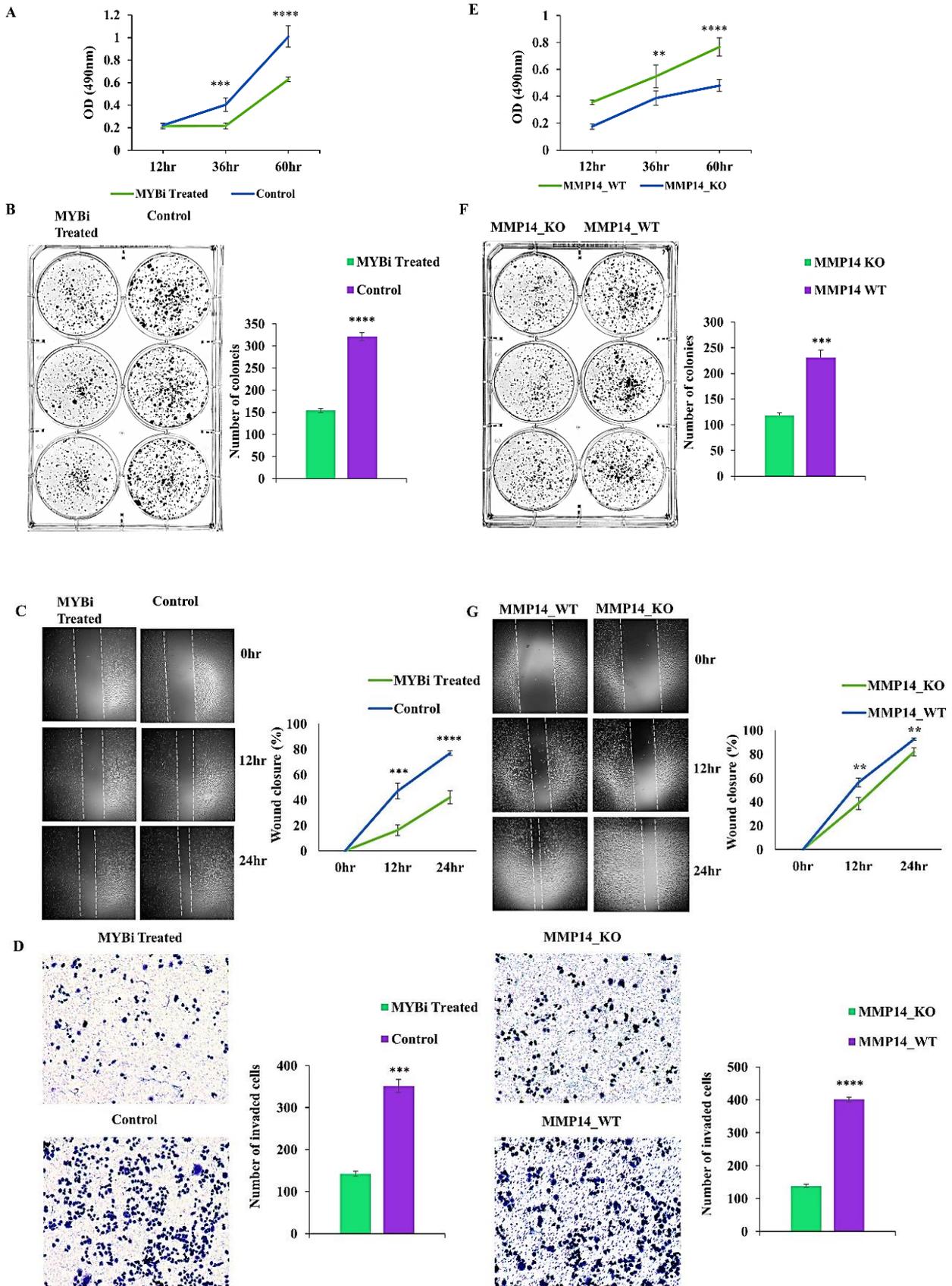


Figure 4.17.1 Effect of MYB inhibition and CRISPR-Cas9-mediated transcription factor binding site deletion on the tumorigenic properties of G415 cells. (A) The graph indicates the OD values from the MTS experiment for G415 cells treated with the MYB inhibitor (MYBi Treated, 1 M, Monensin sodium) and the control group (DMSO) ($P = 2.45E-05$). (B) The image and related bar graph display the number of colonies developed in the MYBi-treated and control (DMSO) groups ($P = 2.33E-05$). (C) The pictures and related graph show wound healing in MYBi-treated and untreated (DMSO) G415 cells ($P = 3.61E-05$). (D) The images and related bar graph illustrate the transwell invasion ability of MYBi treated and control (DMSO) G415 cells ($P = 0.0009$). (E) The graph displays the OD values for G415 cells with a promoter deletion (MMP14 P Del-199bp) compared to wild-type cells using the MTS assay ($P = 0.0001$). (F) The image and related bar graph display the number of colonies developed by cells with a promoter deletion (MMP14 P Del-199bp) and by cells with a wild-type promoter ($P = 4.52E-04$). (G) The images and related graph demonstrate wound healing in both wild-type and promoter-deficient G415 cells (MMP14 P Del-199bp; $P = 0.0022$). (H) The images and related bar graph demonstrate transwell invasion ability in wild-type (WT) and G415 cells with promoter deletion (MMP14 P Del-199bp) ($P = 0.0001$). The experiments were carried out in triplicate, and the results are shown as mean SD. The statistical significance of the difference between the groups was determined using the Student's *t*-test.

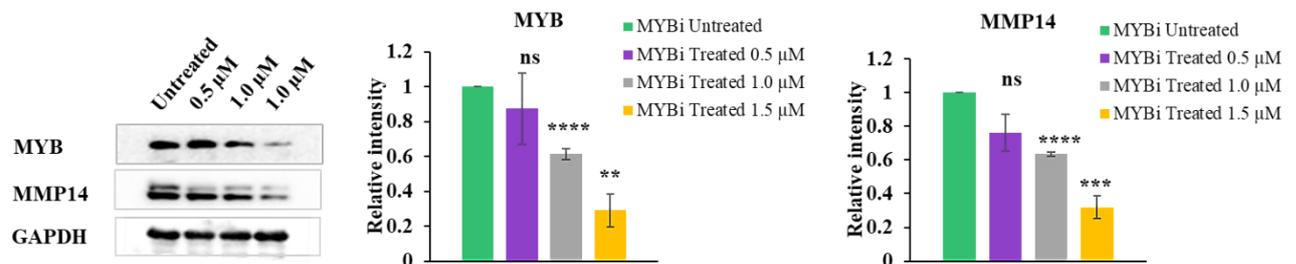


Figure 4.17.2 The MMP14 expression level following MYB inhibition treatment in G415 cells. MMP14 expression levels are decreased in a dose-dependent way in Western blot analysis after MYB inhibitor treatment at various concentrations, as shown by a graphical representation of the densitometry data. The values are shown as mean standard deviation. Statistics were calculated between the groups using the student's *t*-test. $P > 0.05$, **, $P 0.01$, ***, $P 0.001$, and ****, $P 0.0001$, among other statistics.

4.18 Deletion of the transcription factor binding site reduces the tumorigenic properties in GBC cells

Deletion of the transcription factor binding site in GBC cells resulted in a decrease in their tumorigenic properties. We conducted an experiment where we modified the promoter to eliminate the binding site for transcription factors and assessed the tumorigenic and metastatic

potential of G415 cells. The GBC cells with the deleted promoter exhibited a notable reduction in cell proliferation and colony formation ability when compared to the wild-type cells (**Figure 4.17.1E and F**). Furthermore, the cells with the deleted promoter demonstrated significantly diminished migration and invasion capabilities compared to the wild-type cells (**Figure 4.17.1G and H**).

4.19 Discussion

MMP14 are members of the zinc-dependent endopeptidases protein family (330). It has been shown that the MMP14 promote tumour invasion, proliferation, differentiation, and apoptosis (10, 276, 331). Phosphorylation of tyrosine and threonine residues in MMP14's cytoplasmic tail promotes activation of intracellular signalling pathways such as Akt, ERK1/2, and Rac1, which play important roles in tumour cell proliferation and migration (332, 333). Various chronic inflammation-induced malignancies (334, 335), including gastric adenocarcinoma (336), hepatocellular carcinoma (40), colorectal carcinoma (337), cholangiocarcinoma (338), and gallbladder cancer (51), have been associated with increased MMP14 expression. In this study, we explored the relationships between genetic variations in the MMP14 promoter region and gallbladder cancer and their putative role in regulating MMP14 expression via altering SOX10 and MYB transcription factor binding sites. Aggressive metastatic tumour cells are known to modify adjacent extracellular matrix to enhance invasion into stromal cells, organs, and lymph nodes (339). The MMP14 has important role in stromal remodelling and over expressed and colocalised to invadopodia in many cancers (333). In gallbladder tumours and the stromal cells surrounding them, we observed a markedly increased expression of MMP14.

Numerous studies have reported association of MMP14 variants as risk factor for various diseases and cancer, including hepatocellular carcinoma (38), ovarian cancer (39), oral cancer

(291), Winchester syndrome (292), osteonecrosis of the femoral head (293), and intervertebral disc degeneration (294). Although studies are available in other cancers, we are the first to report association of MMP14 promoter variants with gallbladder cancer.

The current study we explored the association of MMP14 promoter variants rs1003349 and rs1004030 with GBC. We found significant association of variants rs1004030 and rs1003349 with GBC. Previous studies on MMP14 in the Zhuang ethnic group from Guangxi, China, revealed that the risk alleles 'G' and 'C' at loci rs1003349 and rs1004030, respectively, were strongly linked to osteoporosis (296). Additionally, Astrid Munkert et al. found that in the German population, the risk allele 'G' at rs1003349 was strongly linked to both focal and segmental glomerulosclerosis and mesangiocapillary glomerulonephritis (294). Similarly, in our population we found presence of alleles 'G' and 'C' at loci rs1003349 and rs1004030, increase the risk for GBC. In addition to these findings, we found that alleles 'G' and 'C' at loci rs1003349 and rs1004030 'G' alters the expression of MMP14 and leads to GBC pathogenesis. This might be due to presence of regulatory promoter variants might alter binding site of transcriptional regulators, thereby affecting the expression levels. In the prior study reported the presence of binding sites for Sp1 and RR1, flanking SNPs rs1003349 and rs1004030 and, show allele dependent regulation of MMP14 expression (295). The ENCODE data also indicates that both variants are found in the cis-Regulatory Element, which is emphasized by the presence of an active promoter signature and significant DNase and H3K4me3 signals. In addition, our in-silico analysis found that 'C' allele of rs1004030 and 'G' allele of rs1003349 create binding site for SOX10 and c-Myb, respectively.

To validate these in-silico predictions, we carried out in depth *in-vitro* and *in-vivo* studies in GBC. Our study provides clear evidence of functional implication of SNPs. We verified the findings using EMSA, ChIP, luciferase assays, and CRISPR-cas9-based deletion of a 119-bp

genomic area surrounding the locus, in which loci rs1004030 and rs1003349 with alleles 'C' and 'G' were found to facilitate binding of transcription factors SOX10 and MYB, respectively. Suggests the presence of functional SNPs rs1004030 and rs1003349 in the promoter region may have a direct regulatory role in MMP14 expression by creating the binding site for SOX10 and MYB.

MYB (proto-oncogene or c-Myb) is overexpressed in leukemias, breast cancer, colon cancer, adenoid cystic carcinoma, and osteosarcoma (340, 341). Additionally, the transcriptome gene enrichment functional analysis identifies additional genes that are involved in controlling the stress response, cell adhesion, and cell differentiation or morphogenesis (342, 343). Gallbladder mucosa from patients with gallbladder disease exhibits oxidative stress (344). The cells under stress regulate MYB protein levels and increase transcriptional activity via post-translational SUMOylation (Small Ubiquitin-like Modifier protein) (345). MYB expression enhances NOX1 mediated p38 MAPK pathway and protects colorectal carcinomas cells against cisplatin and doxorubicin-induced apoptosis (346).

SOX-10 overexpression has been identified as a differential diagnostic marker for metastasis and survival outcomes in triple-negative breast cancer (347), bladder cancer (348), and nasopharyngeal carcinoma (349). The frequent mutation and small sample size may be responsible for the decreased expression of SOX10 in the bile duct TCGA cohort. Also, the SOX10 transcription factors are essential for neural crest cell (NCC) lineage development and function. NCC lineage is a highly migratory stem cell population (350). Given that cancer metastasis and NCC migration are comparable, it is anticipated that MMP14 plays an important role in both cancer metastasis and the epithelial-to-mesenchymal transition of neural crest stem cells (351, 352). In HCC, SOX10 enhances Wnt/ β -catenin signalling by facilitating the binding of TCF4 to β -catenin and forming a stable SOX10/TCF4/ β -catenin complex (353, 354). The canonical Wnt signalling pathway regulates SOX10 expression levels in melanoma in a β -

catenin-dependent manner (355). In non-cancerous cells, β -catenin prevents MMP14 from localizing to the membrane and activating pro-MMP2 by proteolysis. However, in cancer cells, β -catenin promotes MMP14 expression and Wnt-3a-mediated signalling via the transcription factors Tcf-4/Lef (356). The interaction of c-Myb with β -catenin increases breast cancer invasion and metastasis via increasing Wnt/ β -catenin/Axin2 signalling (357). In colorectal tumors (CRC), MYB also facilitated Wnt signalling activation and elevated MYC expression, which illustrates the complex regulation of MMP14 activity in cancer and non-cancer cells mediated by transcription factors MYB and SOX10 in association with Wnt signalling (358). However, xenograft or transgenic animal models might be helpful in understanding the molecular mechanism in an in-vivo system. The prognosis of the patient and expression studies in larger patient samples will assist in the identification of a biomarker for invasive advanced gallbladder cancers.

Overall, the results of the current investigation point to a novel association between the MMP14 promoter variants rs1004030 and rs1003349 and gallbladder cancer. MMP14 expression was found to be increased in gallbladder cancer tissue, which may aid in tumour metastasis. Our results emphasize the regulatory functions of MMP14 promoter variations, which promote SOX10 and MYB-mediated MMP14 expression in GBC pathogenesis. Validate the allele-specific functional implication of rs1004030 and rs1003349 on MMP14 expression levels.



CHAPTER 5

Results and Discussion



Vinay J
NISER, Bhubaneswar

Objective 2

To determine the association of MMP7 promoter variants with GBC and its effect on gene expression

Findings of this section has been communicated.

- **Vinay J, Singh SP, Dixit M.** Matrix metalloproteinase-7 promoter variants association and functional significance in gallbladder cancer. BMC Genomics, (2023).

5.1 Introduction

Matrix metalloproteinase-7 (MMP7; MIM: 178990) is a small secretory proteolytic enzyme with broad substrate selectivity. It is constitutively expressed in various organs, such as the adult ductal epithelium, intestinal glandular epithelium, pancreas, liver, and breast (359). MMP7 expression is frequently elevated in cancers of the esophagus (360), colon (361), liver (46), ovary (45), and pancreatic cancers (44, 362, 363). It is positively correlated with the severity of cancer invasiveness (47). These findings suggest MMP7's involvement in cancer pathogenesis and poor prognosis. Increased MMP7 expression can induce tumorigenesis by facilitating angiogenesis and EMT (47).

The SNPs in the promoter might affect the expression of MMP7, thus altering the risk for cancer. MMP7 variants rs11568818 and rs11225297, with rare alleles 'G' and 'A,' respectively, show poor survival in Chinese women with breast cancer (298). The functional promoter variants A181G (rs11568818) and C153T (rs11568819) are reported to be risk factors for squamous cell carcinoma, gastric adenocarcinoma (299, 300), and non-small cell lung carcinoma (301). They exhibit allele-specific regulation of gene expression by interacting with numerous nuclear proteins (302-305). Patients with hypertension (304) and gastric cancer (299) have been shown to have allele-specific Cyclic AMP Response Element-Binding Protein (CREB) binding at promoter polymorphism rs11568818 (299, 304). In each instance, the risk allele "G" increases MMP7 expression and provides a binding site for the transcription factor CREB. A study identified the MMP7 promoter variant rs11568818 as a risk factor for GBC (52). However, the role of other SNPs and MMP7 expression in GBC remains unclear.

In this chapter we investigate the association of MMP7 promoter variants with gallbladder cancer and its putative role in MMP7 expression. We showed increased expression of MMP7

in tumors and allele-specific expression of MMP7. Overall, the present study emphasizes the functional relevance of these variants in gallbladder cancer.

Results

5.2 Screening of MMP7 promoter SNPs

To evaluate genetic variation in the MMP7 promoter region we scanned the MMP7 promoter sequence covering 1500 bases from the transcript start site. We used two primer sets to screen all SNPs within the region. The screening set consists of 50 GBC and 50 healthy control samples. We found five SNPs in the region, out of which four variants [g.102530902G>A (rs11568819), g.102530930T>C (rs11568818), g.102532089A>C (rs113823671), and g.102532127G>A (rs17098318)] showed allelic variation, while g.102530864G>A (rs17886546) was found non-polymorphic.

5.3 Discovery set allele and genotype frequency analysis

We analysed all MMP7 promoter variants, rs11568819, rs11568818, rs113823671 and rs17098318, to understand population specific effect size or odds of disease exposure and minor allele frequencies. As shown in **Table 5.3.1**, only two of the four polymorphisms were significantly associated with GBC in the discovery set with p value<0.05 (rs113823671 and rs17098318). These variants were tested for genetic association with GBC in the subsequent validation set. The analysis of allelic and genotype frequencies for each SNP in discovery set is discussed in following sections.

rs11568819 G>A variant shows no significant association with GBC

The allelic and genotypic frequency distribution of variant rs11568819, G>A between gallbladder cancer and control population showed 'G' as the major allele and 'A' as the minor allele in our study population (**Table 5.3.1**). rs11568819, G>A showed no significant association with GBC at both, allelic ($p = 1.0$, OR = 2.02, 95% CI, 0.18-22.27) and genotypic levels ($p = 1.0$, OR = 2.04, 95% CI, 0.18-23.27).

The variant rs11568818 T>C shows no significant association with GBC

The allelic and genotypic distribution of the variant between the case and control groups shows 'T' as a major allele and 'C' as a minor allele in our study population (**Table 5.3.1**). rs11568818 T>C showed no significant association with GBC at both, allelic ($p = 0.5443$, OR = 1.20, 95% CI, 0.66-2.18) and genotypic levels ($p = 0.6468$, OR = 1.36, 95% CI, 0.37-5.09).

rs17098318 G>A shows significant association with GBC

The genotype and allele frequencies of gallbladder cancer and the control population were compared, the promoter variant rs17098318 was found to be significantly associated with GBC in discovery set. The minor allele 'A' was more frequent in GBC and showed significant increase in risk for GBC ($p = 0.032$, OR = 2.39, 95% CI, 1.06-05.39). rs17098318 variant was tested for genetic association with GBC in the subsequent validation set (**Table 5.3.1**).

The variant rs113823671 A>C shows significant association with GBC

The allelic and genotypic distribution of the variants between the case and control showed 'A' as a major allele and 'C' as a minor allele in our study population. MMP7 promoter variant rs113823671 showed significant association with GBC. The allelic ($p = 0.033$, OR = 5.44, 95% CI, 1.16-25.52) and genotypic models ($p = 0.028$, OR = 06, 95% CI, 1.24-28.99) showed

To determine the association of MMP7 promoter variants with GBC and its effect on gene expression

significant association with GBC in discovery set. The variant rs113823671 was tested for genetic association with GBC in the subsequent validation set (**Table 5.3.1**).

Table 5.3.1 Allelic and genotypic frequency distribution of MMP7 promoter variants in GBC and control subjects (Discovery set)

	rs11568818	Allele or genotype count (frequency)			
	Alleles/ Genotypes	Patients (N = 50)	Control (N = 50)	P value	OR (95% CI)
Genotypic model	TT ^{ref}	22 (0.44)	25 (0.50)		
	TC	22 (0.44)	20 (0.40)	0.597	1.25 (0.54-2.88)
	CC	06 (0.12)	05 (0.10)	0.647	1.36 (0.37-5.09)
Allelic model	T ^{ref}	66 (0.66)	70 (0.70)	0.544	1.20 (0.66-2.18)
	C	34 (0.34)	30 (0.30)		
	rs11568819				
Genotypic model	GG ^{ref}	48 (0.96)	49 (0.98)		
	GA	02 (0.04)	01 (0.02)	1.0*	2.04 (0.18-23.27)
	AA	0 (0.00)	0 (0.00)	-	-
Allelic model	G ^{ref}	98 (0.98)	99 (0.99)	1.0*	2.02 (0.18-22.65)
	A	02 (0.02)	01 (0.01)		
	rs17098318				
Genotypic model	GG ^{ref}	32 (0.64)	41 (0.82)		
	GA	15 (0.30)	08 (0.16)	0.074	2.40 (0.91-06.37)
	AA	3 (0.06)	01 (0.02)	0.325*	3.84 (0.38-38.72)
Allelic model	G ^{ref}	79 (0.79)	90 (0.90)	0.032	2.39 (1.06-05.39)
	A	21 (0.21)	10 (0.10)		
	rs113823671				
Genotypic model	AA ^{ref}	40 (0.80)	48 (0.96)		
	AC	10 (0.20)	02 (0.04)	0.028*	06 (1.24-28.99)
	CC	0 (0.00)	0 (0.00)	-	-
Allelic model	A	90 (0.90)	98 (0.98)	0.033*	5.44 (1.16-25.52)
	C	10 (0.10)	02 (0.02)		

OR-odds ratio, **CI**-confidence interval, **N**-sample size, *ref*-reference

P value of Chi-square test with Yates continuity correction

P_{adjusted} and **OR_{adjusted}**, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

5.4 Analysis of Hardy-Weinberg equilibrium and study power

Variants rs113823671 and rs17098318 allelic frequencies were tested for Hardy-Weinberg equilibrium using a significance cut-off of $p < 0.05$. The SNPs rs113823671 ($p = 0.7920$) and rs17098318 ($p = 0.0791$) passed the HWE test indicating genotype frequencies in control population satisfy Hardy-Weinberg assumptions (**Table 5.4.1**). The study power was calculated for rs113823671 and rs17098318 using PS statistical power analysis programme. The set parameter to calculate power of the study were, 80% beta power and 0.05 alpha significance level (**Table 5.4.1**).

Table 5.4.1 Sample size and HWE calculations for MMP7 promoter SNPs.

Sl. No.	SNP	Calculated Sample size*		HWE p-value#
		Case	control	
1	rs17098318	244	244	0.0791
2	rs113823671	288	288	0.7920

* PS statistical power program-based analysis at 80% beta power and 0.05 alpha significance level

HWE was calculated based on the study's actual number of control samples.

5.5 Demographic profile of study subjects

This study included 300 GBC patients and 300 healthy controls. In **Table 4.5.1**, the demographic characteristics of the study subjects are listed. The mean age of GBC incidence in our study population was found to be 53.54 years ($SD \pm 11.21$). The analysis of gender wise incidence of GBC, high incidence was observed in females compared to male patients ($P = 0.0001$).

Table 5.5.1 Demographic profile of the gallbladder cancer patients and controls recruited in the study

Characteristics	Cases (n = 300)	Control (n = 300)	P value
Age ^a	53.54 \pm 11.21	45.13 \pm 11.37	< 0.0001
Gender ^b			
Female	171	114	< 0.0001

To determine the association of MMP7 promoter variants with GBC and its effect on gene expression

Male	129	186	
-------------	-----	-----	--

^a Student's T-test was used to compare mean values of age

^b Chi-Square test was used to compare the difference in frequency of male and female

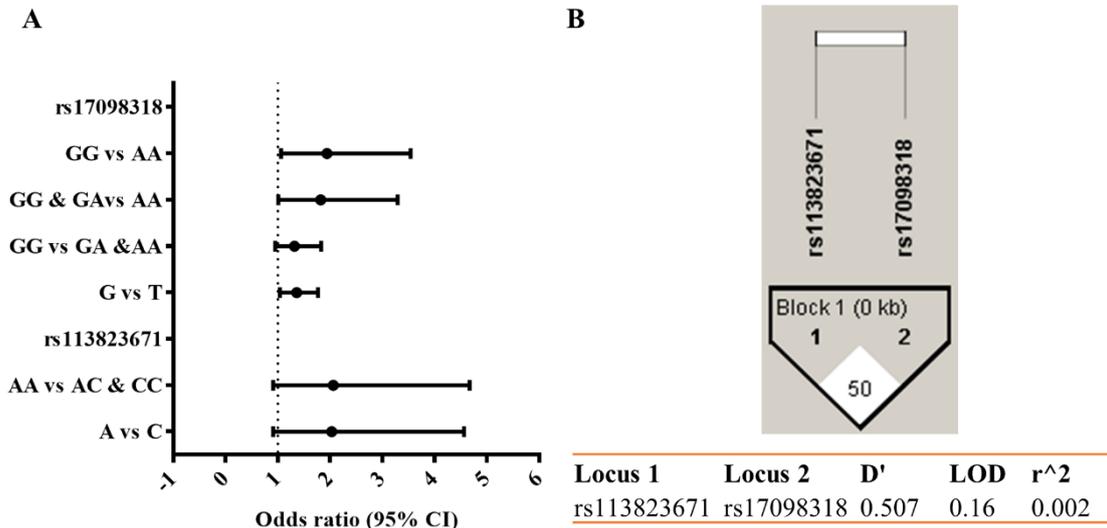


Figure 5.5.1 Gallbladder cancer genetic risk forest plot and linkage disequilibrium plot for MMP7 promoter variants rs113823671, and rs17098318. (A) The odds ratios and 95% confidence intervals of each variants genotype and its individual allelic effect on gallbladder cancer are shown as dark spots and matching horizontal lines. * Reference genotype or allele, odds ratio (OR), and confidence interval (CI). (B) A linkage disequilibrium plot created with the software programme Haploview. Pairwise D' values are used to illustrate linkage disequilibrium (LD). The amplitude and relevance of pairwise LD are shown by shading, with a red-to-white gradient showing higher-to-lower LD values.

5.6 rs113823671 allele C is associated with an increased risk for gallbladder cancer

The allele and genotypic distribution between gallbladder cancer and the control population showed allele 'A' as the major and 'C' as the minor allele. In rs113823671, we did not find any significant association with GBC at both allelic ($p = 0.080$, OR = 2.03, 95% CI, 0.91-4.56) and genotypic levels ($p = 0.076$, OR = 2.06, 95% CI, 0.91-4.67) (

Table 5.6.1, Figure 5.5.1A). After the adjustment of the confounding factors such as age and gender, in the allelic model, individuals carrying allele 'C' showed an increased risk for GBC compared with allele 'A' carriers ($p = 0.034$, OR = 2.54, 95% CI, 1.08-5.98). Similarly, in the genotypic model, we found individuals with genotype 'AC' showed an increased risk for GBC susceptibility with a p-value of 0.031 (OR = 2.60, 95% CI, 1.09-6.21) when compared with wildtype 'AA' genotype. Altogether, the variant rs113823671 shows an increased risk for GBC once confounding factors such as age and gender are adjusted, which emphasizes the confounding factor's contribution toward GBC as a risk factor.

Table 5.6.1 Allelic and genotypic frequency distribution of rs113823671 in GBC and control subjects

	rs113823671	Allele or genotype count (frequency)		P value	OR (95% CI)	P ^{adjusted}	OR ^{adjusted} (95% CI)
		Patients (N = 300)	Control (N = 300)				
Genotypic model	AA ^{ref}	282 (0.94)	291 (0.97)				
	AC.	18 (0.06)	9 (0.03)	0.0763	2.06 (0.91-4.67)	0.0314	3.10 (1.09-6.21)
	CC	0 (0.00)	0 (0.00)	-	-	-	-
Allelic model	A ^{ref}	582 (0.97)	591 (0.99)	0.0798	2.03 (0.91-4.56)	0.0336	2.54 (1.08-5.98)
	C	18 (0.03)	9 (0.01)				

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P^{adjusted} and OR^{adjusted}, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

5.7 Allele A of variant rs17098318 is associated with Gallbladder cancer

The study of allelic and genotypic distribution at rs17098318 between GBC patients and the control population found allele 'G' as major and 'A' as minor allele in our study population. The frequency of allele 'A' was more in GBC patients compared to the control population. The association at the allelic level showed that patients with minor allele 'A' increase the risk for gallbladder cancer ($p = 0.022$, OR = 1.36, 95% CI, 1.04-1.77) (**Table 5.7.1, Figure 5.5.1A**).

To determine the association of MMP7 promoter variants with GBC and its effect on gene expression

In the genotypic model, it was found that the individuals carrying rs17098318 'AA' genotype were at significantly higher risk for GBC susceptibility with a p-value of 0.029 (OR = 1.94, 95% CI, 1.06-3.54) when compared with wild-type 'GG' genotype. To figure out the effect of allelic interaction, various models were analyzed and it was found that rs17098318 follows a recessive model in which 'AA' genotype increases the risk for GBC (p = 0.042, OR = 1.82, 95% CI, 1.01-3.29). On the other hand, in the case of the dominant model (p = 0.096) and the additive model (p = 0.079), no significant association with GBC was observed. The confounding factors were adjusted using binomial logistic regression analysis and found that the allele 'A' retained its significance with GBC (p = 0.024, OR = 1.39, 95% CI, 1.04-1.86) but, we found no significant association in case of recessive model (p = 0.052). Overall, the variant rs17098318 showed significant association with GBC. The individual with allele 'A' increased risk for GBC that is independent of confounding factors like age and gender.

Table 5.7.1 Allelic and genotypic frequency distribution of rs17098318 in GBC and control subjects

	rs17098318	Allele or genotype count (frequency)		P value	OR (95% CI)	P _{adjusted}	OR _{adjusted} (95% CI)
		Patients (N = 300)	Control (N = 300)				
Genotypic model	GG ^{ref}	170 (0.57)	190 (0.63)				
	GA	97 (0.32)	91 (0.30)	0.3297	1.19 (0.84-1.70)	0.0348	2.04 (1.05-3.94)
	AA	33 (0.11)	19 (0.06)	0.0286	1.94 (1.06-3.54)	0.1484	1.67 (0.83-3.33)
Allelic model	G ^{ref}	437 (0.73)	471 (0.79)	0.0222	1.36 (1.04-1.77)	0.0242	1.39 (1.04-1.86)
	A	163 (0.27)	129 (0.22)				
Dominant model	GG ^{ref}	170 (0.57)	190 (0.63)	0.0956	1.32 (0.95-1.83)	0.0940	1.36 (0.95-1.94)
	GA+AA	130 (0.43)	110 (0.37)				
Recessive model	GG+GA ^{ref}	267 (0.89)	281 (0.94)	0.0422	1.82 (1.01-3.29)	0.0519	1.90 (0.99-3.62)
	AA	33 (0.11)	19 (0.06)				
Additive model	GG vs GA vs AA			0.0792			

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P^{adjusted} and OR^{adjusted} , adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

5.8 Haplotype analysis

The cumulative effect of MMP7 promoter variations rs113823671 and rs17098318 on genetic predisposition in GBC was evaluated by the study of linkage disequilibrium (LD) and haplotypic association. The frequency distribution of the haplotype revealed haplotype A-G (rs113823671, and rs17098318, respectively) as the most frequent haplotype (

Table 5.8.1). Found no significant association in any of the haplotype. Additionally, the Monte Carlo test for empirical significance with 10,000 permutations was carried out to account for genotyping errors and unknown factors. However, no significant improvement was observed after permutation in both haplotype A-A and C-G. However, Also, the correlation coefficient r^2 between rs113823671, and rs17098318 was 0.002 which suggests weak linkage, suggesting independent effects on the alteration of risk for GBC (**Figure 5.5.1B**).

Table 5.8.1 Haplotype association of the variants rs113823671 and rs17098318 with gallbladder cancer.

Haplotype	Frequencies		OR (95% CI)	p-value	p-value ^b
	Case (N = 300)	Control (N = 300)			
A-G ^a	0.702 (211)	0.771 (231)	1	-	-
A-A	0.268 (80)	0.214 (64)	1.37 (0.94-1.99)	0.1029	0.0712
C-G	0.026 (8)	0.014 (4)	2.19 (0.65-7.38)	0.1949	0.2687

^a Reference haplotype

^b Empirical p value at N= 10,000 permutations

5.9 Promoter variant rs113823671 risk allele C shows increased luciferase activity

Since the variants are in the MMP7 promoter region, we hypothesized that they might play a regulatory role in the expression of MMP7. Through an in-silico analysis of the MMP7 promoter region, we found that these variations could facilitate the binding of transcription factors **Table 5.9.1**. To test our hypothesis, we performed luciferase reporter assays in two mammalian cell lines, HEK293T and TGBC1TKB. Allele-based regulatory role luciferase reporter assays were carried out by cloning a 37 bp genomic region surrounding rs113823671 (either 'A' or 'C') and rs17098318 (either 'G' or 'A') into pGL4.23 luciferase minimal promoter vector. Constructs with the rs113823671 'C' allele in TGBC1TKB cells (8.49 ± 1.52) showed a

significant increase in reporter activity ($p = 0.0003$) compared to the 'A' allele (4.39 ± 1.33) (**Figure 5.9.1A**). Similarly, allele 'C' (4.21 ± 0.38) in HEK293T cells showed a significant increase in luciferase activity compared to allele 'A' (2.08 ± 0.13) ($p = 7.86E-11$) (**Figure 5.9.1B**). However, the reporter luciferase activity at rs17098318 revealed a significant difference between the 'G' and 'A' alleles in TGBC1TKB ($p = 0.030$) (**Figure 5.9.1A**), but not in HEK293T ($p = 0.725$) (**Figure 5.9.1B**). These observations suggest that the risk allele 'C' of rs113823671 may contribute to increased MMP7 expression.

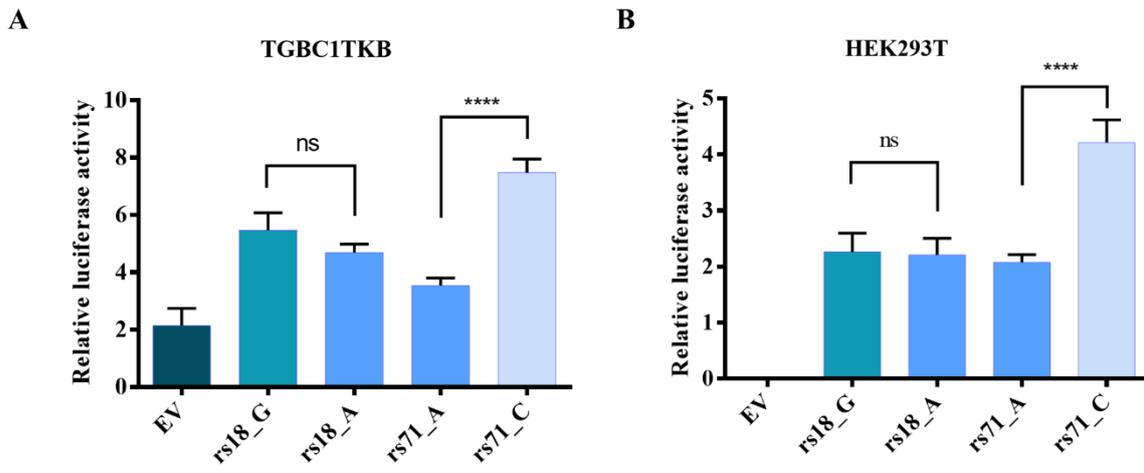


Figure 5.9.1 Allele-specific luciferase activity of variants rs113823671 and rs17098318.

(A and B) Relative luciferase activity of MMP7 promoter variants in TGBC1TKB and HEK293T cell lines, X-axis shows pGL4.23 empty vector (EV), pGL4.23 with promoter insert containing allele 'A' and 'C' of rs113823671 (rs71) and rs17098318 (rs18) with alleles 'G' and 'A.' The Y axis shows the relative luciferase activity of each construct. An unpaired student's t-test was used to study statistical differences between the groups. All experiments were replicated independently three times.

Table 5.9.1 Putative transcription factor binding at variants rs113823671 and rs17098318.

SNPs	Sequence	Putative TF Binding Site
rs17098318	ACAT A TGATAA	USF-1 [T00877]
rs113823671	TT C ACTAGAG	SREBP-1c [T01562]

5.10 Gallbladder cancer tissue expression of MMP7 and its association with rs113823671 and rs17098318 genotypes

MMP7 expression was studied using immunohistochemistry in 27 samples of gallbladder cancer tissue and surrounding normal tissue. MMP7 expression was significantly higher in

gallbladder tumour tissue (mean Allred score 5.96) than in adjacent normal tissue (mean Allred score 4.43, $p = 0.0061$) (**Figure 5.10.1A and B**). TCGA data set (Project ID: TCGA-CHOL) was used to validate further the MMP7 expression (tcga-data.nci.nih.gov, accessed 4 March 2021). The findings showed that tumour samples expressed more MMP7 than control samples ($p = 0.0001$) (**Figure 5.10.1C**). Additionally, we studied MMP7 expression in GBC tissue samples having different genotype. Found significant difference between MMP7 expression and the genotype of the promoter variants rs113823671 and rs17098318. In contrast to patients with genotype 'AA' (Mean Allred score 5.88 ± 0.24), those with genotype 'AC' (Mean Allred score 7.0 ± 0.0) at rs113823671 had a significant increase in MMP7 expression ($p = 0.0001$) (**Figure 5.10.2A**).

Likewise, at locus rs17098318, we observed that patients with the 'AA' genotype (Mean Allred score 7.40 ± 0.25) had significantly higher MMP7 expression than patients with the 'GG' genotype ($p = 0.0001$). (Mean Allred score 4.83 ± 0.17). Also, patients with the 'GA' genotype (Mean Allred score 6.60 ± 0.16) had higher MMP7 expression ($p = 0.0001$) than those with the 'GG' genotype (**Figure 5.10.2B**). Overall, we found that each copy of the risk alleles at promoter variants rs113823671 and rs17098318 led to increased MMP7 expression. Moreover, our findings support the functional significance of MMP7 promoter variants in regulating MMP7 gene expression.

To determine the association of MMP7 promoter variants with GBC and its effect on gene expression

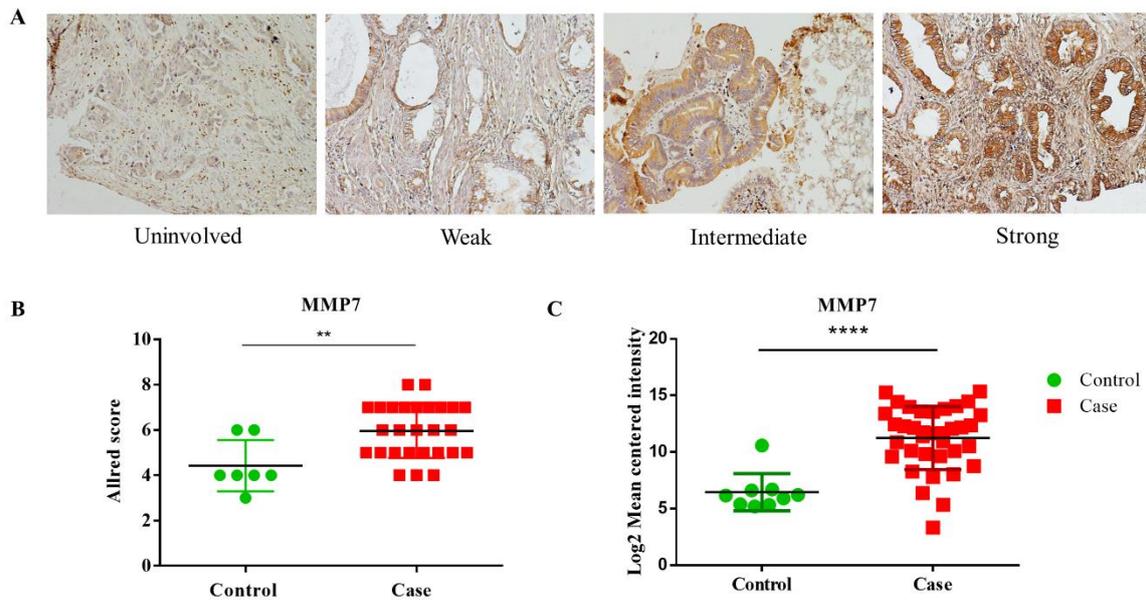


Figure 5.10.1 The MMP7 protein expression in gallbladder tumour tissue and adjacent normal tissue. (A) Representative images of gallbladder cancer tissue immunohistochemistry (IHC) with weak or negative, intermediate, and strong staining intensity. The scale bar represents 50 μm with a 10X objective and 10X magnification. (B) MMP7 IHC scoring of gallbladder cancer tissue and adjacent normal tissue samples, X-axis shows sample type, case, and control. Y axis shows Allred score. The unpaired non-parametric two-sided Mann-Whitney test was used to analyze the statistical difference between the groups (C) RNA-seq expression data of MMP7 from the TCGA database (TCGA-CHOL), where the X axis represents sample type, case, and control, and the Y axis represents Log2 mean intensity value. Statistical differences between the groups were analyzed by unpaired non-parametric two-sided Mann-Whitney test.

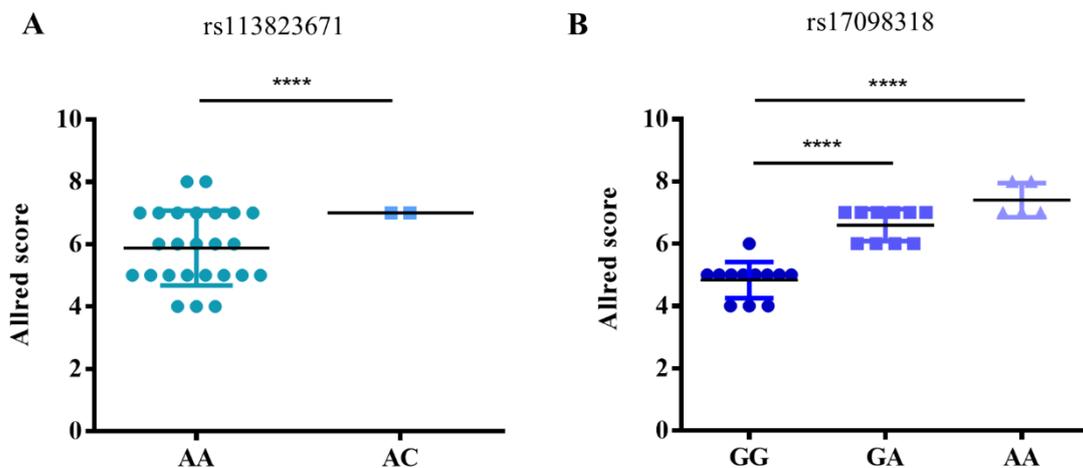


Figure 5.10.2 MMP7 protein expression in gallbladder cancer patients with different genotypes of variants rs113823671 and rs17098318. (A) Analysis of MMP7 expression in

GBC patient's between rs113823671 genotype, X-axis shows rs113823671 genotypes and the Y-axis shows respective Allred score. (B) Similarly, for rs17098318; X-axis shows rs17098318 genotypes and Y-axis shows respective Allred score. The statistical test used was a two-sided unpaired student's t-test with Welch's correction.

5.11 Discussion

MMP7 is expressed in the ductal and epithelial linings of major exocrine glands. MMP7 is commonly associated with tumourinvasion and is up-regulated in colon, liver, esophageal, and pancreatic cancers (44-47, 360-363). Despite several findings on MMP7's role in the pathogenesis and poor prognosis of many cancers, there have been limited studies on GBC. We found significantly increased MMP7 expression in patients with GBC. In this study, we explored the association of MMP7 promoter variants with gallbladder cancer and their possible role in regulating MMP7.

Previous studies have shown that genetic variants in the MMP7 gene are a risk factor for several cancers, including breast cancer (298), squamous cell carcinoma, gastric cancer (299), cardiac adenocarcinoma (300), and non-small cell lung carcinoma (301). rs11568818, T181C polymorphism in gastric cancer is significantly associated with lymph node metastases, vascular invasion, and poor overall survival of patients (364). T181C (rs11568818) and G153A (rs11568819) variations in hypercholesterolemic individuals show allele-specific regulation of MMP7 promoter activity, which is directly related to coronary artery luminal size (302). In addition, T181C (rs11568818) was associated with essential hypertension in a Caucasian population in central Russia (365). It emphasizes the functional implications of these variants, further leading us to understand the mechanistic role of gallbladder cancer better.

The current study we investigated the association of MMP7 promoter variants rs11568819 G>A, rs11568818 T>C, rs113823671 A>C, and rs17098318 G>A with gallbladder cancer and its putative role in MMP7 expression. The variants rs11568819 G>A and rs11568818 T>C

showed no association between allelic and genotypic levels in the discovery set. The variant rs113823671 A>C showed genetic association only at the allelic level. As both age and gender are known independent risk factors for GBC, adjusting for confounding factors like age and gender significantly elevated the risk in both allelic and dominant models. The variant rs17098318 G>A showed significant association in both allelic and recessive models. The reporter luciferase assay and genotype-phenotype correlation studies were performed to validate the potential role of the associated SNPs rs113823671 A>C and rs17098318 G>A in the expression of MMP7. The risk allele 'C' at rs113823671 significantly increased luciferase activity in both TGBC1TKB and HEK293T cell lines. In contrast, rs17098318 in HEK293T cells showed no significant change in luciferase activity, whereas TGBC1TKB cells showed marginal change in luciferase activity. Thus, luciferase data suggests that rs113823671 has a more significant influence, which requires further mechanistic study. Further, rs113823671 genotype-phenotype data complements the genetic association data in GBC. In order to delineate the mechanistic role of these promoter variants, we carried out in-silico promoter analysis. In the analysis we found that the SREBP-1c (Sterol-regulatory element binding protein-1c) and USF-1 (Upstream Transcription Factor 1) binding sites are created by the 'C' allele of rs113823671 and the 'A' allele of rs17098318, respectively (**Table 5.9.1**). These in-silico predictions need to be validated through in-depth, in-vitro, and in-vivo studies. In an allele-dependent manner, the USF-1 allelic variants regulate the expression of USF-1 in atherosclerotic plaque and lipoprotein metabolism (366). SREBP-1c is an intracellular cholesterol sensor in the endoplasmic reticulum that promotes colorectal cancer cell invasion by increasing MMP7 expression and NF-B pathway activation (367). Overall, our results suggest that promoter polymorphism may regulate MMP7 expression, which could therefore modulate GBC susceptibility.

In summary, we found MMP7 promoter variants rs113823671 A>C and rs17098318 G>A as a risk factor for GBC in the population of the Eastern Indian state Odisha. Additionally, we discovered that the variant rs113823671 has an allele-specific functional impact on MMP7 expression levels, which requires further validation through detailed mechanistic studies.



CHAPTER 6

Results and Discussion



Vinay J
NISER, Bhubaneswar

Objective 3

To discover the genetic association of MMP2 promoter SNPs with GBC and genotype-phenotype correlation

Findings of this section has been communicated.

- **Vinay J**, Singh SP, Dixit M. Association and functional significance of MMP2 promoter variants in gallbladder cancer.

6.1 Introduction

Matrix Metalloproteinase 2, also known as Gelatinase A (MMP2, 72kDa Gelatinase, 72kDa Type IV Collagenase, [MIM: 120360]), is a zinc-dependent endopeptidase that facilitates matrix protein and type IV collagen degradation in basement membranes (368). The MMP2 gene is frequently upregulated and associated with the progression of malignancies such as melanoma (15), gliomas (16), breast (18), lung (20), ovarian (21), colon (22), pancreatic (23), hepatocellular (24), and gastric carcinoma (25) progression and positively correlated with tumour metastasis. Also, in a study the MMP2 knock out showed decreased tumour burden, lung metastasis, blood vessel density, and increased survival rate, which emphasize the tumorigenic role in development of metastatic prostatic neuroendocrine cancer (48). Altogether, the evidence points to MMP2 involvement in pathogenesis and poor prognosis of cancer. However, the molecular basis of MMP2 up-regulation in GBC remains unclear.

Many reports on MMP2 promoter polymorphisms suggest their role in regulating gene expression and enzymatic activity of MMPs various diseases including cancer (306, 312, 313). *MMP2* functional promoter variant rs242866 G>A (-1575) and rs243865 C>T (-1306) is reported to be a risk factor for coronary triple-vessel disease (TVD) (309), end-stage kidney disease (ESKD) (308), and metabolic syndrome (MetS) (307). Also, rs243865 C>T (-1306) is a risk factor for intracranial aneurysms (IA) (310) and Alzheimer's Disease (AD) (311). Likewise, *MMP2* functional promoter variants rs243865 C>T (-1306) and rs2285053 C>T (-735) with allele 'C' and 'C,' respectively, show an increased risk for lung (312), nasopharyngeal (313) and esophageal carcinoma (314, 315). The functional variants rs243865 C>T (-1306) and rs2285053 C>T (-735) are in strong linkage disequilibrium (LD) and affect the binding of the Sp1 binding site in an allele-dependent manner. Allele C>T transition at both the loci destroys the Sp1 binding element and significantly reduces the *MMP2* promoter activity (314, 316). The

presence of haplotype T-T (rs243865 and rs2285053, respectively) shows a 3.7-fold decrease in *MMP2* transcription compared to the C-C haplotype in esophageal carcinoma, suggesting allele-dependent regulation of *MMP2* expression (314). However, the role of these *MMP2* promoter variants has yet to be explored in GBC. Therefore, this study aimed to investigate the association of promoter variants rs2285053, rs243864, rs2285052, rs17859821, rs17859816, rs1488656253, rs1391392808, rs17859817, and rs243865 with gallbladder cancer and its putative role in *MMP2* expression.

In the current chapter, we have shown the genetic association of *MMP2* promoter variants with GBC. Also, aberrant expression of *MMP2* in gallbladder tumour tissue suggests its role in GBC pathogenesis. Altogether, the present study emphasizes the functional relevance of promoter variants in expressing *MMP2* in gallbladder cancer.

Results

6.2 Screening of *MMP2* promoter variants in GBC

We sequenced 50 GBC and 50 control samples as screening sets to identify SNPs in the *MMP2* promoter region. Two primer sets were used to cover all variants in the 1500 bases upstream of the *MMP2* transcription start site. During the sequencing, we discovered 12 SNPs, of which 10 showed allelic variation: NC_000016.10: g.55478465C>T (rs2285053), g.55478410T>G (rs243864), g.55478245C>A (rs2285052), g.55478141G>A (rs17859821), g.55477726C>T (rs17859816), g.55477727C>T (rs1488656253), g.55477784C>T (rs1391392808), g.55477785C>T (rs1961996235), g.55477894C>T (rs243865), and g.55477899C>T (rs1961998763). However, the variant g.55477787T>C (rs17859817) was found to be non-polymorphic. Furthermore, we identified a novel variant, NC 000016.10:g.55477735G>A, in our study population.

6.3 Alleles and genotypes frequency in the discovery set

We carried out the analysis of all MMP2 promoter variants using two primer sets. Primer set 1 covered variants rs2285053, rs243864, rs2285052, and rs17859821, while primer set 2 covered SNPs rs17859816, rs1488656253, rs1391392808, rs1961996235, rs243865, and rs1961998763, as well as a novel variant (NC 000016.10:g.55477735G>A). The screening set was used to determine the minor allele frequencies and population-specific effect size or odds of disease exposure.

As shown in Error! Reference source not found., SNPs covered in primer set 1 did not show a significant association with GBC. In contrast, rs1961998763 and the new variant v55477735, G>A, in primer set 2 demonstrated a significant association with GBC in the discovery set, with a p-value < 0.05. These variants were subsequently tested for genetic association with GBC in the validation set. The analysis of allelic and genotype frequencies for each SNP in the discovery set is discussed in the following sections.

In discovery set analysis of primer set 1 SNPs (rs2285053, rs243864, rs2285052, and rs17859821)

The allelic and genotypic frequency distribution of variants rs2285053, rs243864, rs2285052, and rs17859821 are given in Error! Reference source not found..

For rs2285053, the allelic model showed no significant difference in frequency between patients and controls (P = 0.7913, OR 0.87, 95% CI, 0.30-2.48). The genotypic model also did not show a significant association with GBC (P = 0.7773, OR 0.85, 95% CI, 0.28-2.57). Similarly, for rs243864, the allelic model did not show a significant difference in frequency between patients and controls (P = 0.2207, OR 1.66, 95% CI, 0.73-3.74). The genotypic model also did not show a significant association with GBC (P = 0.5598, OR 1.32, 95% CI, 0.53-3.31).

For rs2285052, the allelic model showed no significant difference in frequency between patients and controls (P = 0.1785, OR 1.58, 95% CI, 0.81-3.07). The genotypic model also did

not show a significant association with GBC ($P = 1$, OR 0.99, 95% CI, 0.43-2.27). Similarly, for variant rs17859821, the allelic ($P = 0.7913$, OR 0.87, 95% CI 0.30-2.48) and genotypic models ($P = 0.7773$, OR 0.85, 95% CI=0.28-2.57) showed no significant association with GBC.

Over all, the allelic and genotypic frequency distribution analysis of variants rs2285053, rs243864, rs2285052, and rs17859821 in discovery set showed no significant association with GBC.

Table 6.3.1 Allelic and genotypic frequency distribution of MMP2 promoter variants in primer set 1 (Discovery set)

1	rs2285053	Allele or genotype cou (frequency)		P value	OR (95% CI)
		Patients (N = 50)	Control (N = 50)		
Genotypic model	CC ^{ref}	43 (0.86)	42 (0.84)	0.7773	0.85 (0.28-2.57)
	CT	7 (0.14)	8 (0.16)		
	TT	0 (0.00)	0 (0.00)		
Allelic mod	C ^{ref}	93 (0.93)	92 (0.92)	0.7913	0.87 (0.30-2.48)
	T	07 (0.07)	08 (0.08)		
2					
Genotypic model	TT ^{ref}	35 (0.70)	39 (0.78)	0.5598	1.32 (0.53-3.31)
	TG	13 (0.26)	11 (0.22)		
	TT	2 (0.04)	0 (0.00)		
Allelic mod	T ^{ref}	83 (0.83)	89 (0.89)	0.2207	1.66 (0.73-3.74)
	G	17 (0.17)	11 (0.11)		
3					
Genotypic model	CC ^{ref}	28 (0.56)	31 (0.62)	1	0.99 (0.43-2.27)
	AC	17 (0.34)	19 (0.38)		
	AA	05 (0.10)	0 (0.00)		
Allelic mod	C ^{ref}	73 (0.73)	81 (0.81)	0.1785	1.58 (0.81-3.07)
	A	27 (0.27)	19 (0.19)		
4					
Genotypic model	GG ^{ref}	43 (0.86)	42 (0.84)	0.7773	0.85 (0.28-2.57)
	GA	7 (0.14)	8 (0.16)		
	AA	0 (0.00)	0 (0.00)		
Allelic mod	G ^{ref}	93 (0.93)	92 (0.92)	0.7913	0.87 (0.30-2.48)
	A	07 (0.07)	08 (0.08)		

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

* P value of Fisher's Exact test

The discovery set analysis of primer set 2 SNPs (rs17859816, rs1488656253, rs1391392808, rs1961996235, rs243865, rs1961998763) including a novel variant (NC 000016.10:g.55477735G>A)

An analysis was conducted to evaluate the distribution of allelic and genotypic frequencies of rs17859816, rs1488656253, rs1391392808, rs1961996235, rs243865, rs1961998763, and a novel variant (NC 000016.10:g.55477735G>A) (**Table 6.3.2**).

Variant rs1488656253, rs17859816, rs243865, rs1961996235 and rs1391392808 showed no significant association with the GBC under study in both the genotypic and allelic models. Also, variant rs17859817 was found to be non-polymorphic.

For rs1961998763, genotype 'CT' showed a significant association with GBC ($p = 0.0009$, OR 1.25, 95% CI 1.09-1.44). In the allelic level, the 'T' allele was significantly associated with GBC ($p = 0.0012$, OR 1.11, 95% CI 0.04-1.19).

The novel variant (NC 000016.10:g.55477735G>A), genotype 'GA' showed a significant association with GBC ($p = 0.0309$, OR 9.33, 95% CI 1.12-77.7). Also in the allelic model, the allele 'A' was significantly associated with GBC ($p = 0.0349$, OR 8.61, 95% CI 1.06-70.17).

Overall, rs1961998763 and the novel variant (NC 000016.10:g.55477735G>A) showed a significant association with GBC in the discovery set, with a p -value < 0.05 . The association of these variants with GBC was subsequently tested in the validation set.

Table 6.3.2 Allelic and genotypic frequency distribution of MMP2 promoter variants in primer set 2 (Discovery set)

1	rs1961998763	Allele or genotype count (frequency)		P value	OR (95% CI)
	Alleles/ Genotypes	Patients (N = 50)	Control (N = 50)		
Genotypic model	CC ^{ref}	40 (0.80)	50 (1.00)	0.0009	-
	CT	10 (0.20)	0 (0.00)		
	TT	0 (0.00)	0 (0.00)		
Allelic model	C ^{ref}	90 (0.90)	100 (1.00)	0.0012	-
	T	10 (0.10)	0 (0.00)		
2	rs243865				
Genotypic model	CC ^{ref}	34 (0.68)	39 (0.78)	0.4062	1.49 (0.58-3.83)
	CT	13 (0.26)	10 (0.20)		
	TT	3 (0.06)	01 (0.02)		
Allelic model	C ^{ref}	81 (0.81)	88 (0.88)	0.1715	1.72 (0.79-3.77)
	T	19 (0.19)	12 (0.12)		

3	rs17859817	Non polymorphic			
4	rs1961996235				
Genotypic model	CC^{ref}	49 (0.98)	50 (1.00)		
	CT	1 (0.02)	0 (0.00)	1.0*	-
	TT	0 (0.00)	0 (0.00)	-	-
Allelic model	C^{ref}	99 (0.99)	100 (1.00)	1.0*	-
	T	1 (0.01)	0 (0.00)		
5	rs1391392808				
Genotypic model	CC^{ref}	48 (0.96)	50 (1.00)		
	CT	2 (0.04)	0 (0.00)	0.4949*	-
	TT	0 (0.00)	0 (0.00)	-	-
Allelic model	C^{ref}	98 (0.98)	100 (1.00)	0.4974*	-
	T	2 (0.02)	0 (0.00)		
6	55477735, G>A				
Genotypic model	GG^{ref}	42 (0.84)	49 (0.98)		
	GA	8 (0.16)	01 (0.02)	0.0309*	9.33 (1.12-77.7)
	AA	0 (0.00)	0 (0.00)	-	-
Allelic model	G^{ref}	92 (0.92)	99 (0.99)	0.0349*	8.61 (1.06-70.17)
	A	8 (0.08)	01 (0.01)		
7	rs1488656253				
Genotypic model	GG^{ref}	48 (0.96)	50 (1.00)		
	GA	2 (0.04)	0 (0.00)	0.4949*	-
	AA	0 (0.00)	0 (0.00)	-	-
Allelic model	G^{ref}	98 (0.98)	100 (1.00)	0.4975*	-
	A	2 (0.02)	0 (0.00)		
8	rs17859816				
Genotypic model	GG^{ref}	48 (0.96)	50 (1.00)		
	GA	2 (0.04)	0 (0.00)	0.4949*	-
	AA	0 (0.00)	0 (0.00)	-	-
Allelic model	G^{ref}	98 (0.98)	100 (1.00)	0.4975*	-
	A	2 (0.02)	0 (0.00)		

OR-odds ratio, **CI**-confidence interval, **N**-sample size, *ref*-reference
P value of Pearson Chi-square test

* **P value** of Fisher's Exact test

6.4 Analysis of Hardy-Weinberg equilibrium and study power

The association of primer set 2 variants were subsequently tested in the validation set. The allelic and genotypic frequencies of SNPs in primer set 2 were tested for HWE using a default significance cut-off $p < 0.05$. The SNPs rs1961998763 ($P = 1$), rs243865 ($P = 0.0659$), rs1961996235 ($P = 1$), rs1391392808 ($P = 1$), v55477735 ($P = 0.6912$), rs1488656253 ($P = 1$), and rs17859816 ($P = 1$) passed the HWE test (**Table 6.3.3**). The study power was calculated for the above variants using PS statistical power analysis program. The parameters used to calculate the power of the study were 80% beta power and a 0.05 alpha significance level.

Table 6.3.3 Sample size and HWE calculations for MMP2 promoter SNPs

Sl. No.	SNP	Calculated Sample size*		HWE p-value [#]
		Case	control	
1	rs1961998763	268	268	1
2	rs243865	178	178	0.0659
3	rs1961996235	177	177	1
4	rs1391392808	205	205	1
5	v55477735	231	231	0.6912
6	rs1488656253	208	208	1
7	rs17859816	182	182	1

* PS statistical power program-based analysis at 80% beta power and 0.05 alpha significance level

[#] HWE was calculated based on the study's actual number of control samples.

6.5 Demographic profile of study subjects

In this study, 300 gallbladder cancer patients and 300 healthy controls were included, and their demographic characteristics were listed in **Table 6.3.4**. The mean age of GBC incidence in the study population was found to be 53.54 years ($SD \pm 11.07$), which was significantly higher than the mean age of the control group, 48.46 years ($SD \pm 10.81$), with a p -value < 0.0001 .

Gender-wise analysis of GBC incidence showed a significantly higher frequency of GBC incidence among females (175 cases) compared to male patients (125 cases) with a p-value of 0.0001. Among the control group, there were 119 females and 181 males. These findings suggest that gender and age may be important risk factors for GBC incidence in our study population.

Table 6.3.4 Demographic profile of the gallbladder cancer patients and controls recruited in the study

Characteristics	Cases (n = 300)	Control (n = 300)	P value
Age ^a	53.48 ± 11.07	48.46 ± 10.81	< 0.0001
Gender ^b			
Female	175	119	< 0.0001
Male	125	181	

^a

Student's T-test was used to compare mean values of age

^b Chi-Square test was used to compare the difference in frequency of male and female

6.6 Association of the rs1961998763 variant with increased risk of gallbladder cancer

The allelic and genotypic frequencies of rs1961998763 were compared between 300 GBC patients and 300 healthy controls, as shown in the **Table 6.3.5**. In the genotypic model, the 'CC' genotype was the reference, and the 'CT' genotype was significantly associated with GBC, with a p-value of 1.90E-05 and an OR of 1.06 (95% CI: 1.03-1.10). The 'TT' genotype was not observed in either the patient or control groups. In the allelic model, the 'C' allele was the reference, and the 'T' allele was significantly associated with GBC, with a p-value of 1.91E-05 and an OR of 1.03 (95% CI: 1.02-1.05).

After adjusting for confounding factors such as age and gender, the p-value for both models was 0.9982. It suggests that age and gender are important factors that need to be considered in GBC. It could also indicate that the genetic variant is not directly related to the disease, but rather its association with the disease is mediated through age and gender. However, due to the

lack of the 'T' allele in control group, the adjusted OR for the allelic and genotypic model could not be calculated. Also, further studies are required in large samples to validate these findings.

Table 6.3.5 Allelic and genotypic frequency distribution of rs1961998763 in GBC and control subjects

rs1961998763		Allele or genotype count (frequency)					
	Alleles/ Genotypes	Patients (N = 300)	Control (N = 300)	P value	OR (95% CI)	P _{adjusted}	OR _{adjusted} (95% CI)
Genotypic model	CC ^{ref}	282 (0.94)	300 (1)				
	CT	18 (0.06)	0 (0.0)	1.90E-05	1.06 (1.03-1.10)	0.9982	-
	TT	0 (0.00)	0 (0.00)	-	-	-	-
Allelic model	C ^{ref}	582 (0.97)	600 (1.00)	1.91E-05	1.03 (1.02-1.05)	0.9982	-
	T	18 (0.03)	0 (0.00)				

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P_{adjusted} and OR_{adjusted}, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

6.7 Allele 'T' of variant rs243865 is associated with increased risk of gallbladder cancer

The allelic and genotypic frequencies of rs243865 were compared between GBC patients and healthy controls, as presented in the **Table 6.3.6**. In the genotypic model, the 'CC' genotype was used as a reference, and the 'TT' genotype was found to be significantly associated with GBC, with a p-value of 0.0305 and an OR of 2.25 (95% CI: 1.06-4.75). The 'CT' genotype was not significantly associated with the GBC (p=0.1302, OR=1.33, 95% CI: 0.92-1.94).

In the allelic model, the 'C' allele was the reference, and the 'T' allele was significantly associated with GBC, with a p-value of 0.0087 and an OR of 1.49 (95% CI: 1.11-2.0). In the dominant model, the 'CT' and 'TT' genotypes combined were compared to the 'CC' genotype, and a significant association was observed, with a p-value of 0.0336 and an OR of 1.46 (95% CI: 1.03-2.07). In the recessive model, the 'CC' and 'CT' genotypes combined were compared to the 'TT' genotype, and a significant association was observed, with a p-value of 0.0489 and an OR of 2.08 (95% CI: 0.99-4.37).

After adjusting for confounding factors such as age and gender, the p-values for the genotypic, allelic, dominant, and recessive models changed to 0.0711, 0.0353, 0.0967, and 0.0945, respectively. The ORs for the adjusted models ranged from 1.40 to 2.07, which were consistent with the unadjusted ORs. This suggests that age and gender are important factors that need to be considered in GBC, but their adjustment did not alter the associations between rs243865 and the GBC.

Overall, these findings suggest that the 'TT' genotype and 'T' allele of rs243865 may be risk factors for GBC. However, further studies are needed to confirm these findings and to determine the underlying mechanisms by which rs243865 may contribute to the development of GBC.

Table 6.3.6 Allelic and genotypic frequency distribution of rs243865 in GBC and control subjects

rs243865		Allele or genotype count (frequency)					
	Alleles/ Genotypes	Patients (N = 300)	Control (N = 300)	P value	OR (95% CI)	P ^{adjusted}	OR ^{adjusted} (95% CI)
Genotypic model	CC ^{ref}	196 (0.65)	220 (0.73)				
	CT	82 (0.27)	69 (0.23)	0.1302	1.33 (0.92-1.94)	0.0711	2.07 (0.94-4.58)
	TT	22 (0.07)	11 (0.04)	0.0305	2.25 (1.06-4.75)	0.2376	1.66 (0.72-3.82)
Allelic model	C ^{ref}	474 (0.79)	509 (0.85)	0.0087	1.49 (1.11-2.0)	0.0353	1.40 (1.02-1.91)
	T	126 (0.21)	91 (0.15)				
Dominant model	CC ^{ref}	196 (0.65)	220 (0.73)	0.0336	1.46 (1.03-2.07)	0.0967	1.36 (0.95-1.97)
	CT+TT	104 (0.35)	80 (0.27)				
Recessive model	CC+CT ^{ref}	278 (0.93)	289 (0.96)	0.0489	2.08 (0.99-4.37)	0.0945	1.96 (0.89-4.29)
	TT	22 (0.07)	11 (0.04)				
Additive model	CC vs CT vs TT			0.0457			

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P^{adjusted} and OR^{adjusted}, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

6.8 Association of rs1961996235 variant with increased risk of gallbladder cancer

We investigated the association of the rs1961996235 variant with the risk of GBC (Table 6.3.7). In the genotypic model, the 'CT' genotype was found to be significantly associated with GBC, with a p-value of 0.030 and an OR of 1.02 (95% CI: 1.0-1.04) when compared to the reference 'CC' genotype. No 'TT' genotype was observed in either group.

The allelic model revealed that the 'T' allele was significantly associated with GBC, with a p-value of 0.031 and an OR of 1.01 (95% CI: 1.0-1.02) compared to the reference 'C' allele.

When adjusting for confounding factors such as age and gender, the p-values for the genotypic and allelic models changed to p-values^{adjusted} = 0.9989 and 0.9998, respectively, indicating the importance of considering these factors in GBC analysis. It could also suggest that the genetic variant's association with GBC is mediated through age and gender. However, due to the lack of the 'TT' genotype, the OR^{adjusted} for the genotypic model could not be calculated.

Overall, our findings suggest that the 'CT' genotype and 'T' allele of rs1961996235 may be risk factors for GBC. However, further studies with larger sample sizes are needed to confirm these results.

Table 6.3.7 Allelic and genotypic frequency distribution of rs1961996235 in GBC and control subjects

rs1961996235		Allele or genotype count (frequency)					
	Alleles/ Genotypes	Patients (N = 300)	Control (N = 300)	P value	OR (95% CI)	P ^{adjusted}	OR ^{adjusted} (95% CI)
Genotypic model	CC ^{ref}	294 (0.98)	300 (1)				
	CT	6 (0.02)	0 (0.0)	0.030*	1.02 (1.0-1.04)	0.9989	-
	TT	0 (0.00)	0 (0.00)	-	-	-	-
Allelic model	C ^{ref}	594 (0.99)	600 (1.00)	0.031*	1.01 (1.00-1.02)	0.9998	-
	T	6 (0.01)	0 (0.00)				

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P^{adjusted} and OR^{adjusted}, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

* P value of Fisher's Exact test

6.9 rs1391392808 allele 'T' is associated with an increased risk for gallbladder cancer

The allelic and genotypic frequencies of rs1391392808 were compared between 300 patients with GBC patients and 300 healthy controls, as presented in the **Table 6.3.8**. In the genotypic model, the 'CC' genotype was used as a reference, and the 'CT' genotype was found to be significantly associated with GBC, with a p-value of 0.0005 and an OR of 1.04 (95% CI: 1.02-1.06). The 'TT' genotype was not observed in either group.

In the allelic model, the 'C' allele was the reference, and the 'T' allele was significantly associated with GBC, with a p-value of 0.0005 and an OR of 1.02 (95% CI: 1.01-1.03).

After adjusting for confounding factors such as age and gender, the p-values changed for the genotypic and allelic models with 0.9984 and 0.9985, respectively. This means that age and gender may have been influencing the observed association between rs1391392808 and GBC. It is important to note that this does not necessarily rule out a potential association between rs1391392808 and GBC, but rather highlights the interaction of confounding factors with the variant. Adjusting for these factors revealed that there may not be a significant association between rs1391392808 and GBC. This may be due to lack of the 'T' allele in control group and small sample size. However, further studies with larger sample sizes are needed to validate our findings.

Table 6.3.8 Allelic and genotypic frequency distribution of rs1391392808 in GBC and control subjects

rs1391392808		Allele or genotype count (frequency)					
	Alleles/ Genotypes	Patients (N = 300)	Control (N = 300)	P value	OR (95% CI)	P ^{adjusted}	OR ^{adjusted} (95% CI)
Genotypic model	CC ^{ref}	288 (0.96)	300 (1)				
	CT	12 (0.04)	0 (0.0)	0.0005	1.04 (1.02-1.06)	0.9984	-
	TT	0 (0.00)	0 (0.00)	-	-	-	-
Allelic model	C ^{ref}	588 (0.98)	600 (1.00)	0.0005	1.02 (1.01-1.03)	0.9985	-
	T	12 (0.02)	0 (0.00)				

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P^{adjusted} and OR^{adjusted}, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

6.10 The variant v.55477735G>A is associated with an increased risk for gallbladder cancer

The allelic and genotypic frequencies of v.55477735G>A were compared between 300 patients with gallbladder cancer and 300 healthy controls, as presented in **Table 6.3.9**. In the genotypic model, the 'GG' genotype was used as a reference, and the 'GA' genotype was found to be significantly associated with GBC, with a p-value of 3.02E-05 and an OR of 3.80 (95% CI: 1.95-7.39). The 'AA' genotype was not observed in either group.

In the allelic model, the 'G' allele was the reference, and the 'A' allele was significantly associated with GBC, with a p-value of 4.61E-05 and an OR of 3.59 (95% CI: 1.87-6.91).

After adjusting for confounding factors such as age and gender for the genotypic and allelic models' significance level increased to 6.15E-05 and 9.22E-05, respectively. This suggests that age and gender may have been influencing the observed association between v.55477735G>A and GBC. Adjusting for these factors revealed that the association between v.55477735G>A and GBC significantly increased.

Overall, our findings suggest that the 'GA' genotype and 'A' allele of v.55477735G>A may be risk factors for GBC.

Table 6.3.9 Allelic and genotypic frequency distribution of a novel variant NC_000016.10:g.55477735G>A in GBC and control subjects

g.55477735G>A		Allele or genotype count (frequency)					
	Alleles/ Genotypes	Patients (N = 300)	Control (N = 300)	P value	OR (95% CI)	P ^{adjusted}	OR ^{adjusted} (95% CI)
Genotypic model	GG ^{ref}	259 (0.86)	288 (0.96)				
	GA	41 (0.14)	12 (0.04)	3.02E-05	3.80 (1.95-7.39)	6.15E-05	4.11 (2.06-8.20)
	AA	0 (0.00)	0 (0.00)	-	-	-	-
Allelic model	G ^{ref}	559 (0.93)	588 (0.98)	4.61E-05	3.59 (1.87-6.91)	9.22E-05	3.86 (1.96-7.61)
	A	41 (0.07)	12 (0.02)				

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P^{adjusted} and OR^{adjusted}, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

6.11 Association of rs1488656253 variant with increased risk of gallbladder cancer

The allelic and genotypic frequencies between 300 GBC patients and 300 healthy controls were compared for variant rs1488656253 (**Table 6.3.10**). In the genotypic model, the 'CT' genotype was significantly associated with GBC, with a p-value of 2.88E-05 and an OR of 1.06 (95% CI: 1.03-1.09) when compared to the 'CC' genotype. The 'TT' genotype was not observed in either group. Similarly, the 'T' allele was significantly associated with GBC, with a p-value of 3.29E-05 and an OR of 1.03 (95% CI: 1.02-1.04) when compared to the 'C' allele in the allelic model.

After adjusting for confounding factors such as age and gender, the p-values for the genotypic and allelic models' significance decreased. However, it should be noted that adjusting for these factors revealed that there may not be a direct significant association between rs1488656253 and GBC, indicating the indirect interaction of confounding factors with the variant. We should also make a note of the fact that lack of the 'T' allele in control group therefore further studies are required to confirm the potential association between rs1488656253 and GBC.

Table 6.3.10 Allelic and genotypic frequency distribution of variant rs1488656253 in GBC and control subjects

rs1488656253		Allele or genotype count (frequency)					
	Alleles/ Genotypes	Patients (N = 300)	Control (N = 300)	P value	OR (95% CI)	P ^{adjusted}	OR ^{adjusted} (95% CI)
Genotypic model	CC ^{ref}	283 (0.94)	300 (1)				
	CT	17 (0.06)	0 (0.0)	2.88E-05	1.06 (1.03-1.09)	0.9982	-
	TT	0 (0.00)	0 (0.00)	-	-	-	-
Allelic model	C ^{ref}	583 (0.97)	600 (1.00)	3.29E-05	1.03 (1.02-1.04)	0.9982	-
	T	17 (0.03)	0 (0.00)				

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P^{adjusted} and OR^{adjusted}, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

6.12 Allele 'T' of variant rs17859816 is associated with Gallbladder cancer

We investigated the allelic and genotypic frequencies of rs17859816 in 300 patients with GBC and 300 healthy controls. The results are shown in **Table 6.3.11**.

In the genotypic model, the 'CC' genotype was used as a reference, and the 'CT' genotype was found to be significantly associated with GBC, with a p-value of 5.03E-05 and an OR of 1.06 (95% CI: 1.03-1.09). The 'TT' genotype was not observed in either group. In the allelic model, the 'C' allele was the reference, and the 'T' allele was significantly associated with the GBC, with a p-value of 5.65E-05 and an OR of 1.03 (95% CI: 1.01-1.04).

After adjusting for confounding factors such as age and gender, the p-values for the genotypic and allelic models reduced to 0.9983 and 0.9982, respectively. This means that age and gender may have been influencing the observed association between rs17859816 and GBC. Adjusting for these factors revealed that there may not be a significant association between rs17859816 and GBC. It is important to note that this does not necessarily rule out a potential association between rs17859816 and GBC, but rather highlights the interaction of confounding factors with the variant. Hence, the findings should be required to validate in large samples size to compensate the lack of 'T' allele in control population.

Table 6.3.11 Allelic and genotypic frequency distribution of variant rs17859816 in GBC and control subjects

rs17859816		Allele or genotype count (frequency)					
	Alleles/ Genotypes	Patients (N = 300)	Control (N = 300)	P value	OR (95% CI)	P ^{adjusted}	OR ^{adjusted} (95% CI)
Genotypic model	CC ^{ref}	284 (0.95)	300 (1)				
	CT	16 (0.05)	0 (0.0)	5.03E-05	1.06 (1.03-1.09)	0.9983	-
	TT	0 (0.00)	0 (0.00)	-	-	-	-
Allelic model	C ^{ref}	584 (0.97)	600 (1.00)	5.65E-05	1.03 (1.01-1.04)	0.9982	-
	T	16 (0.03)	0 (0.00)				

OR-odds ratio, CI-confidence interval, N-sample size, *ref*-reference

P value of Pearson Chi-square test

P^{adjusted} and OR^{adjusted}, adjusted P values and odds ratio adjustment for confounding factors gender and age derived from binomial logistic regression analysis

6.13 Haplotype analysis

The haplotype frequencies and their association with the disease were investigated in a haplotype block comprising of five SNPs, including rs17859816, rs1488656253, g.55477735G>A, rs1391392808, and rs1961996235, in 300 patients with the disease and 300 healthy controls. The haplotypes and their frequencies in cases and controls are presented in **Table 6.3.12**.

The haplotype "CCGCC" was the most frequent haplotype in both cases (90%) and controls (98%). The haplotype "CCACC" was found to be significantly associated with the disease (p-value = 4.23E-05) with an OR of 5.4 (95% CI: 2.23-13.3). After the permutation correction the p-value was 0.0142, which suggests that the observed association between the haplotype "CCACC" and the disease is unlikely to have occurred by random chance or sampling error. These findings suggest that the haplotype "CCACC" in haplotype block may be a risk factor for the disease.

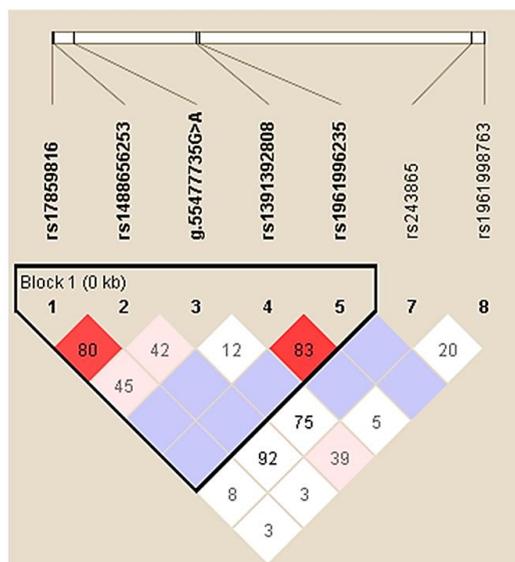
Additionally, the linkage disequilibrium (LD) between the five SNPs was evaluated. The results are presented in the LD plot (Error! Reference source not found.), which shows that the strongest LD was observed between rs17859816 and rs1488656253 ($D' = 0.809$, $LOD = 19.88$, $r^2 = 0.616$), followed by rs1391392808 and rs1961996235 ($D' = 0.831$, $LOD = 7.85$, $r^2 = 0.343$). LD between the remaining SNPs, is generally low, with an r^2 value ranging from 0.061 to 0.051.

These results suggest that the SNPs in haplotype block are in partial LD, with the strongest association observed between rs17859816 and rs1488656253, indicating that they are closely linked. However, LD between the remaining SNPs was relatively weak, suggesting a potential for independent effects on the modulation of risk for GBC.

Table 6.3.12 Haplotype association of the MMP2 promoter variants with gallbladder cancer

Haplotype block									
rs17859816	rs1488656253	g.55477735G>A	rs1391392808	rs1961996235	Frequencies		OR (95% CI)	p- Value	p- Value ^a
					Case (300)	Control (300)			
					C	C	G	C	C
C	C	A	C	C	0.10 (30)	0.02 (6)	5.4 (2.23-13.3)	4.23E-05	0.0142

^a p-Value after permutation correction where N=10000.



Locus 1	Locus 2	D'	LOD	r ²
rs17859816	rs1488656253	0.809	19.88	0.616
rs1391392808	rs1961996235	0.831	7.85	0.343
rs17859816	g.55477735G>A	0.455	4.15	0.061
rs1488656253	g.55477735G>A	0.423	3.9	0.056
g.55477735G>A	rs1961998763	0.395	3.67	0.051 [†]

Figure 6.3.1 Linkage disequilibrium plot for MMP2 promoter variants. Linkage disequilibrium plot is created with the software programme Haploview. Pairwise D' values are used to illustrate linkage disequilibrium (LD). The amplitude and relevance of pairwise LD are shown by shading, with a red-to-white gradient showing higher-to-lower LD values.

6.14 MMP2 expression in gallbladder cancer tissue and its association with promoter variants

The immunohistochemical analysis of 27 gallbladder cancer tissue and adjacent normal tissue (**Figure 6.3.2A**), we observed that MMP2 expression was significantly increased in GBC tissue (mean Allred score 5.96) compared to adjacent normal tissue (mean Allred score 4.43, $p = 0.0124$) (**Figure 6.3.2B**). In order to validate our findings, we further analyzed MMP2 expression at the transcript level using TCGA data set (Project ID: TCGA-CHOL), which revealed a significant increase in MMP2 expression in tumoursamples compared to normal ($p = 0.0291$) (**Figure 6.3.2C**). We also examined the association between MMP2 expression and the genotype of seven promoter SNPs, including rs17859816, rs1488656253, g.55477735G>A,

rs1391392808, rs1961996235, rs243865, and rs1961998763 in GBC patients, to confirm the functional significance of these variants in the regulation of MMP2 gene expression.

Our results showed that patients with the 'TT' genotype at rs243865 had significantly higher MMP2 expression compared to those with the 'CC' genotype, with a mean Allred score of 6.50 ± 0.50 and 4.47 ± 0.14 , respectively. However, the p-value for this association was not statistically significant. Similarly, patients with the 'CT' genotype (mean Allred score of 5.50 ± 0.19) demonstrated increased MMP2 expression relative to patients carrying the 'CC' genotype, with a significant p-value of 0.0006 (**Figure 6.3.3 A**). In addition, we observed a significant increase in MMP2 expression ($p = 0.0115$) in patients carrying the 'GA' genotype at locus g55477735G>A, with a mean Allred score of 4.46 ± 0.22 , compared to individuals with the 'GG' genotype, with a mean Allred score of 5.25 ± 0.19 (**Figure 6.3.3 E**). However, no significant changes in MMP2 expression were observed with respect to individual genotypes of other MMP2 promoter variants (**Figure 6.3.3**).

Overall, our findings suggest that each copy of the rare allele at variants rs243865 and g55477735G>A is associated with a change in MMP2 expression, highlighting the functional significance of these promoter variants in MMP2 gene expression.

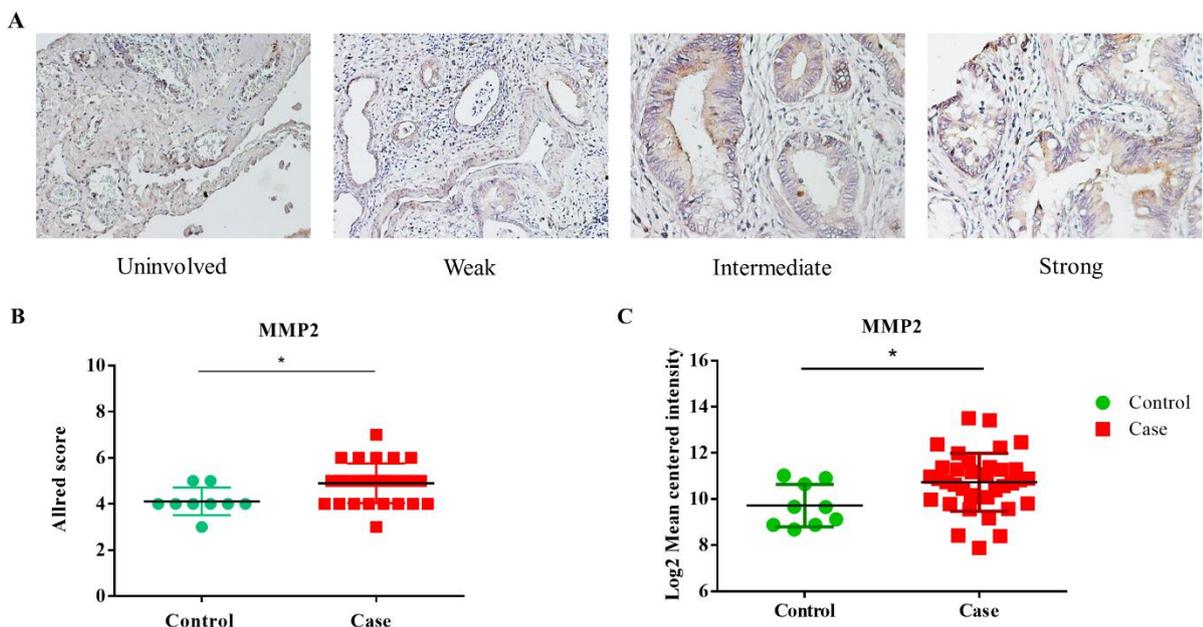
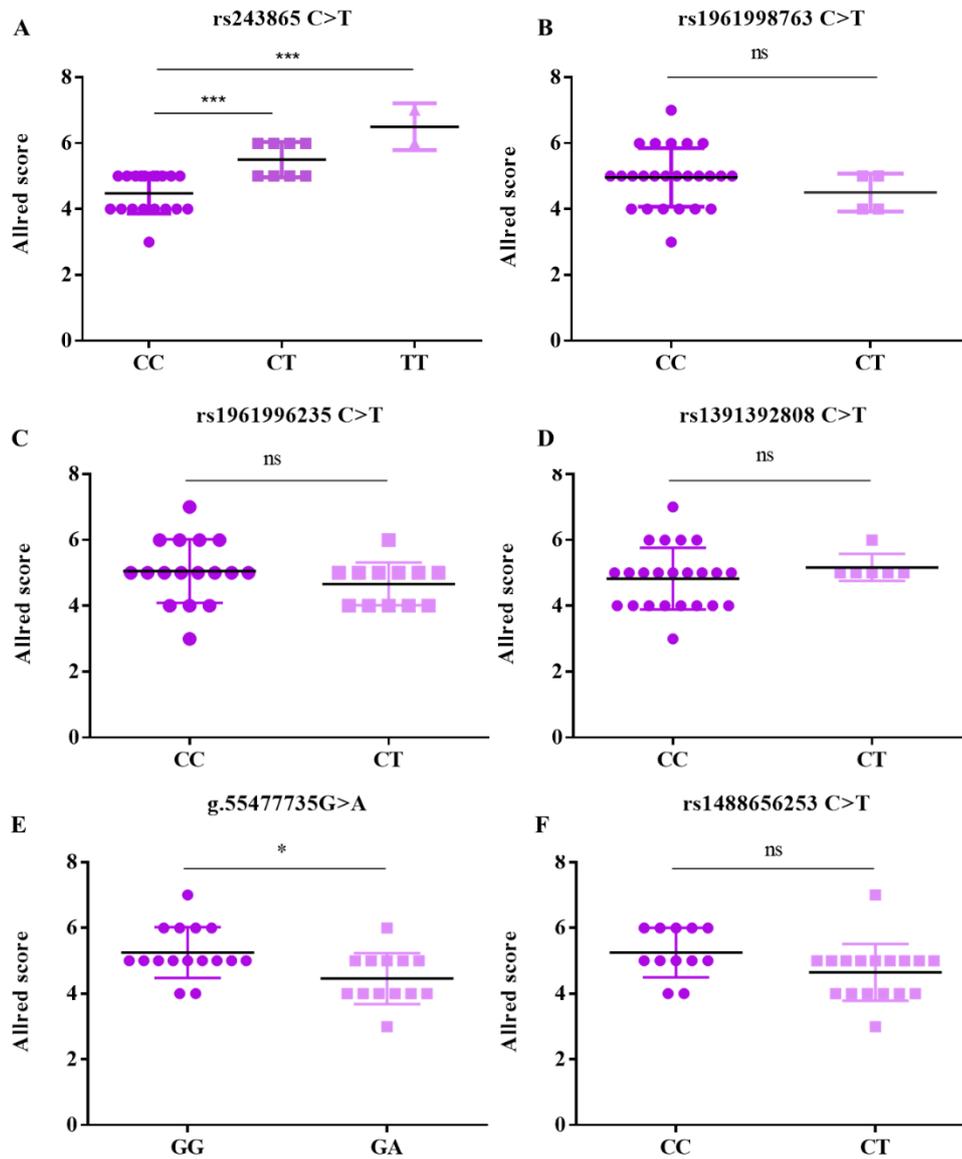


Figure 6.3.2 The MMP2 protein expression in gallbladder tumour tissue and adjacent normal tissue. (A) Representative images of gallbladder cancer tissue IHC with weak or negative, intermediate, and strong staining intensity. The scale bar represents $50 \mu\text{m}$ with $10\times$ objective. (B) MMP2 IHC scoring of gallbladder cancer tissue and adjacent normal tissue

To discover the genetic association of MMP2 promoter SNPs with GBC and genotype-phenotype correlation

samples, X-axis shows sample type, case, and control. Y axis shows Allred score. The unpaired non-parametric two-sided Mann-Whitney test was used to analyze the statistical difference between the group (C) RNA-seq expression data of MMP2 from the TCGA database (TCGA-CHOL), where the X axis represents sample type, case, and control, and the Y axis represents Log2 mean intensity value. Statistical differences between the groups were analyzed by unpaired non-parametric two-sided Mann-Whitney test.



this finding may provide functional evidence for the association between the rs243865C>T locus and MMP2 expression observed in previous studies.

Our results showed a significant association between the 'T' allele of rs243865 and higher MMP2 expression in GBC tissues, which was consistent with our luciferase assay results.

Overall, our findings provide evidence for the regulatory role of the rs243865C>T locus in MMP2 gene expression in GBC. The identification of this regulatory region may have important implications for understanding the molecular mechanisms underlying GBC pathogenesis and for the development of targeted therapies.

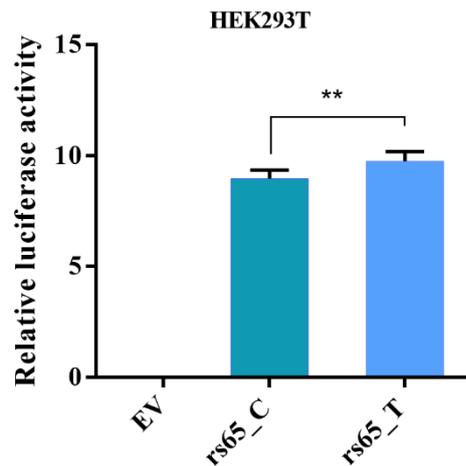


Figure 6.3.4 Allele-specific luciferase activity of MMP2 promoter variant rs243865.

Relative luciferase activity of MMP promoter variants in HEK293T cell lines, X-axis shows pGL4.23 empty vector (EV), pGL4.23 with promoter insert containing allele 'C' and 'T' of rs243865. The Y axis shows the relative luciferase activity of each construct. An unpaired student's t-test was used to study statistical differences between the groups. All experiments were replicated independently three times.

6.16 Discussion

MMP2 is ubiquitously expressed in vascular smooth muscle, cardiomyocytes, endothelial cells, respiratory epithelial cells, macrophages, fibroblasts, and stromal cells in many tissues (369). MMP2 gene expression is associated with poor prognosis and metastasis of lung, colon, breast, ovarian, and renal cell carcinoma (370). Although many studies reported the role of MMP2 in pathogenesis and prognosis in multiple cancer, there were limited studies available on GBC. We found a significant increase in MMP2 protein levels in GBC patients. In the current study, we explored the association of *MMP2* promoter variants with gallbladder cancer and their possible role in regulating MMP2 expression by affecting the transcription factor binding site.

Earlier study in asthma showed, two MMP2 regulatory variants (rs243865C>T and rs2285053C>T) have been found to be risk factors, as they increase the MMP2 expression (371). Similarly, these variants have been associated with increased susceptibility to nasopharyngeal carcinoma (313) and lung cancer (312) in Chinese populations. Two independent studies have shown that these MMP2 variant rs243865 C>T (-1306) has allele-specific regulation of MMP2 expression by disrupting the binding site for the Sp1 transcription factor (314, 316). Additionally, the presence of a specific haplotype (C-C) (rs243865-rs2285053) in esophageal cancer patients has been shown to have an additive effect on MMP2 mRNA levels compared to a haplotype (T-T) (314). These findings emphasize the functional implications of these variants and further highlight the need to investigate their mechanistic role in gallbladder cancer.

The current study explored the potential association and functional significance of MMP2 promoter variants in GBC. Our findings indicate that individuals with the 'T' allele at MMP2 promoter variants rs17859816, rs1488656253, rs1391392808, rs1961996235, rs243865, and rs1961998763 are more susceptible to GBC than those with the 'C' allele in the population of

East Indian state, Odisha. Additionally, we identified a novel single nucleotide polymorphism at loci NC_000016.10:g.55477735G>A, with the 'A' allele significantly associated with GBC. The allelic interaction of all associated promoter variants followed a dominant model, highlighting that even a single copy of the risk allele 'T' could pose a risk for GBC. In our study, the patients carrying the 'T' allele at rs243865 have an increased risk for gallbladder cancer (OR 1.49). These findings are in parallel to studies in other cancers including, prostate (OR 1.52) (372), bladder (OR 1.76) (373), and oral cancer (OR 1.46) (374). However, many studies have contradicting findings showing 'T' allele at rs243865, as protective to esophageal (OR 0.95) (314), lung (OR 0.53) (312), head and neck cancer (OR 0.53) (375). Interestingly, our results were consistent with a recent meta-analysis in which Asian populations with the 'T' allele had a higher risk (OR 1.48) for cancer than the 'C' allele (376). These population-specific changes in allelic predisposition suggest that specific environmental and ethnic factors play a role in developing different cancer types.

Furthermore, haplotype analysis demonstrated that the C-C-A-C-C haplotype (rs17859816, rs1488656253, and g.55477735G>A, rs1391392808, rs1961996235) was significantly associated with GBC, with g.55477735G>A variant serving as a high-risk contributor towards GBC pathogenesis. While genotype-phenotype analysis and luciferase assay were carried out to assess the functional relevance of MMP2 promoter SNPs. The genotype-phenotype analysis found no significant difference with other SNPs, except for variant rs243865, which showed an increase in MMP2 expression with each copy of the risk allele 'T'. This observation was parallel to our case-control association study. The role of rs243865 was further validated through a luciferase assay, where risk allele 'T' showed a significant increase in luciferase activity. Moreover, we identified the 'T' allele of rs243865, a risk allele for gallbladder cancer, was found to increase MMP2 expression independently of Sp1 binding, as it disrupts the Sp1 binding site (316). Sp1 belongs to a family of C2H2 zinc finger transcription factors and has a dual role in

activating and suppressing genes based on an RNA binding motif in the 3' UTR region of mRNA (377).

To fully understand the molecular mechanism of MMP2 regulation in gallbladder cancer, further in-depth in vitro and in vivo studies are needed. Our data suggest that promoter variants are risk factor for gallbladder cancer and may contribute to its susceptibility. We identified a novel genetic variant, NC_000016.10:g.55477735G>A, and a de novo association of MMP2 promoter variants, rs17859816C>T, rs1488656253C>T, rs1391392808 C>T, rs1961996235C>T, rs243865C>T, and rs1961998763C>T, with GBC in our study population. However, future studies in multiple large cohorts are needed to better understand these promoter variants' functional relevance.



CHAPTER 7

Results and Discussion



Vinay J
NISER, Bhubaneswar



Objective 4

To find out MMP associated pathways altered in GBC using publicly available databases



Vinay J
NISER, Bhubaneswar

7.1 Introduction

GBC is a highly aggressive malignancy with poor prognosis and limited treatment options (7). The complex interplay between genetic and epigenetic alterations, as well as alterations in cellular signalling pathways, contributes to the development and progression of GBC (378, 379). MMPs are a family of zinc-dependent endopeptidases that play important roles in extracellular matrix remodeling, cell migration, and invasion (380). Dysregulation of MMP expression and activity has been implicated in the pathogenesis of several cancers, including GBC (381). MMP2, MMP7, and MMP14 have been shown to be overexpressed in GBC and may play a key role in the pathogenesis of this disease (51). However, the molecular pathways and biological processes that are regulated by these MMPs in GBC remain largely unknown.

To gain insights into the biological pathways altered in GBC that are associated with MMPs, we performed pathway enrichment analysis using publicly available databases. We utilized the TCGA and GEO datasets to investigate the gene expression profiles of GBC patients and compared them with normal samples. We also used R program DESeq2, GSEA, DAVID Bioinformatics Resources, Cytoscape and EnrichmentMap, UCSC Xena browser, and cBioportal to identify and analyze the molecular pathways and networks associated with MMPs. Additionally, we used SMART App to investigate the epigenetic contribution of MMP2, MMP7, and MMP14 in GBC.

Our analysis revealed several key pathways and biological processes that are dysregulated in GBC and are associated with MMP expression. These pathways comprise ECM remodeling, angiogenesis, cell adhesion, cytoskeleton organization, and immune response. Epigenetic analysis of MMP2, MMP7, and MMP14 showed a reduction in methylation levels in GBC samples, with a negative correlation observed between MMP7 and MMP14 transcript levels

and methylation status in GBC patients. These findings emphasize the potential epigenetic contribution of MMPs in GBC pathogenesis.

Our study aims to provide a comprehensive analysis of the pathways altered in GBC that are associated with MMPs. By identifying these pathways, we can gain a better understanding of the molecular mechanisms underlying the pathogenesis of GBC and potentially identify new therapeutic targets and biomarkers for this disease. Additionally, our study highlights the utility of publicly available databases and bioinformatics tools in investigating complex biological systems and can serve as a roadmap for other researchers investigating the role of MMPs in GBC.

Results

7.2 Differential Gene Expression Analysis of RNA-Seq data from GEO database

To investigate the common pathways altered in GBC, we performed a bioinformatics analysis using publicly available RNA-Seq data from two distinct studies in the NCBI Gene Expression Omnibus (GEO) database (accession number: GSE139682 and GSE132223), which contained primary and metastatic GBC samples. After normalization of raw counts using the DESeq2 package, we obtained a list of differentially expressed genes (DEGs), representing the top 25% of the total differentially expressed genes. We then performed pathway and network analysis using DAVID and GSEA tools.

7.3 Enrichment analysis of DEGs in primary GBC patients

The present study analysed the functions and pathway enrichment of candidate differentially expressed genes (DEGs) in primary GBC patients using the DAVID website. The DEGs were categorized into three functional datasets, namely, cellular components, molecular functions, and biological processes. The top 20 enriched pathways in each dataset were listed in **Table**

7.3.1 and **Figure 7.3.1**. The biological process group showed enrichment in various pathways, including angiogenesis, cell adhesion, signal transduction, positive regulation of angiogenesis, and transmembrane receptor protein tyrosine kinase signalling pathway, among others. In the cellular component group, the DEGs were mainly enriched in the plasma membrane, extracellular matrix, cell surface, cell-cell junction, and receptor complex, among others. In the molecular function group, protein binding, integrin binding, calcium ion binding, and heparin binding, among others, were the main areas of enrichment. It is important to note that the processes of integrin binding, cell adhesion, ECM remodelling, and cytoskeleton organization are interconnected and play a significant role in regulating cellular behaviour and affecting the tumour microenvironment (382) The analysis identified several key pathways and biological processes that are dysregulated in GBC and are associated with MMP expression and activity.

Additionally, we conducted enrichment analysis of DEGs in Disease dataset and three pathways datasets, namely, KEGG pathways, Reactome pathways, and Wiki pathways (**Figure 7.3.1**). The DEGs in KEGG pathways were mainly enriched in Ras signalling, calcium signalling, tight junction, focal adhesion, and cell adhesion. The Wiki pathways showed enrichment in focal adhesion, burn wound healing, and angiogenesis, while Reactome pathways showed enrichment in extracellular matrix organization. The analysis emphasized that these pathways are downstream effectors of MMP and may play a critical role in GBC progression initial stage of cancer.

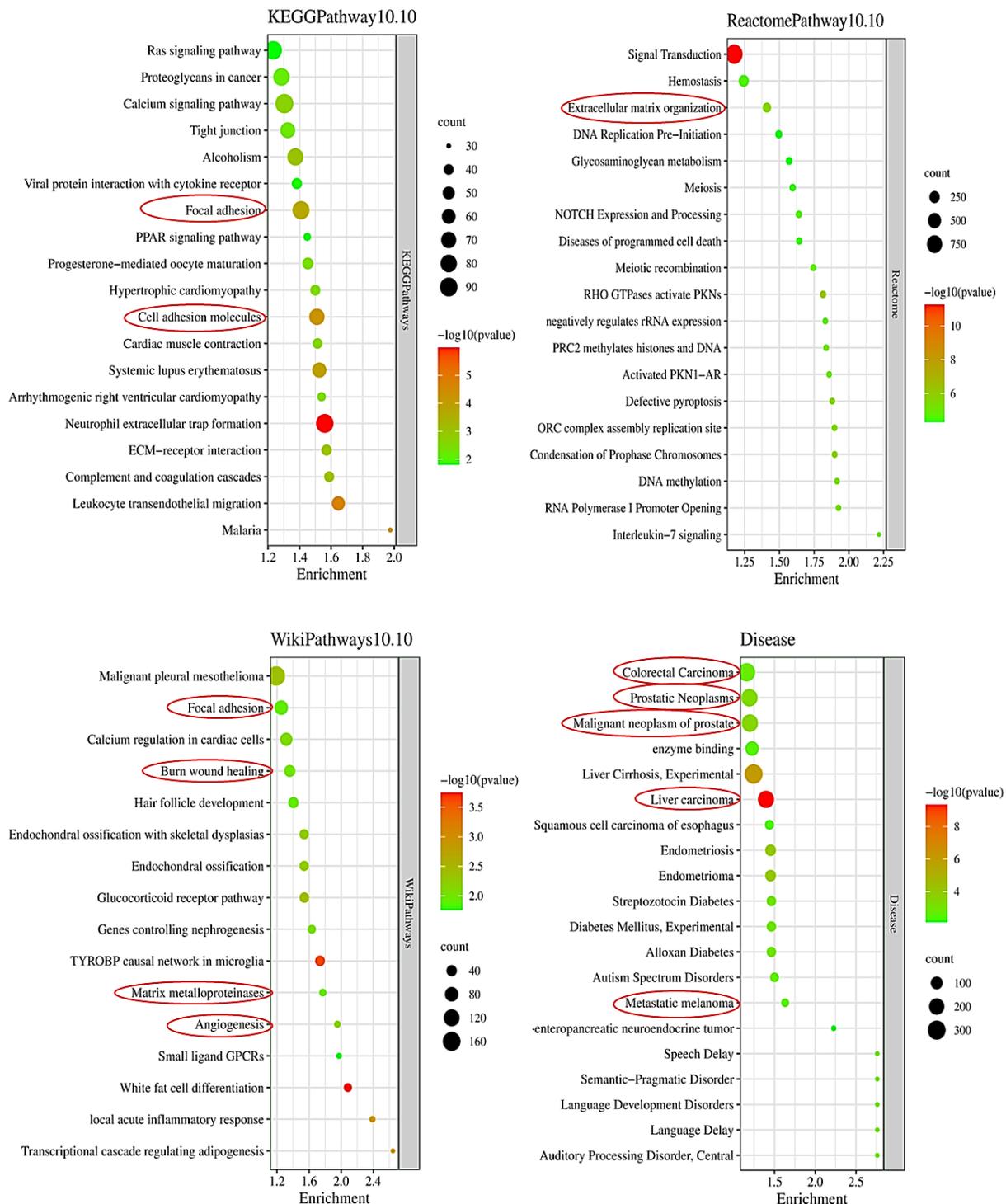


Figure 7.3.1 The Bubble plot data of functional enrichment analysis of primary GBC patients from DAVID web database. The figure shows a bubble plot representing the enrichment of differentially expressed genes in Disease dataset and three pathways datasets including KEGG pathways, Reactome pathways, and Wiki pathways. The enrichment score is represented on the horizontal axis while the vertical axis shows the pathway name. The size of the bubble represents the number of differential genes enriched in each pathway, and the colour

To find out MMP associated pathways altered in GBC using publicly available databases

represents the log p-value. The smaller the log p-value, the closer the colour is to the green colour and higher the value red colour. The top 20 enriched pathways are showed in the bubble plot.

Table 7.3.1 Gene ontological classifications of differentially expressed genes in primary GBC patients

Term	Description	Count	PValue	FDR
Biological processes				
GO:0001525	Angiogenesis	127	1.59E-12	1.40E-08
GO:0007155	Cell adhesion	230	4.49E-10	1.98E-06
GO:0007165	Signal transduction	460	6.73E-09	1.98E-05
GO:0098609	Cell-cell adhesion	92	2.81E-08	6.19E-05
GO:0045766	Positive regulation of angiogenesis	81	7.03E-08	0.0001
GO:0007169	Transmembrane receptor protein tyrosine kinase signalling pathway	64	4.18E-06	0.0062
GO:0071625	Vocalization behavior	14	2.27E-05	0.0286
GO:0008284	Positive regulation of cell proliferation	203	4.27E-05	0.0471
GO:0008360	Regulation of cell shape	69	5.77E-05	0.0566
GO:0001570	Vasculogenesis	34	8.34E-05	0.0691
GO:0007043	Cell-cell junction assembly	24	9.30E-05	0.0691
GO:0007275	Multicellular organism development	90	0.0001	0.0691
GO:0042493	Response to drug	116	0.0001	0.0691
GO:0018108	Peptidyl-tyrosine phosphorylation	62	0.0001	0.0842
GO:0070830	Bicellular tight junction assembly	28	0.0003	0.1692
GO:0001938	Positive regulation of endothelial cell proliferation	38	0.0003	0.1705
GO:0071456	Cellular response to hypoxia	61	0.0003	0.1705
GO:0061028	Establishment of endothelial barrier	15	0.0004	0.1907
GO:0001701	In utero embryonic development	86	0.0004	0.1915
GO:0008277	Regulation of G-protein coupled receptor protein signalling pathway	26	0.0004	0.1918
Cellular components				
GO:0005886	plasma membrane	1699	3.32E-18	4.36E-15
GO:0031012	extracellular matrix	128	3.87E-12	2.54E-09
GO:0005576	extracellular region	749	9.00E-12	3.94E-09
GO:0005887	integral component of plasma membrane	545	1.21E-10	3.97E-08
GO:0016020	membrane	870	2.26E-10	5.93E-08
GO:0009986	cell surface	256	5.08E-10	1.11E-07
GO:0005911	cell-cell junction	93	1.45E-08	2.73E-06
GO:0043235	receptor complex	102	1.97E-08	3.24E-06
GO:0070062	extracellular exosome	747	7.64E-08	1.12E-05
GO:0042383	sarcolemma	57	1.67E-07	2.19E-05
GO:0009897	external side of plasma membrane	180	7.80E-07	9.02E-05
GO:0005829	cytosol	1717	8.23E-07	9.02E-05
GO:0016324	apical plasma membrane	148	6.72E-06	0.0007
GO:0005925	focal adhesion	163	1.54E-05	0.0014

GO:0070161	anchoring junction	181	2.21E-05	0.0019
GO:0014069	postsynaptic density	109	2.62E-05	0.0022
GO:0043197	dendritic spine	74	3.39E-05	0.0026
GO:0097060	synaptic membrane	22	6.35E-05	0.0046
GO:0005737	cytoplasm	1712	0.0001	0.0084
GO:0005856	cytoskeleton	196	0.0001	0.0084
Molecular functions				
GO:0005515	protein binding	3993	1.61E-16	4.34E-13
GO:0005178	integrin binding	80	2.43E-07	0.0003
GO:0005509	calcium ion binding	285	1.59E-06	0.0014
GO:0008201	heparin binding	82	4.46E-06	0.0029
GO:0030527	structural constituent of chromatin	52	5.44E-06	0.0029
GO:0019901	protein kinase binding	198	1.27E-05	0.0057
GO:0005201	extracellular matrix structural constituent	64	6.59E-05	0.0254
GO:0042803	protein homodimerization activity	260	0.0001	0.0357
GO:0017147	Wnt-protein binding	20	0.0002	0.0464
GO:0050840	extracellular matrix binding	20	0.0003	0.0776
GO:0017046	peptide hormone binding	25	0.0006	0.1374
GO:0038023	signalling receptor activity	93	0.0008	0.1699
GO:0005539	glycosaminoglycan binding	17	0.0008	0.1699
GO:0005096	GTPase activator activity	110	0.0015	0.2719
GO:0030165	PDZ domain binding	42	0.0016	0.2719
GO:0003700	transcription factor activity, sequence-specific DNA binding	199	0.0016	0.2719
GO:0019838	growth factor binding	21	0.0018	0.2931
GO:0045296	cadherin binding	119	0.0026	0.3913
GO:0004714	transmembrane receptor protein tyrosine kinase activity	50	0.0032	0.4289
GO:0019899	enzyme binding	143	0.0032	0.4289

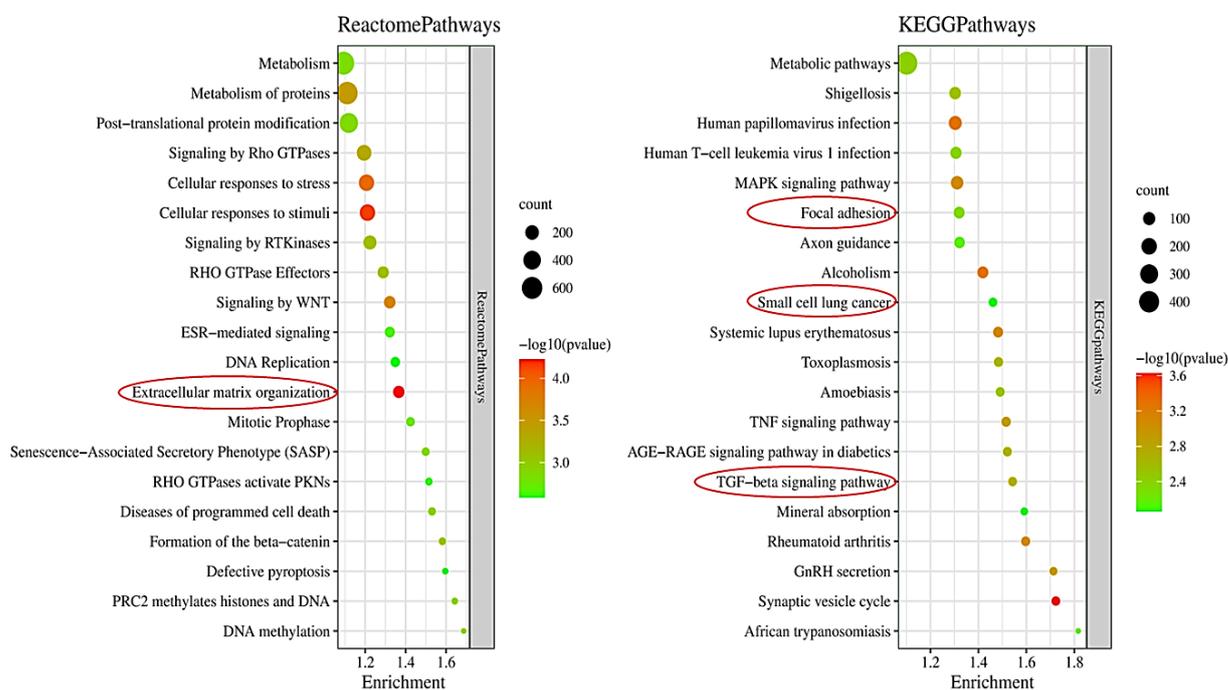
7.4 Enrichment analysis of DEGs in metastatic GBC patients

Similarly, we performed an enrichment analysis of differentially expressed genes (DEGs) in metastatic GBC patients using the DAVID website. The DEGs were classified into three functional datasets, namely cellular components, molecular functions, and biological processes. Top 20 enriched pathways in each dataset are listed in **Table 7.4.1** and **Figure 7.4.1**. In the biological process group, the DEGs were mainly enriched in immune response, signal transduction, angiogenesis, response to drug, bile acid signalling pathway, and cell adhesion, among others. The cellular component group revealed that the aberrantly expressed genes were mainly enriched in the extracellular exosome, plasma membrane, extracellular matrix, and cell surface. In the molecular function group, protein binding, integrin binding, iron ion binding, collagen binding, actin filament binding, growth factor activity, and transmembrane signalling receptor activity were the main areas of enrichment. Our analysis in metastatic sample also

To find out MMP associated pathways altered in GBC using publicly available databases

showed that several key pathways and biological processes were dysregulated in GBC and associated with MMP expression.

Moreover, the enrichment analysis of DEGs was performed in the Disease dataset and three pathways datasets including KEGG, Reactome, and Wiki pathways (**Figure 7.4.1**). KEGG pathways were mainly enriched in MAPK signalling, TNF signalling, TGF- β signalling, focal adhesion, and small cell lung cancer. Wiki pathways showed enrichment in focal adhesion, burn wound healing, TGF- β signalling and endothelial ossification. Reactome pathways showed enrichment in extracellular matrix organization. These pathways are downstream effectors of MMP and may play a critical role in GBC progression and metastasis.



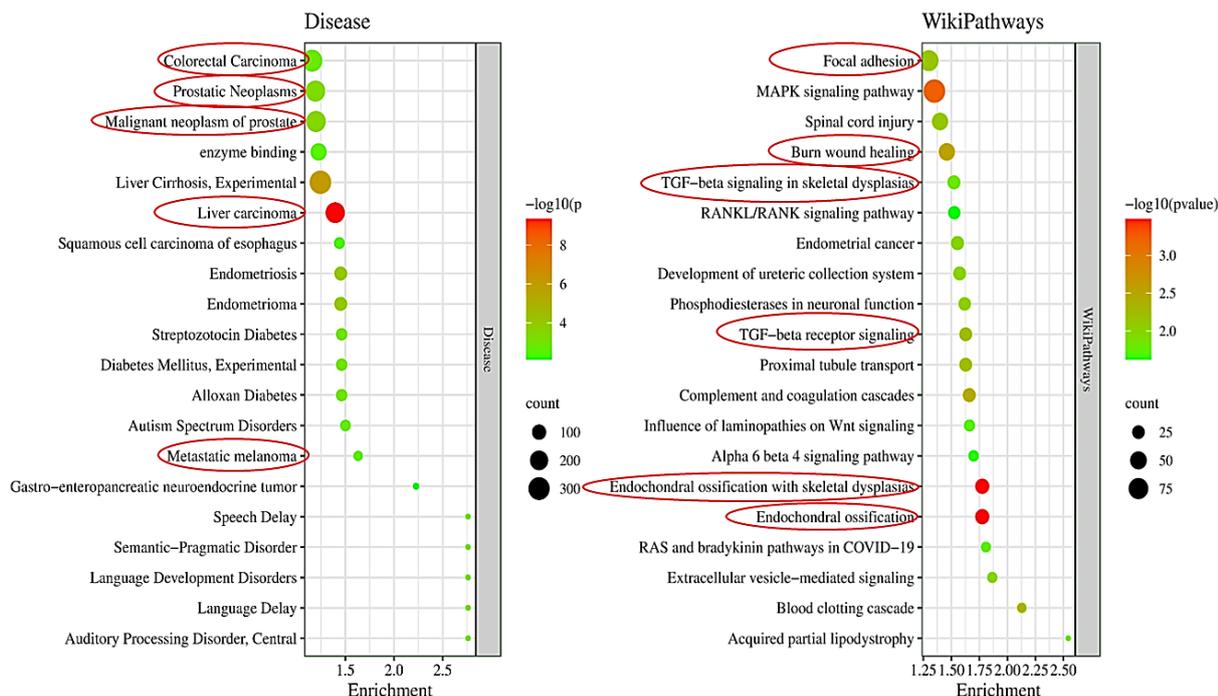


Figure 7.4.1 Bubble plot analysis data from DAVID web database showing functional enrichment analysis of metastatic GBC patients. The figure displays a bubble plot representing the enrichment of differentially expressed genes in Disease dataset and three pathways datasets including KEGG, Reactome, and Wiki pathways. The pathway name is depicted on the vertical axis, whereas the horizontal axis represents the enrichment score. The log p -value is indicated by the size of the bubble, which is color-coded. A smaller log p -value is associated with a greener color, while a higher log p -value is represented by a redder color. The size of each bubble is proportional to the number of differential genes enriched in the respective pathway. The top 20 enriched pathways are presented in the bubble plot.

Table 7.4.1 Gene ontological classifications of differentially expressed genes in metastatic GBC patients using DAVID web database

Term	Discription	Count	P Value	FDR
Biological processes				
GO:0006955	immune response	231	2.66E-16	2.39E-12
GO:0002250	adaptive immune response	203	3.13E-12	1.41E-08
GO:0002377	immunoglobulin production	60	2.43E-11	7.28E-08
GO:0006805	xenobiotic metabolic process	59	9.91E-10	2.23E-06
GO:0006958	complement activation, classical pathway	62	3.14E-06	0.0056
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	425	5.00E-06	0.0075
GO:0007165	signal transduction	445	7.15E-06	0.0092
GO:0001934	positive regulation of protein phosphorylation	92	1.12E-05	0.0126
GO:0006956	complement activation	19	1.42E-05	0.0128
GO:0001525	angiogenesis	107	1.42E-05	0.0128

To find out MMP associated pathways altered in GBC using publicly available databases

GO:0042493	response to drug	121	1.81E-05	0.0148
GO:0038183	bile acid signalling pathway	13	2.12E-05	0.0159
GO:0006882	cellular zinc ion homeostasis	23	2.30E-05	0.0159
GO:0008202	steroid metabolic process	31	3.61E-05	0.0232
GO:0001649	osteoblast differentiation	59	4.88E-05	0.0292
GO:0007507	heart development	92	7.76E-05	0.0436
GO:0007155	cell adhesion	205	0.0002	0.0837
GO:0008203	cholesterol metabolic process	41	0.0002	0.0873
GO:0006954	inflammatory response	156	0.0002	0.0873
GO:0006536	glutamate metabolic process	12	0.0002	0.0873
Cellular components				
GO:0005576	extracellular region	853	4.96E-30	6.61E-27
GO:0072562	blood microparticle	94	2.22E-18	1.48E-15
GO:0009897	external side of plasma membrane	216	2.65E-16	1.18E-13
GO:0005615	extracellular space	709	1.77E-13	5.91E-11
GO:0019814	immunoglobulin complex	59	3.73E-10	9.94E-08
GO:0005737	cytoplasm	1805	1.10E-09	2.45E-07
GO:0070062	extracellular exosome	769	3.66E-09	6.97E-07
GO:0005886	plasma membrane	1635	4.42E-08	7.37E-06
GO:0016020	membrane	848	1.48E-06	0.0002
GO:0005782	peroxisomal matrix	33	2.40E-06	0.0003
GO:0005788	endoplasmic reticulum lumen	127	4.89E-06	0.0006
GO:0009986	cell surface	239	5.44E-06	0.0006
GO:0031012	extracellular matrix	110	6.51E-06	0.0007
GO:0005777	peroxisome	53	8.66E-05	0.0082
GO:0005829	cytosol	1710	0.0001	0.0104
GO:0015629	actin cytoskeleton	104	0.0001	0.0121
GO:0070161	anchoring junction	177	0.0003	0.0210
GO:0034364	high-density lipoprotein particle	18	0.0003	0.0210
GO:0032991	macromolecular complex	248	0.0004	0.0235
GO:0042101	T cell receptor complex	63	0.0004	0.0235
Molecular functions				
GO:0005515	protein binding	3993	1.54E-10	4.44E-07
GO:0003823	antigen binding	79	3.21E-09	4.62E-06
GO:0042802	identical protein binding	605	2.44E-06	0.0023
GO:0004888	transmembrane signalling receptor activity	88	3.75E-06	0.0027
GO:0008201	heparin binding	82	8.03E-06	0.0046
GO:0042803	protein homodimerization activity	270	1.06E-05	0.0051
GO:0046982	protein heterodimerization activity	150	3.15E-05	0.0129
GO:0004497	monooxygenase activity	41	4.08E-05	0.0139
GO:0005102	receptor binding	158	4.34E-05	0.0139
GO:0005506	iron ion binding	66	7.37E-05	0.0205
GO:0005178	integrin binding	73	7.90E-05	0.0205
GO:0016491	oxidoreductase activity	102	8.56E-05	0.0205
GO:0071949	FAD binding	25	9.67E-05	0.0214
GO:0005201	extracellular matrix structural constituent	64	0.0001	0.0214
GO:0051015	actin filament binding	95	0.0002	0.0324
GO:0008083	growth factor activity	73	0.0002	0.0449

GO:0005518	collagen binding	35	0.0005	0.0898
GO:0015347	sodium-independent organic anion transmembrane transporter activity	15	0.0006	0.1025
GO:0020037	heme binding	65	0.0009	0.1364
GO:0016705	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	30	0.0019	0.2680

7.5 GSEA

In order to validate the findings from the DAVID analysis of differentially expressed genes between primary and metastatic GBC samples, a Gene Set Enrichment Analysis (GSEA) was performed using curated gene sets in the Molecular Signatures Database (MSigDB), consisting of CGP (chemical and genetic perturbations, 3405-gene sets) and CP (canonical pathways, 3090-gene sets). The list of canonical pathways gene sets includes CP-BIOCARTA (BioCarta gene sets, 292-gene sets), CP-KEGG (KEGG gene sets, 186-gene sets), CP-PID (PID gene sets, 196-gene sets), CP-REACTOME (Reactome gene sets, 1654-gene sets), and CP-WIKIPATHWAYS (WikiPathways gene sets, 733-gene sets). In the primary GBC group, 149 gene sets were significantly enriched and linked by 1556 edges, while in the metastatic group, 108 gene sets were significantly enriched and linked by 500 edges. The results of GSEA with $FDR < 0.05$ and $p < 0.001$ were visualized using the Enrichment Map plugin available in Cytoscape (**Figure 7.5.1**). The results of GSEA analysis with the top 20 gene sets involved in primary samples and metastatic GBC samples are presented in **Table 7.5.1** and **Table 7.5.2**, respectively.

In primary GBC samples, we observed the enrichment of gene sets involved in cell cycle, proliferation, and tumorigenesis (**Table 7.5.1, Figure 7.5.1A to C and Figure 7.5.2A**). These gene sets showed high redundancy, as indicated by tight clusters of nodes and edges with overlapping genes, particularly for pathways associated with tumour progression and metastasis, epigenetic regulation, and cell cycle (**Figure 7.5.1B and C**, red coloured Nodes and blue coloured Edge lines). The most significant enrichment was seen in the cervical cancer proliferation cluster (NES = 3.15, FDR < 0.001 and rank at max = 2016), cholangiocarcinoma cluster (NES = 2.92, FDR < 0.001 and rank at max = 2182), breast cancer grade 1 vs 3 up (NES = 2.92, FDR < 0.001 and rank at max = 2182), and liver cancer subclass proliferation up (NES = 2.57, FDR = 0.0003 and rank at max = 1984) (**Table 7.5.1 and Figure 7.5.2A**).

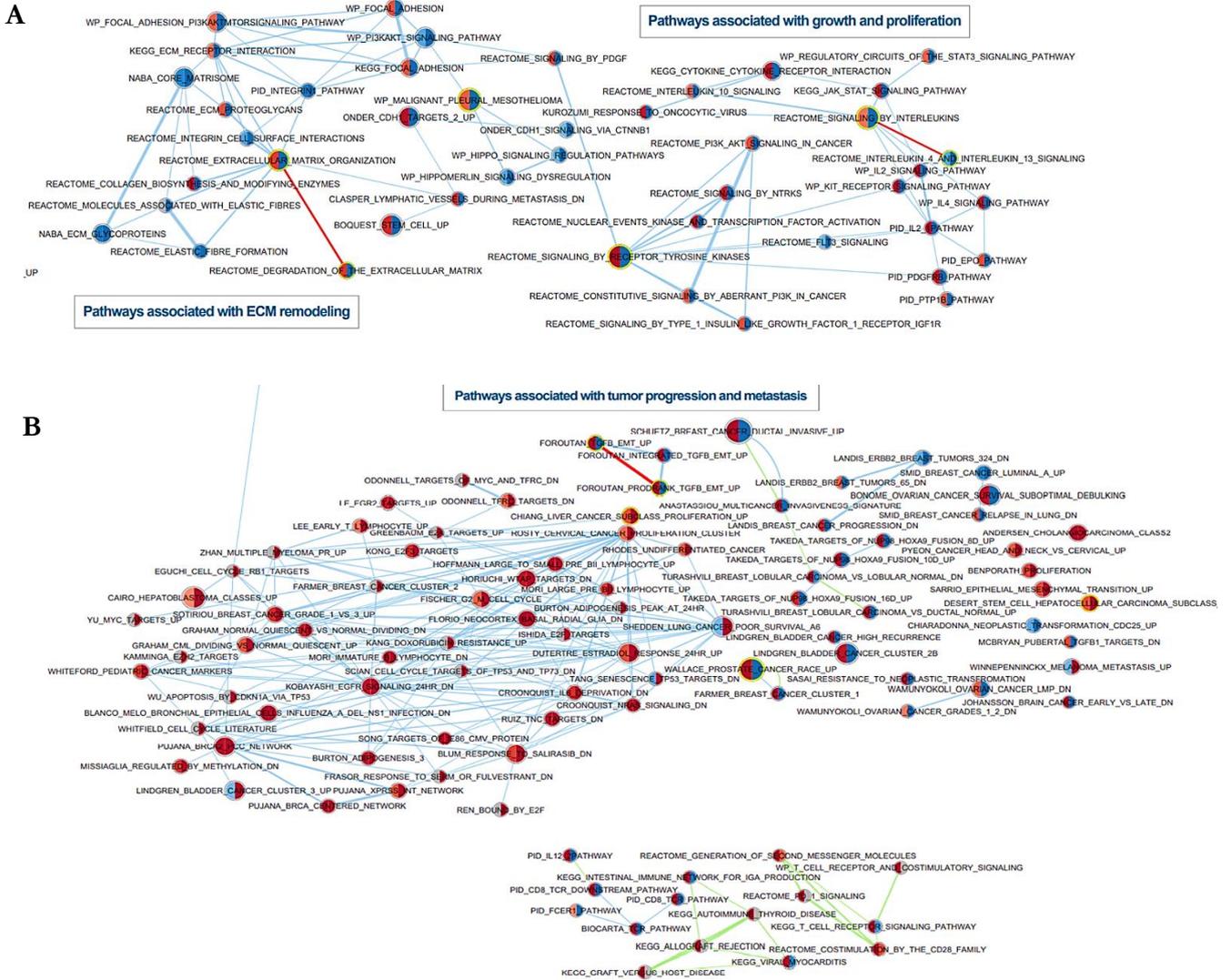
In metastatic GBC samples, gene sets with significant enrichment were involved in inflammation, immune-response, and pathways associated with metabolism (**Table 7.5.2, Figure 7.5.1D and E**). We also observed gene sets enriched in receptor tyrosine kinases (RTKs), TGF- β signalling, ECM degradation, and organization pathways (**Figure 7.5.1A to B**). These gene sets showed high redundancy, as indicated by tight clusters of nodes and edges with overlapping genes (**Figure 7.5.1D and E**, indicated with red/blue colored nodes and green colored edge lines). Among these most significant enrichment was seen in inflammation and immune response gene sets, such as CD22 mediated B cell receptor (BCR) regulation (NES = 3.34), antigen activated BCR (NES = 3.31), Fc ϵ RI mediated MAPK activation (NES = 3.22), Fc ϵ RI mediated CA²⁺ mobilization (NES = 3.12), role of LAT2/NTAL/LAB on calcium mobilization (NES = 3.22), Fc-gamma receptor (Fc γ R) activation (NES = 3.14) and Fc γ R dependent phagocytosis (NES = 3.16) (**Table 7.5.2 and Figure 7.5.1E**).

Table 7.5.1 Gene set enrichment analysis showing top gene-sets in GBC with primary samples, ranked according to NES

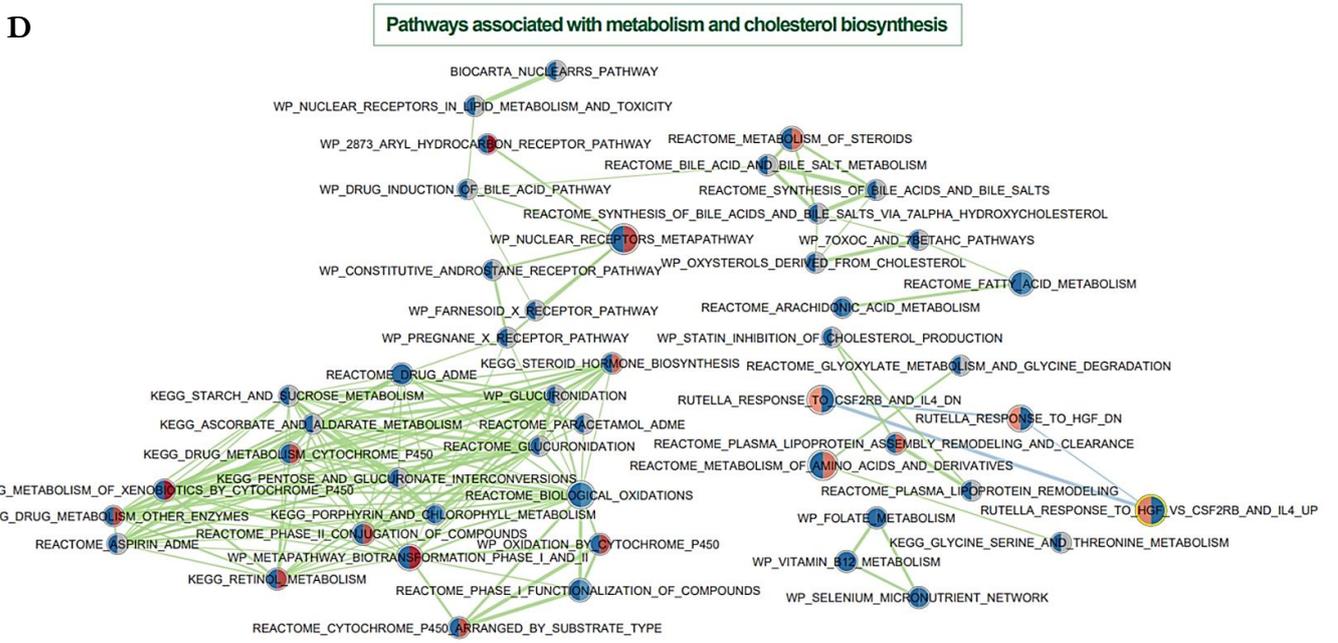
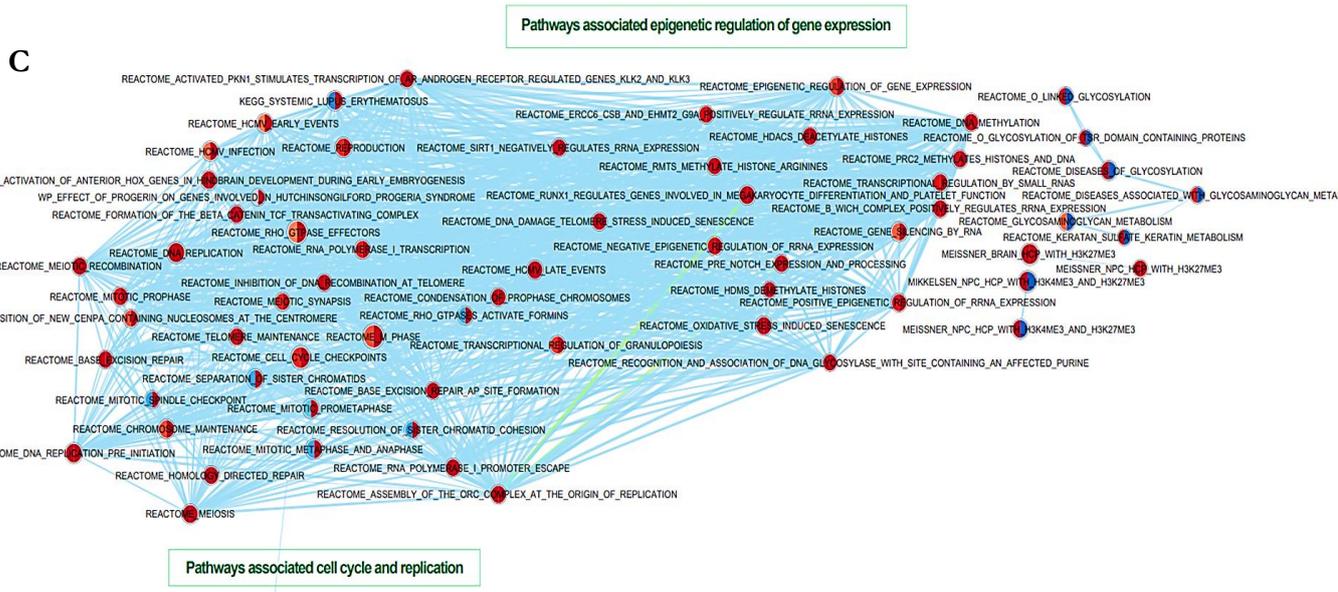
Name of gene set	Size of gene set	NES	FDR	Rank at max
Top 20 C2: curated gene sets				
Rosty_cervical_cancer_proliferation_cluster	92	3.15	<0.001	2016
Andersen_cholangiocarcinoma_class2	84	2.92	<0.001	2182
Sotiriou_breast_cancer_grade_1_vs_3_up	92	2.83	<0.001	1994
Kong_E2F3_targets	55	2.78	<0.001	2016
Blanco_bronchial_epithelial_cells_influenza_infection_dn	101	2.75	<0.001	2156
Odonnell_TFRC_targets_dn	63	2.75	<0.001	1705
Fischer_G2_M_cell_cycle	111	2.74	<0.001	1955
Kobayashi_EGFR_signalling_24hr_dn	116	2.63	0.0003	2626
Croonquist_IL6_deprivation_dn	56	2.60	0.0003	2175
Ishida_E2F_targets	32	2.60	0.0003	1540
Zhan_multiple_myeloma_pr_up	23	2.57	0.0003	1540
Chiang_liver_cancer_subclass_proliferation_up	95	2.57	0.0003	1984
Kang_doxorubicin_resistance_up	39	2.56	0.0003	1634
Whitfield_cell_cycle_literature	29	2.56	0.0003	2175
Odonnell_targets_of_MYC_and_TFRC1_dn	25	2.53	0.0004	1573
Whiteford_pediatric_cancer_markers	66	2.51	0.0005	2608
Vantveer_breast_cancer_metastasis_dn	59	2.50	0.0005	1914
Florio_neocortex_basal_radial_glia_dn	109	2.49	0.0006	1960
Croonquist_NRAS_signalling_dn	43	2.48	0.0006	1540
Villanueva_liver_cancer_KRT19_up	70	2.46	0.0008	1745

NES-Normalized enrichment score,

FDR-False discovery rate,
 Rank at max-Position of geneset at maximum enrichment score



To find out MMP associated pathways altered in GBC using publicly available databases



E

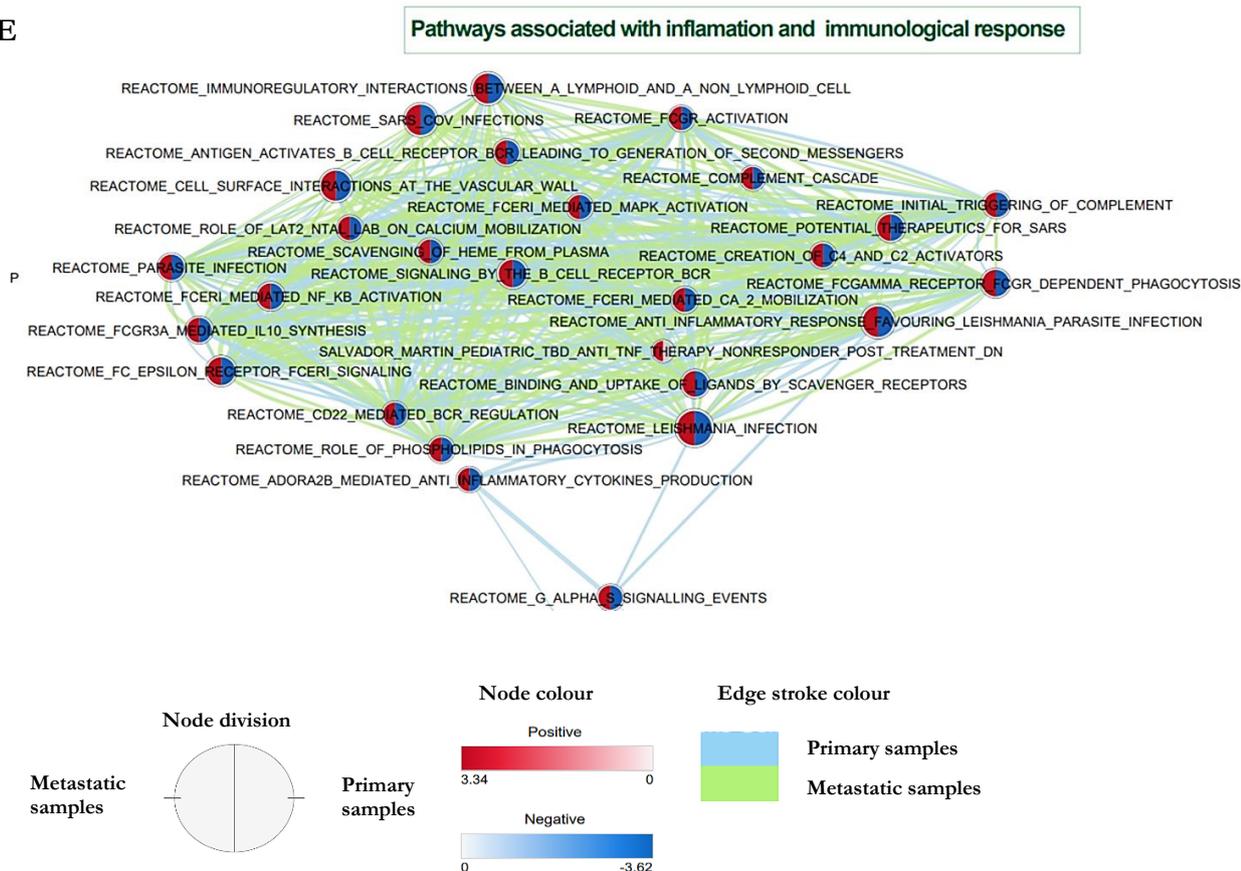


Figure 7.5.1 Visualisations of gene set enrichment analysis (GSEA) using Cytoscape Enrichment Map. Results of primary and metastatic GBC samples gene set enrichment analysis (GSEA) using Cytoscape Enrichment Map. Each node (circle) corresponds to a gene set either up-regulated (red) or down-regulated (blue) in each sample type, left hemisphere corresponds to metastatic samples and right hemisphere corresponds to primary samples. Edges (blue lines for primary samples and green lines for metastatic samples) link sets with shared genes, and thickness of lines correlates with the number of genes in common between two sets. Only gene sets with $FDR < 0.05$ and $p < 0.01$ were included in visualizations; disconnected nodes and small clusters were removed.

Table 7.5.2 Gene set enrichment analysis showing top gene-sets in GBC with metastatic samples, ranked according to NES

Name of gene set	Size of gene set	NES	FDR	Rank at max
Top 20 C2: curated gene sets				
Reactome_CD22_mediated_BCR_regulation	51	3.34	<0.001	1184
Reactome_antigen_activates_B_cell_receptor	58	3.31	<0.001	964
Reactome_FcεRI mediated MAPK activation	56	3.22	<0.001	1184
Reactome role of LAT2/NTAL/LAB on calcium mobilization	53	3.22	<0.001	1184

To find out MMP associated pathways altered in GBC using publicly available databases

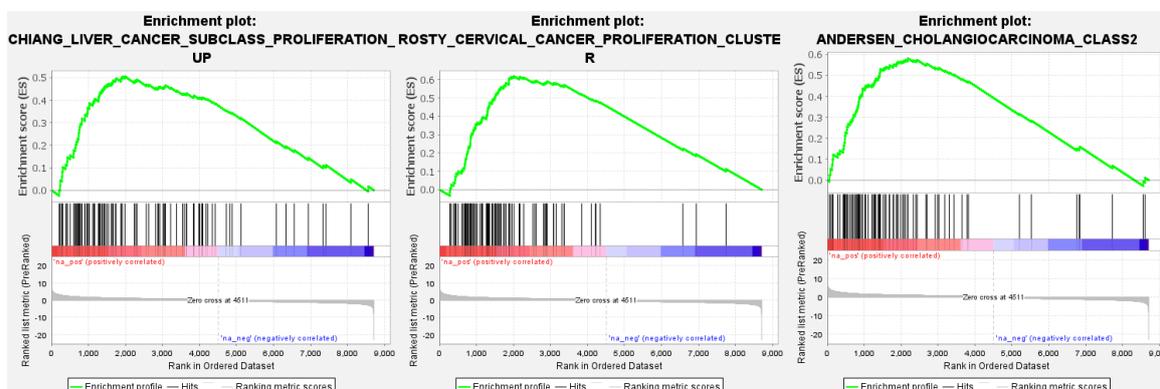
Reactome_parasite_infection	68	3.19	<0.001	964
Reactome_role_of_phospholipids_in_phagocytosis	58	3.16	<0.001	964
Reactome_Fc-gamma_receptor(FcγR)_activation	54	3.14	<0.001	1184
Reactome_FcεRI_mediated_CA ²⁺ _mobilization	56	3.12	<0.001	1184
Reactome_Fc-gamma_receptor(FcγR)_dependent_phagocytosis	74	3.09	<0.001	964
Reactome_potential_therapeutics_for_SARS	76	3.07	<0.001	1770
Reactome_FcγR3a_mediated_IL10_synthesis	61	3.07	<0.001	964
Reactome_scavenging_of_HEME_from_plasma	55	2.96	<0.001	1356
Reactome_signalling_by_the_B_cell_receptor(BCR)	77	2.94	<0.001	964
Reactome_anti_inflammatory_response_favouring_leishmaniasis	98	2.88	<0.001	964
Reactome_leishmania_infection	125	2.87	<0.001	964
Reactome_FcεRI_mediated_nf_kb_activation	68	2.87	<0.001	964
Reactome_Fc_epsilon_receptor_I(FcεRI)_signalling	80	2.79	<0.001	964
Reactome_creation_of_C4_and_C2_activators	59	2.61	<0.001	964
Salvador_pediatic_TBD_anti_TNF_therapy_nonresponder_dn	15	2.61	<0.001	1356
Reactome_SARS_Cov_infections	143	2.53	<0.001	1221

NES-Normalized enrichment score,

FDR-False discovery rate,

Rank at max-Position of geneset at maximum enrichment score

A.



B.

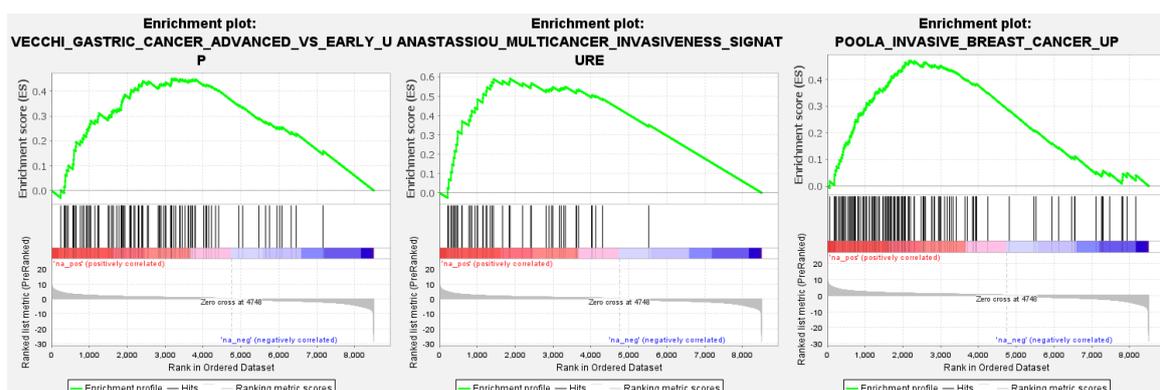


Figure 7.5.2 Gene set enrichment analysis (GSEA). *Results of primary and metastatic GBC samples gene set enrichment analysis (GSEA). (A) Top gene sets associated with primary GBC samples analysis include, growth and proliferation gene set clusters of gastric, cervical and cholangiocarcinoma. (B) similarly for metastatic samples include, gene sets pertaining to advance stage and invasive signature of gastric and breast cancer. The y-axis represents enrichment score (ES) and on the x-axis are genes (vertical black lines) represented in gene sets rank in ordered datasets. The green line connects points of ES and genes. ES is the maximum deviation from zero as calculated for each gene going down the ranked list, and represents the degree of over-representation of a gene set at the top or the bottom of the ranked gene list. The colored band at the bottom represents the degree of correlation of genes with the primary and metastatic samples datasets (red for positive and blue for negative correlation). Significance threshold set at $FDR < 0.05$.*

7.6 The role of matrix metalloproteinases (MMP-2,-7 and -14) in the pathogenesis of GBC

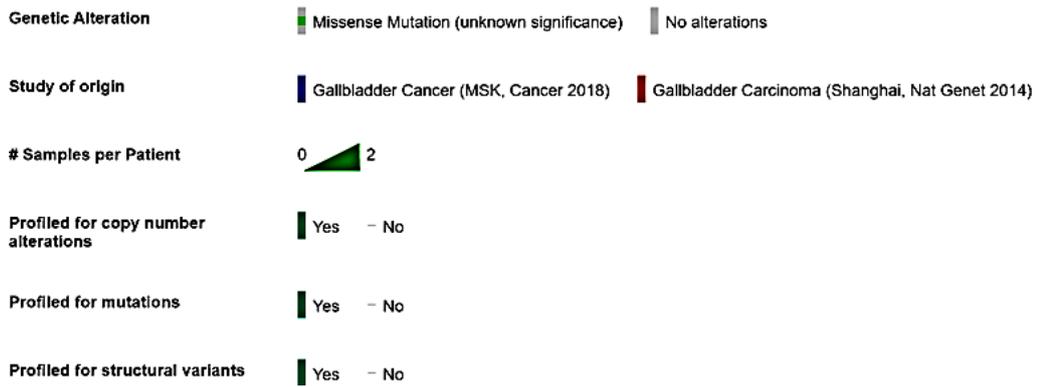
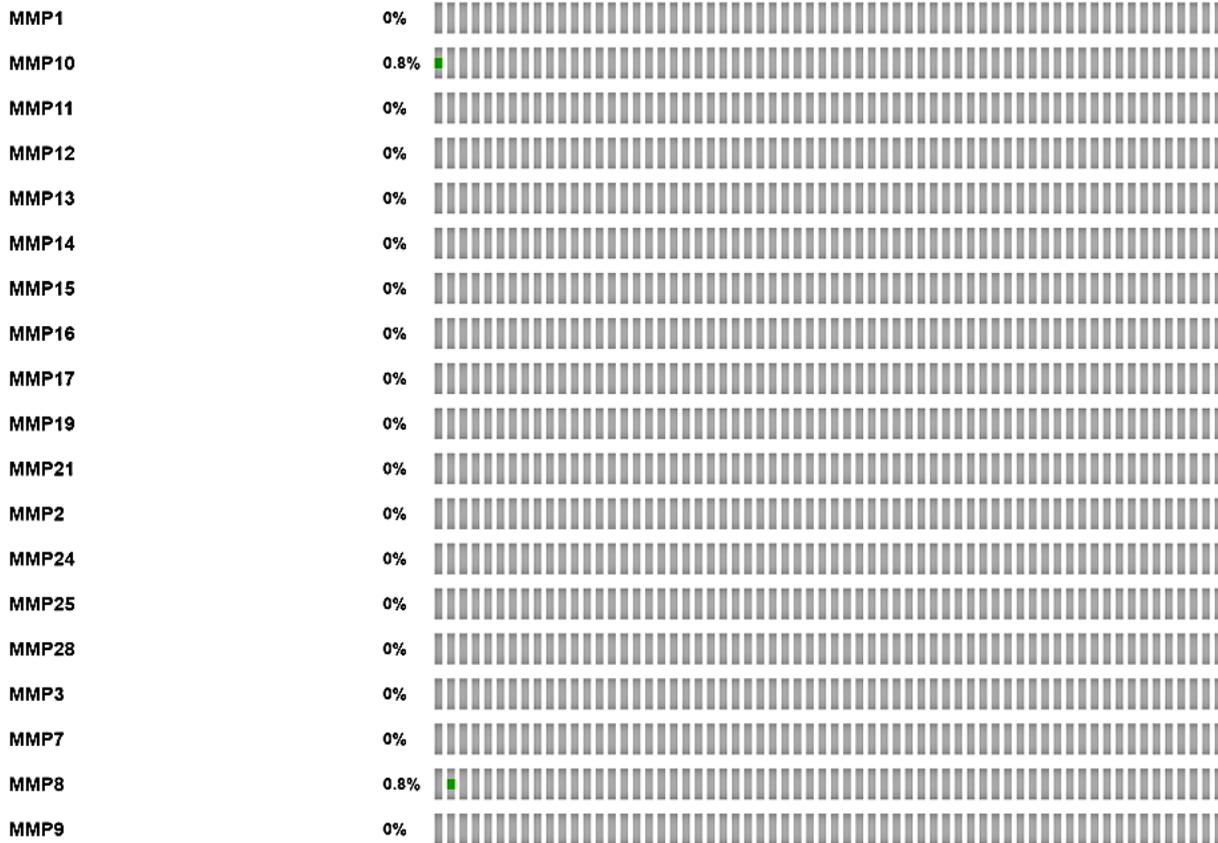
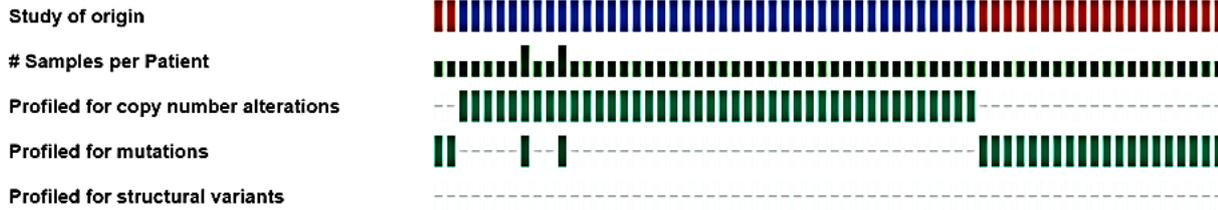
We aimed to explore the role of MMPs in GBC pathogenesis by analysing pathway alterations in primary and metastatic GBC patients. To achieve this, we conducted a pathway analysis of RNA-Seq data from publicly available databases of primary and metastatic GBC samples. Our analysis identified several dysregulated pathways and biological processes associated with MMP expression, including ECM remodelling, angiogenesis, cell adhesion, cytoskeleton organization, and immune response (**Figure 7.3.1** and **Figure 7.4.1**).

We further performed functional annotation and network analysis of primary and metastatic GBC samples, and found that MMPs were involved in pathways such as extracellular matrix organization, IL18 signalling, degradation of the extracellular matrix, malignant pleural mesothelioma, and TGF- β signalling (**Figure 7.5.1A** and **B**, red/blue coloured nodes are highlighted in yellow colour and red coloured edge lines). Notably, these pathways were positively enriched in both primary and metastatic samples (**Table 7.3.1**, **Table 7.4.1**), highlighting the importance of MMP-2, -7, and -14 in tumour initiation and metastasis of GBC. Overall, our findings shed light on the potential role of MMPs in GBC pathogenesis and may provide valuable insights for the development of targeted therapies for this deadly disease.

7.7 Somatic mutations, copy number variation and methylation profile of MMP-2, -7 and -14, in GBC patients

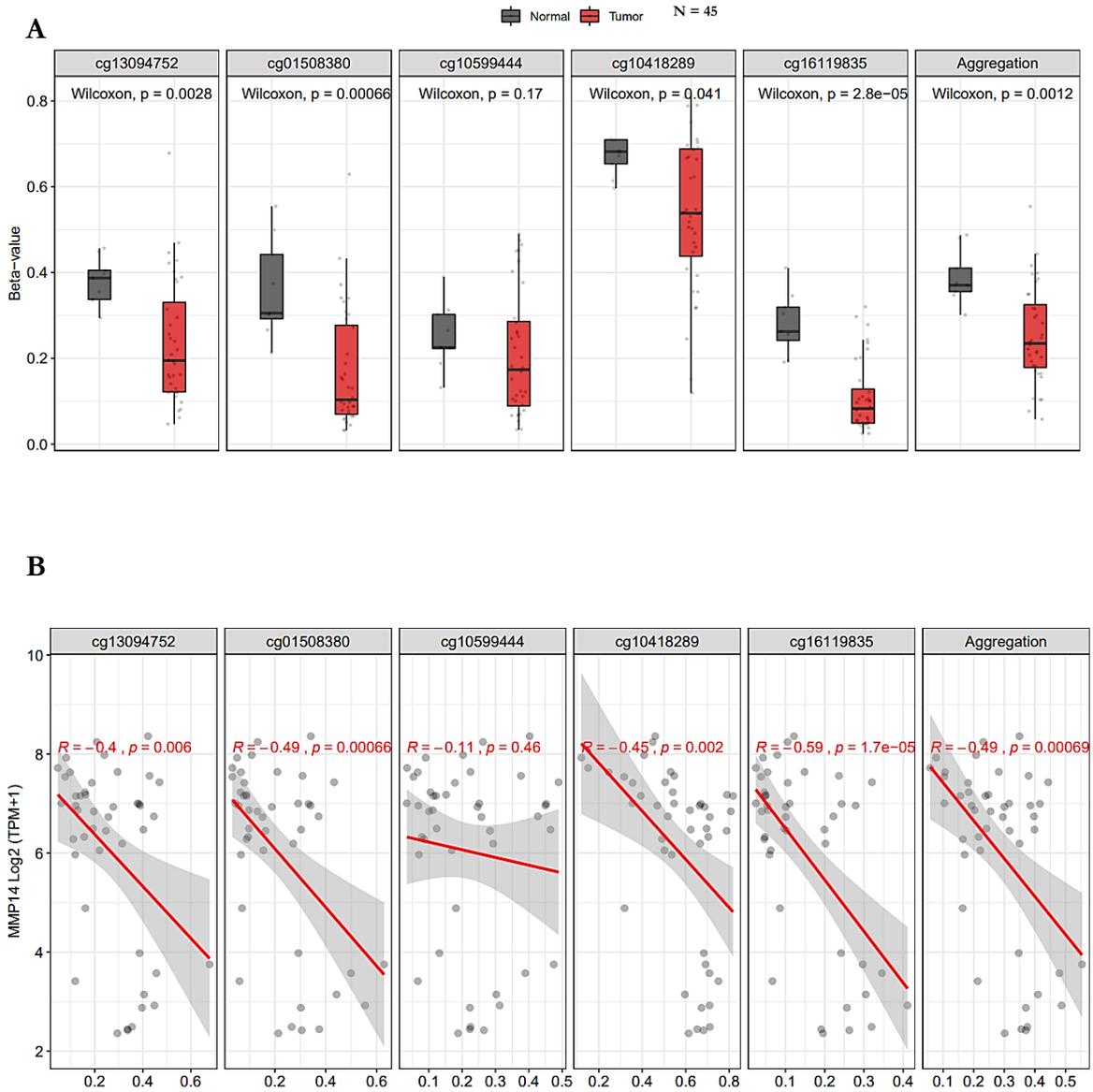
Further, we aimed to investigate the role of MMP-2, -7, and -14 in GBC pathogenesis by analysing somatic mutations, copy number variations, and methylation profile using TCGA data and SMART platform (accessed on 15th Sept 2022). Our analysis revealed that the frequency of somatic mutations and copy number variations of these MMPs in GBC patients was very low (**Figure 7.7.1**). However, we observed a significant change in the methylation levels of MMP7 and MMP14 in GBC patients, with a reduction in methylation level compared to normal samples ($p = 9.6E-05$ and $p = 0.0012$, respectively). On the other hand, MMP2 did not show any significant difference in methylation levels between normal and GBC patients ($p = 0.1900$). Furthermore, we found a negative correlation between MMP7 and MMP14 transcript levels with methylation status in GBC patients ($p = 0.0004$ and $p = 0.0007$, respectively), while no significant correlation was observed between MMP2 transcript level and methylation status ($p = 0.095$) (**Figure 7.7.2**). These findings emphasize the potential role of epigenetic modifications of MMP7 and MMP14 in GBC pathogenesis.

N = 133

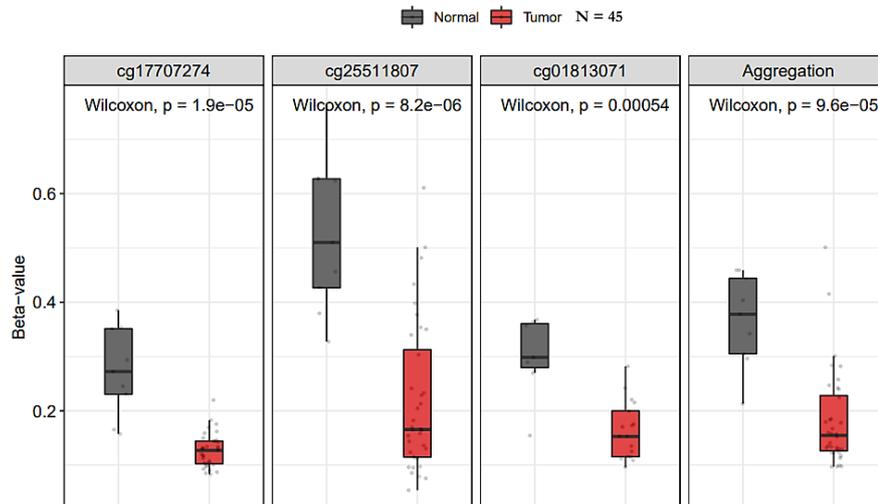


To find out MMP associated pathways altered in GBC using publicly available databases

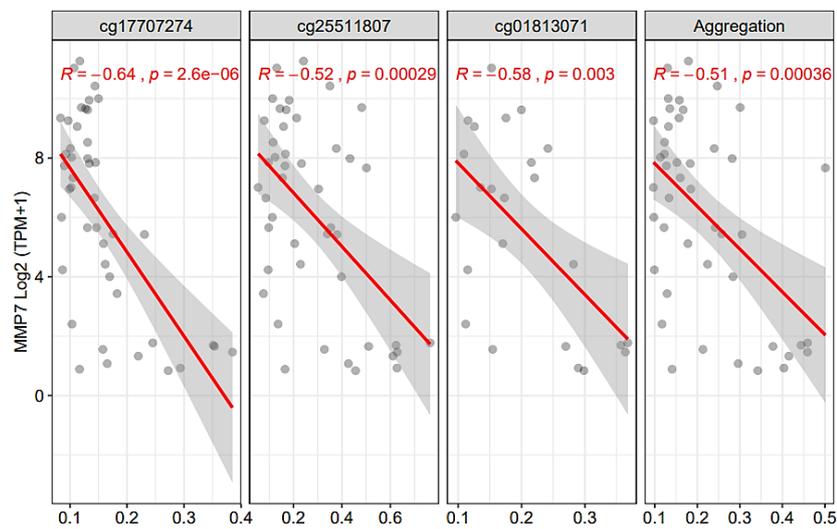
Figure 7.7.1 The OncoPrint plots showing frequency of Genetic alterations in various MMPs in gallbladder cancer. OncoPrint plots depicting the various types of mutations observed in MMP genes among gallbladder cancer patients. The plots were generated using cBioPortal by accessing the gallbladder cancer study in TCGA, which comprised of 133 patient samples.



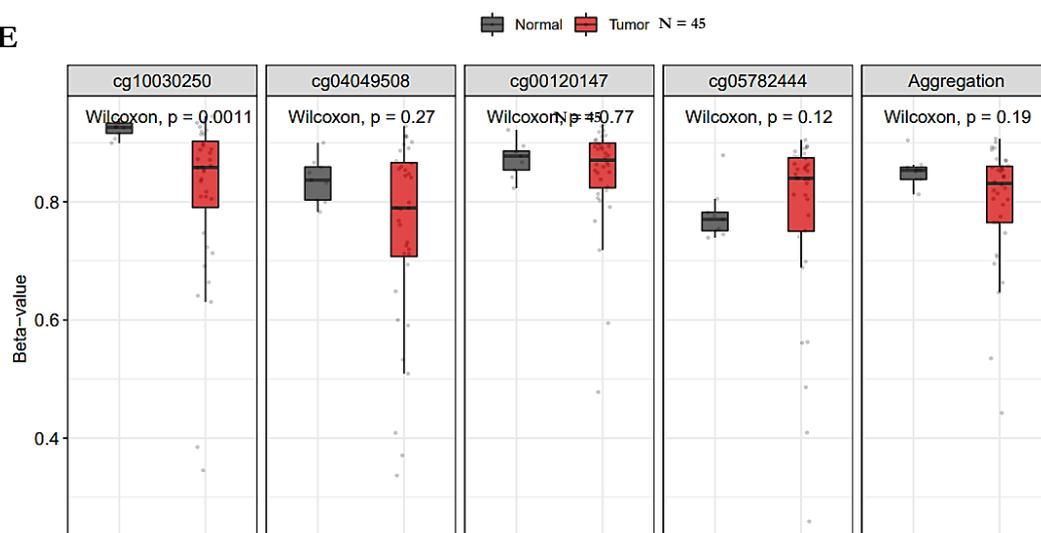
C



D



E



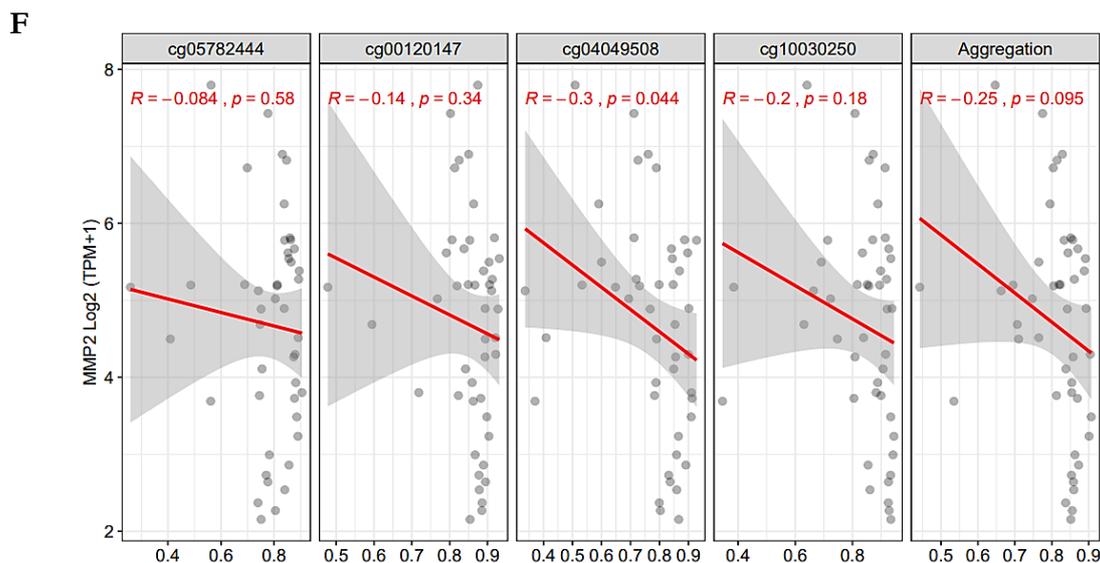


Figure 7.7.2 Methylation status of MMP-2, -7 and -14 promoter region and its correlation with its mRNA expression. The TCGA-CHOL datasets were utilized to investigate the correlation between methylation and mRNA expression of MMP-2, -7, and -14. Using the SMART tool or UCSC Xena browser, we conducted differential methylation level analysis of case vs. control and correlation analysis with the transcript levels of (A and B) MMP14, (C and D) MMP7, and (E and F) MMP2. The x-axis represents the methylation level of each probe in beta value, while the y-axis represents mRNA expression in $\log_2 (TPM+1)$.

7.8 Discussion

GBC is a lethal disease with poor prognosis and limited treatment options (383). The pathogenesis of GBC is influenced by a complex interplay between genetic and epigenetic alterations, as well as alterations in cellular signalling pathways (384). MMPs play crucial roles in extracellular matrix remodelling, cell migration, and invasion, and dysregulation of MMP expression and activity has been implicated in the pathogenesis of several cancers, including GBC (380, 385). MMP2, MMP7, and MMP14 have been shown to be overexpressed in GBC and may play a key role in the pathogenesis of this disease. However, the molecular pathways and biological processes that are regulated by these MMPs in GBC remain largely unknown.

In this study, we aimed to provide a comprehensive analysis of the pathways altered in GBC that are associated with MMPs. To achieve this, we performed pathway enrichment analysis

using publicly available databases, including the TCGA and GEO datasets (accession number: GSE139682 and GSE132223). We also used various bioinformatics tools such as DESeq2 (386), GSEA (387), DAVID Bioinformatics Resources (<http://david.abcc.ncifcrf.gov/>), Cytoscape and EnrichmentMap (388), UCSC Xena browser (<https://xena.ucsc.edu/>), and cBioportal (<https://www.cbioportal.org/>) to identify and analyse the molecular pathways and networks associated with MMPs. Additionally, we used SMART App (389) to investigate the epigenetic contribution of MMP2, MMP7, and MMP14 in GBC.

Our analysis revealed several key pathways and biological processes that are dysregulated in GBC and associated with MMP expression, including ECM remodelling, angiogenesis, cell adhesion, cytoskeleton organization, and immune response. Also, we identified several pathways that were dysregulated in both primary and metastatic GBC patients, such as Integrin binding, angiogenesis, cell adhesion, signal transduction, and extracellular matrix organization.

It is worth noting that the processes of integrin binding, cell adhesion, ECM remodelling, and cytoskeleton organization are interconnected and play a significant role in regulating cellular behaviour and affecting the tumour microenvironment (382). Integrins play a crucial role in connecting the ECM to the cytoskeleton via intracellular focal adhesions (FAs), which include kinases, scaffold, and adaptor proteins. The binding of integrins to FAs and ECM molecules facilitates cell adhesion to the ECM and transfers cytoskeletal forces onto the ECM, allowing for cell migration and signal transduction from the extracellular environment to the intracellular pathways (382, 390). This transmission is mediated by a cascade of signalling molecules, such as focal adhesion kinase (FAK), phosphatidylinositol 3-kinase (PI3K), extracellular signal-regulated kinase 1/2 (ERK1/2), mitogen-activated protein kinase (MAPK), small GTPases such as Rac and Rho, and the β -catenin pathway (391). Integrins are also involved in adhesion, migration, and invasion, which aid in the progression of the metastatic process by activating

MMPs that degrade the basement membrane (392). Our findings support these reports, indicating that ECM remodelling dysregulation is associated with MMP expression in GBC.

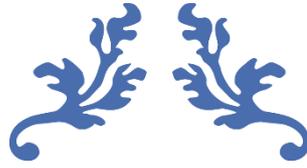
Moreover, ECM fragments resulting from cleavage of full-length ECM proteins have pro- and anti-angiogenic functions. These fragments include endostatin, tumstatin, canstatin, arresten, and hexastatin, which are derived from collagen IV and XVIII and can bind to cell receptors such as integrins and EGFR (393). For instance, arresten, which originates from the NC1 domain of the collagen $\alpha 1(\text{IV})$ chain, inhibits angiogenesis by binding to $\alpha 1\beta 1$ integrin and antagonizing MAPK signalling (394). The relevance of angiogenesis in gallbladder cancer is supported by the study's findings.

RNA sequencing of synovial sarcoma and high-dimensional spatial profiling of RNA and protein targets of primary and metastatic biopsies identified gene sets that were more highly enriched in metastasis and recurrence than in primary tumours, including Reactome Fc γ R Activation and Reactome CD22 Mediated BCR Regulation, suggesting the presence of B cells during the progressive stages (395). Our results also showed an enrichment of pathways related to the immune response in metastatic GBC samples (**Table 7.5.2**), which is consistent with these studies.

Our study has several implications for the understanding of the molecular mechanisms underlying GBC pathogenesis. First, our analysis highlights the utility of publicly available databases and bioinformatics tools in investigating complex biological systems. Second, our findings shed light on the potential epigenetic contribution of MMPs in GBC pathogenesis, providing a new perspective for future studies. Third, the identification of dysregulated pathways associated with MMP expression and activity may aid in the development of new therapeutic targets and biomarkers for GBC. Furthermore, our study may serve as a roadmap for other researchers investigating the role of MMPs in GBC.

Despite the significant insights provided by our study, some limitations should be acknowledged. First, our analysis was limited to publicly available datasets, and the sample sizes were relatively small. Future studies with larger sample sizes and comprehensive clinical data may provide more detailed insights into the role of MMPs in GBC pathogenesis. Second, our analysis was limited to bioinformatics tools, and further experimental validation of our findings is required.

In conclusion, our study provides a comprehensive analysis of the pathways altered in GBC that are associated with MMPs and emphasizes the potential epigenetic contribution of MMPs in GBC pathogenesis. Our findings may aid in the development of new therapeutic targets and biomarkers for GBC.



CHAPTER 8

Summary and conclusion



Vinay J
NISER, Bhubaneswar

Summary and conclusion

Decades of experimental and clinical research have established the crucial role of MMPs in tumour invasion and metastasis, as well as in physiological functions such as wound healing, inflammatory response, and angiogenesis (396, 397). Its deregulation in cancer is frequently linked to the breakdown of basement membrane and cell-matrix adhesion molecules, which promotes tumour invasion and metastasis (11). There have been instances of MMPs being upregulated and genetic polymorphisms being linked to a variety of malignancies, including gallbladder cancer. We learned from the literature that MMP-1, -2, -3, -7, and -14 play a part in tumour initiation and progression (41, 398).

The MMP14 promoter variants rs1004030 and rs1003349 were significantly associated with an increased risk for GBC. The earlier studies proposed regulatory binding sites for Sp1 and RR1, flanking SNPs rs1003349 and rs1004030, with allele-dependent regulation of transcription of MMP14 in rat renal mesangial cells (295). The current study identified direct allele-specific binding of MYB and SOX10 at rs1003349 and rs1004030 in GBC. This is further validated by various genetic and molecular assays such as electrophoretic mobility shift assays, luciferase reporter assays, and CRISPR-cas9-based deletion of a 119-bp genomic area surrounding the locus, in which loci rs1004030 and rs1003349 with alleles 'C' and 'G' were found to facilitate binding of transcription factors SOX10 and MYB, respectively. MYB (proto-oncogene or c-Myb) is overexpressed in leukemias, breast cancer, colon cancer, adenoid cystic carcinoma, and osteosarcoma. Additionally, the transcriptome gene enrichment functional analysis identifies MYB-associated additional genes that are involved in controlling the stress response, cell adhesion, and cell differentiation or morphogenesis (342, 343). SOX-10 overexpression was reported in triple-negative breast cancer (347), bladder cancer (348), and nasopharyngeal carcinoma as a differential diagnostic marker for metastasis and survival outcomes (349). SOX10 expression levels in melanoma are regulated in a β -catenin-dependent manner via the

canonical Wnt signalling pathway (355). β -catenin in non-cancer cells prevents membrane localization of MMP14 and thus its proteolytic activation of pro-MMP2. However, β -catenin increases the MMP14 expression and promotes Wnt-3a-mediated signalling via transcription factors Tcf-4/Lef in cancer cells (356). c-Myb interaction with β -catenin promotes invasion and metastasis of breast cancer by activating Wnt/ β -catenin/Axin2 signalling (357). MYB also mediated the activation of Wnt signalling and increased the MYC expression in colorectal cancers (CRC) (358), which emphasizes the complex regulation of MMP14 activity in cancer and non-cancer cells mediated by transcription factors MYB and SOX10 in cooperation with Wnt signalling. Overall, our study demonstrates the genetic association and mechanistic role of variants rs1003349 and rs1004030 in GBC pathogenesis.

MMP7 is a gene commonly associated with tumour invasion and is up-regulated in various types of cancer, including colon, liver, esophageal, and pancreatic cancers (44-47, 360-363). However, there have been limited studies on MMP7's role in GBC. Previous studies have reported the role of genetic variants within the MMP7 gene in various cancers (298-301). For example, the T181C polymorphism in gastric cancer is associated with poor overall survival of patients (364). In this study, variants rs11568819 G>A and rs11568818 T>C showed no association, but the variants rs113823671 A>C and rs17098318 G>A showed significant association with GBC. We also identified a significantly increased MMP7 expression in patients with GBC, which had not been previously reported. The reporter luciferase assay and genotype-phenotype correlation studies validated the potential role of the associated SNPs rs113823671 A>C and rs17098318 G>A in the expression of MMP7. The study found that the risk allele 'C' at rs113823671 significantly increased luciferase activity, suggesting that rs113823671 has a more significant influence, which requires more mechanistic study. The in-silico promoter analysis found that the SREBP-1c and USF-1 binding sites are created by the 'C' allele of rs113823671 and the 'A' allele of rs17098318, respectively. The study suggests that

promoter polymorphism may regulate MMP7 expression, which could therefore modulate GBC susceptibility. In conclusion, the study identified MMP7 promoter variants rs113823671 A>C and rs17098318 G>A as a risk factor for GBC and discovered that rs113823671 has an allele-specific functional impact on MMP7 expression levels. Further validation through detailed mechanistic studies is required.

Similarly, MMP2 is expressed in several cells, including vascular smooth muscle, cardiomyocytes, macrophages, and epithelial cells, and has a dual role in promoting tumour cell invasion and angiogenesis in many cancer (369). However, limited studies are available on its role in GBC. Studies have found that MMP2 variants are associated with an increased risk of various cancers and autoimmune and inflammatory disorders (372-374, 376, 399-402). For example, two regulatory genetic variants of MMP2, rs243865C>T and rs2285053C>T, have been linked to asthma (371), nasopharyngeal carcinoma (313), and lung cancer (312) in Chinese populations. These variants have allele-specific regulation of MMP2 expression by changing the binding site for the Sp1 transcription factor (314, 316). This study found that individuals with the 'T' allele at MMP2 promoter variants rs17859816, rs1488656253, rs1391392808, rs1961996235, rs243865, and rs1961998763 were more susceptible to GBC than those with the 'C' allele in the population of East Indian state, Odisha. Additionally, a novel single nucleotide polymorphism at loci NC_000016.10:g.55477735G>A, with the 'A' allele was significantly associated with GBC.

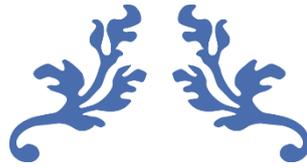
While genotype-phenotype correlation and luciferase assay were carried out to assess the functional relevance of associated MMP2 promoter variants. However, the study did not find any conclusive correlation with other SNPs except for variant rs243865, which showed an increase in MMP2 expression with each copy of the risk allele 'T'. The presence of the 'T' allele of rs243865 was found to increase MMP2. These findings indicate that the MMP2 gene and its variants may have significant implications in the pathogenesis and prognosis of various types

of cancer, including GBC. Further studies are needed to fully understand the molecular mechanisms of MMP2 regulation in GBC.

As, the molecular pathways and biological processes that are regulated by these MMPs in GBC remain largely unknown. We have used publicly available datasets and bioinformatics tools to address the same. Using RNA-Seq data from the GEO database, we performed a differential gene expression analysis of primary and metastatic GBC samples. Pathway and network analyses were conducted using DAVID and GSEA tools, revealing several key pathways and biological processes associated with MMP, such as angiogenesis, cell adhesion, signal transduction, and extracellular matrix organization. The study identified downstream effector pathways of MMPs that may play a critical role in GBC progression at the initial stage of cancer. The study also highlights the utility of publicly available databases and bioinformatics tools in investigating complex biological systems.

One of the key implications of our study includes, identification of transcriptional regulators of MMP14 can lead to the development of not only the biomarkers but also the specific therapeutic targets. So far, many studies have targeted specific MMPs by using function-blocking antibodies, cation zinc chelating hydroxylamine small molecules, and other small molecule inhibitors (403-405). However, focusing on regulation at the transcription level and targeting upstream signalling molecules may overcome these limitations. Defining the role of transcription factors that modulate metalloprotease expression is likely to provide new prospects for understanding its regulation and function and provide more insight into targeted therapies.

Overall, our study highlights the importance of identifying SNPs associated with an increased risk of GBC and the functional implications of these promoter variants in understanding the role of upstream regulatory molecules.



BIBLIOGRAPHY



Vinay J
NISER, Bhubaneswar

8.1 Bibliography

- 1 Sharma, A., Sharma, K.L., Gupta, A., Yadav, A. and Kumar, A. (2017) Gallbladder cancer epidemiology, pathogenesis and molecular genetics: Recent update. *World J. Gastroenterol.*, **23**, 3978-3998.
- 2 Hundal, R. and Shaffer, E.A. (2014) Gallbladder cancer: epidemiology and outcome. *Clin. Epidemiol.*, **6**, 99-109.
- 3 Sikdar, N. (2016) The Geographical, Ethnic Variations and Risk factors of Gallbladder Carcinoma: A Worldwide View. *Journal of Investigative Genomics*, **3**, 49-54.
- 4 Song, X., Hu, Y., Li, Y., Shao, R., Liu, F. and Liu, Y. (2020) Overview of current targeted therapy in gallbladder cancer. *Signal transduction and targeted therapy*, **5**, 230.
- 5 Li, Y., Song, Y., Zhang, Y. and Liu, S. (2022) Progress in gallbladder cancer with lymph node metastasis. *Front. Oncol.*, **12**, 966835.
- 6 Kleiner, D.E. and Stetler-Stevenson, W.G. (1999) Matrix metalloproteinases and metastasis. *Cancer Chemother. Pharmacol.*, **43 Suppl**, S42-51.
- 7 Dwivedi, A.N., Jain, S. and Dixit, R. (2015) Gall bladder carcinoma: Aggressive malignancy with protean loco-regional and distant spread. *World journal of clinical cases*, **3**, 231-244.
- 8 Yang, Y., Tu, Z., Ye, C., Cai, H., Yang, S., Chen, X. and Tu, J. (2021) Site-specific metastases of gallbladder adenocarcinoma and their prognostic value for survival: a SEER-based study. *BMC Surg.*, **21**, 59.
- 9 Ribatti, D., Tamma, R. and Annese, T. (2020) Epithelial-Mesenchymal Transition in Cancer: A Historical Overview. *Transl. Oncol.*, **13**, 100773.
- 10 Kessenbrock, K., Plaks, V. and Werb, Z. (2010) Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell*, **141**, 52-67.
- 11 Egeblad, M. and Werb, Z. (2002) New functions for the matrix metalloproteinases in cancer progression. *Nat. Rev. Cancer*, **2**, 161-174.
- 12 and, M.D.S. and Werb, Z. (2001) How Matrix Metalloproteinases Regulate Cell Behavior. **17**, 463-516.
- 13 Itoh, T., Tanioka, M., Yoshida, H., Yoshioka, T., Nishimoto, H. and Itohara, S. (1998) Reduced angiogenesis and tumor progression in gelatinase A-deficient mice. *Cancer Res.*, **58**, 1048-1051.
- 14 Vihinen, P. and Kahari, V.M. (2002) Matrix metalloproteinases in cancer: prognostic markers and therapeutic targets. *Int. J. Cancer*, **99**, 157-166.
- 15 Rotte, A., Martinka, M. and Li, G. (2012) MMP2 expression is a prognostic marker for primary melanoma patients. *Cell. Oncol. (Dordr.)*, **35**, 207-216.
- 16 Yu, C.-F., Chen, F.-H., Lu, M.-H., Hong, J.-H. and Chiang, C.-S. (2017) Dual roles of tumour cells-derived matrix metalloproteinase 2 on brain tumour growth and invasion. *Br. J. Cancer*, **117**, 1828-1836.
- 17 Hillebrand, L.E., Wickberg, S.M., Gomez-Auli, A., Follo, M., Maurer, J., Busch, H., Boerries, M. and Reinheckel, T. (2019) MMP14 empowers tumor-initiating breast cancer cells under hypoxic nutrient-depleted conditions. *FASEB J*, **33**, 4124-4140.
- 18 Li, H., Qiu, Z., Li, F. and Wang, C. (2017) The relationship between MMP-2 and MMP-9 expression levels with breast cancer incidence and prognosis. *Oncol. Lett.*, **14**, 5865-5870.
- 19 Stawowczyk, M., Wellenstein, M.D., Lee, S.B., Yomtoubian, S., Durrans, A., Choi, H., Narula, N., Altorki, N.K., Gao, D. and Mittal, V. (2017) Matrix Metalloproteinase 14 promotes lung cancer by cleavage of Heparin-Binding EGF-like Growth Factor. *Neoplasia*, **19**, 55-64.

- 20 Han, L., Sheng, B., Zeng, Q., Yao, W. and Jiang, Q. (2020) Correlation between MMP2 expression in lung cancer tissues and clinical parameters: a retrospective clinical analysis. *BMC Pulm. Med.*, **20**, 283.
- 21 Zeng, L., Qian, J., Zhu, F., Wu, F., Zhao, H. and Zhu, H. (2020) The prognostic values of matrix metalloproteinases in ovarian cancer. *J. Int. Med. Res.*, **48**, 300060519825983.
- 22 Buttacavoli, M., Di Cara, G., Roz, E., Pucci-Minafra, I., Feo, S. and Cancemi, P. (2021) Integrated Multi-Omics Investigations of Metalloproteinases in Colon Cancer: Focus on MMP2 and MMP9. *Int. J. Mol. Sci.*, **22**.
- 23 Ellenrieder, V., Alber, B., Lacher, U., Hendler, S.F., Menke, A., Boeck, W., Wagner, M., Wilda, M., Friess, H., Büchler, M. *et al.* (2000) Role of MT-MMPs and MMP-2 in pancreatic cancer progression. *Int. J. Cancer*, **85**, 14-20.
- 24 Liu, D., Kang, H., Gao, M., Jin, L., Zhang, F., Chen, D., Li, M. and Xiao, L. (2020) Exosome-transmitted circ_MMP2 promotes hepatocellular carcinoma metastasis by upregulating MMP2. *Mol. Oncol.*, **14**, 1365-1380.
- 25 Wang, H.L., Zhou, P.Y., Zhang, Y. and Liu, P. (2014) Relationships between abnormal MMP2 expression and prognosis in gastric cancer: a meta-analysis of cohort studies. *Cancer Biother. Radiopharm.*, **29**, 166-172.
- 26 Maybee, D.V., Ink, N.L. and Ali, M.A.M. (2022) Novel Roles of MT1-MMP and MMP-2: Beyond the Extracellular Milieu. *Int. J. Mol. Sci.*, **23**, 9513.
- 27 Ito, T.K., Ishii, G., Saito, S., Yano, K., Hoshino, A., Suzuki, T. and Ochiai, A. (2009) Degradation of soluble VEGF receptor-1 by MMP-7 allows VEGF access to endothelial cells. *Blood*, **113**, 2363-2369.
- 28 Rudolph-Owen, L.A., Chan, R., Muller, W.J. and Matrisian, L.M. (1998) The matrix metalloproteinase matrilysin influences early-stage mammary tumorigenesis. *Cancer Res.*, **58**, 5500-5506.
- 29 Hulboy, D.L., Gautam, S., Fingleton, B. and Matrisian, L.M. (2004) The influence of matrix metalloproteinase-7 on early mammary tumorigenesis in the multiple intestinal neoplasia mouse. *Oncol. Rep.*, **12**, 13-17.
- 30 Brabletz, T., Jung, A., Dag, S., Hlubek, F. and Kirchner, T. (1999) beta-catenin regulates the expression of the matrix metalloproteinase-7 in human colorectal cancer. *Am. J. Pathol.*, **155**, 1033-1038.
- 31 Hankey, W., Frankel, W.L. and Groden, J. (2018) Functions of the APC tumor suppressor protein dependent and independent of canonical WNT signaling: implications for therapeutic targeting. *Cancer Metastasis Rev.*, **37**, 159-172.
- 32 Fowlkes, J.L., Enghild, J.J., Suzuki, K. and Nagase, H. (1994) Matrix metalloproteinases degrade insulin-like growth factor-binding protein-3 in dermal fibroblast cultures. *J. Biol. Chem.*, **269**, 25742-25746.
- 33 Bergers, G., Brekken, R., McMahon, G., Vu, T.H., Itoh, T., Tamaki, K., Tanzawa, K., Thorpe, P., Itohara, S., Werb, Z. *et al.* (2000) Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nat. Cell Biol.*, **2**, 737-744.
- 34 Kirimlioğlu, H., Türkmen, I., Başsüllü, N., Dirican, A., Karadağ, N. and Kirimlioğlu, V. (2009) The expression of matrix metalloproteinases in intrahepatic cholangiocarcinoma, hilar (Klatskin tumor), middle and distal extrahepatic cholangiocarcinoma, gallbladder cancer, and ampullary carcinoma: role of matrix metalloproteinases in tumor progression and prognosis. *Turk. J. Gastroenterol.*, **20**, 41-47.
- 35 Ha, H.Y., Moon, H.B., Nam, M.S., Lee, J.W., Ryoo, Z.Y., Lee, T.H., Lee, K.K., So, B.J., Sato, H., Seiki, M. *et al.* (2001) Overexpression of membrane-type matrix metalloproteinase-1 gene

- induces mammary gland abnormalities and adenocarcinoma in transgenic mice. *Cancer Res.*, **61**, 984-990.
- 36 Karamanou, K., Franchi, M., Vynios, D. and Brézillon, S. (2020) Epithelial-to-mesenchymal transition and invadopodia markers in breast cancer: Lumican a key regulator. *Semin. Cancer Biol.*, **62**, 125-133.
- 37 Zheng, C.L., Lu, Q., Zhang, N., Jing, P.Y., Zhang, J.P., Wang, W.P. and Li, G.Z. (2021) Comprehensive Analysis of the Immune and Prognostic Implication of MMP14 in Lung Cancer. *Dis. Markers*, **2021**, 5917506.
- 38 Chen, T.Y., Li, Y.C., Liu, Y.F., Tsai, C.M., Hsieh, Y.H., Lin, C.W., Yang, S.F. and Weng, C.J. (2011) Role of MMP14 gene polymorphisms in susceptibility and pathological development to hepatocellular carcinoma. *Ann. Surg. Oncol.*, **18**, 2348-2356.
- 39 Wang, Y., Ye, Y., Lin, J., Meyer, L., Wu, X., Lu, K. and Liang, D. (2015) Genetic variants in matrix metalloproteinase genes as disposition factors for ovarian cancer risk, survival, and clinical outcome. *Mol. Carcinog.*, **54**, 430-439.
- 40 Duan, F., Peng, Z., Yin, J., Yang, Z. and Shang, J. (2020) Expression of MMP-14 and prognosis in digestive system carcinoma: a meta-analysis and databases validation. *J. Cancer*, **11**, 1141-1150.
- 41 Sabbota, A.L., Kim, H.R., Zhe, X., Fridman, R., Bonfil, R.D. and Cher, M.L. (2010) Shedding of RANKL by tumor-associated MT1-MMP activates Src-dependent prostate cancer cell migration. *Cancer Res.*, **70**, 5558-5566.
- 42 Makutani, Y., Kawakami, H., Tsujikawa, T., Yoshimura, K., Chiba, Y., Ito, A., Kawamura, J., Haratani, K. and Nakagawa, K. (2022) Contribution of MMP14-expressing cancer-associated fibroblasts in the tumor immune microenvironment to progression of colorectal cancer. *Front. Oncol.*, **12**, 956270.
- 43 Liao, H.-Y., Da, C.-M., Liao, B. and Zhang, H.-H. (2021) Roles of matrix metalloproteinase-7 (MMP-7) in cancer. *Clin. Biochem.*, **92**, 9-18.
- 44 Tan, X., Egami, H., Ishikawa, S., Sugita, H., Kamohara, H., Nakagawa, M., Nozawa, F., Abe, M. and Ogawa, M. (2005) Involvement of matrix metalloproteinase-7 in invasion-metastasis through induction of cell dissociation in pancreatic cancer. *Int. J. Oncol.*, **26**, 1283-1289.
- 45 Zhang, H., Wang, Y., Chen, T., Zhang, Y., Xu, R., Wang, W., Cheng, M. and Chen, Q. (2019) Aberrant Activation Of Hedgehog Signalling Promotes Cell Migration And Invasion Via Matrix Metalloproteinase-7 In Ovarian Cancer Cells. *J. Cancer*, **10**, 990-1003.
- 46 Yamamoto, H., Itoh, F., Adachi, Y., Sakamoto, H., Adachi, M., Hinoda, Y. and Imai, K. (1997) Relation of enhanced secretion of active matrix metalloproteinases with tumor spread in human hepatocellular carcinoma. *Gastroenterology*, **112**, 1290-1296.
- 47 Ii, M., Yamamoto, H., Adachi, Y., Maruyama, Y. and Shinomura, Y. (2006) Role of matrix metalloproteinase-7 (matrilysin) in human cancer invasion, apoptosis, growth, and angiogenesis. *Exp. Biol. Med. (Maywood)*, **231**, 20-27.
- 48 Littlepage, L.E., Sternlicht, M.D., Rougier, N., Phillips, J., Gallo, E., Yu, Y., Williams, K., Brenot, A., Gordon, J.I. and Werb, Z. (2010) Matrix metalloproteinases contribute distinct roles in neuroendocrine prostate carcinogenesis, metastasis, and angiogenesis progression. *Cancer Res.*, **70**, 2224-2234.
- 49 Dhir, V. and Mohandas, K.M. (1999) Epidemiology of digestive tract cancers in India IV. Gall bladder and pancreas. *Indian J. Gastroenterol.*, **18**, 24-28.
- 50 Hussain, M.A., Pati, S., Swain, S., Prusty, M., Kadam, S. and Nayak, S. (2012) Pattern and trends of cancer in odisha, India: a retrospective study. *Asian Pac. J. Cancer Prev.*, **13**, 6333-6336.

- 51 Karadag, N., Kirimlioglu, H., Isik, B., Yilmaz, S. and Kirimlioglu, V. (2008) Expression of matrix metalloproteinases in gallbladder carcinoma and their significance in carcinogenesis. *Appl. Immunohistochem. Mol. Morphol.*, **16**, 148-152.
- 52 Sharma, K.L., Misra, S., Kumar, A. and Mittal, B. (2012) Higher risk of matrix metalloproteinase (MMP-2, 7, 9) and tissue inhibitor of metalloproteinase (TIMP-2) genetic variants to gallbladder cancer. *Liver international : official journal of the International Association for the Study of the Liver*, **32**, 1278-1286.
- 53 Dutta, U., Bush, N., Kalsi, D., Popli, P. and Kapoor, V.K.J.C.C.O. (2019) Epidemiology of gallbladder cancer in India. *2019*, **8**.
- 54 Roa, I., Araya, J.C., Wistuba, I., De Aretxabala, X., Mauriz, A. and Cea, M. (1990) [In situ and invasive cancer of the gallbladder: nucleolar organizer regions]. *Rev. Med. Chil.*, **118**, 369-372.
- 55 Yamamoto, M., Nakajo, S. and Tahara, E. (1989) Dysplasia of the gallbladder. Its histogenesis and correlation to gallbladder adenocarcinoma. *Pathol. Res. Pract.*, **185**, 454-460.
- 56 Albores-Saavedra, J., Henson, D.E. and Klimstra, D.S. (2015) *Tumors of the Gallbladder, Extrahepatic Bile Ducts, and Vaterian System*. American Registry of Pathology.
- 57 Pérez-Moreno, P., Riquelme, I., García, P., Brebi, P. and Roa, J.C. (2022) Environmental and Lifestyle Risk Factors in the Carcinogenesis of Gallbladder Cancer. *Journal of personalized medicine*, **12**.
- 58 Roa, I., Araya, J.C., Villaseca, M., De Aretxabala, X., Riedemann, P., Endoh, K. and Roa, J. (1996) Preneoplastic lesions and gallbladder cancer: an estimate of the period required for progression. *Gastroenterology*, **111**, 232-236.
- 59 Sheth, S., Bedford, A. and Chopra, S. (2000) Primary gallbladder cancer: recognition of risk factors and the role of prophylactic cholecystectomy. *Am. J. Gastroenterol.*, **95**, 1402-1410.
- 60 Lazcano-Ponce, E.C., Miquel, J.F., Munoz, N., Herrero, R., Ferrecio, C., Wistuba, II, Alonso de Ruiz, P., Aristi Urista, G. and Nervi, F. (2001) Epidemiology and molecular pathology of gallbladder cancer. *CA Cancer J. Clin.*, **51**, 349-364.
- 61 Samuel, S., Mukherjee, S., Ammannagari, N., Pokuri, V.K., Kuvshinoff, B., Groman, A., LeVe, C.M. and Iyer, R. (2018) Clinicopathological characteristics and outcomes of rare histologic subtypes of gallbladder cancer over two decades: A population-based study. *PLoS One*, **13**, e0198809.
- 62 Bal, M.M., Ramadwar, M., Deodhar, K. and Shrikhande, S. (2015) Pathology of Gallbladder Carcinoma: Current Understanding and New Perspectives. *Pathology & Oncology Research*, **21**, 509-525.
- 63 Barreto, S.G., Dutt, A. and Chaudhary, A. (2014) A genetic model for gallbladder carcinogenesis and its dissemination. *Ann. Oncol.*, **25**, 1086-1097.
- 64 Wistuba, II and Albores-Saavedra, J. (1999) Genetic abnormalities involved in the pathogenesis of gallbladder carcinoma. *J. Hepatobiliary. Pancreat. Surg.*, **6**, 237-244.
- 65 Wistuba, II and Gazdar, A.F. (2004) Gallbladder cancer: lessons from a rare tumour. *Nat. Rev. Cancer*, **4**, 695-706.
- 66 Roa, J.C., Basturk, O. and Adsay, V. (2021) Dysplasia and carcinoma of the gallbladder: pathological evaluation, sampling, differential diagnosis and clinical implications. *Histopathology*, **79**, 2-19.
- 67 Hansel, D.E., Maitra, A. and Argani, P. (2004) Pathology of the gallbladder: a concise review. *Curr. Diagn. Pathol.*, **10**, 304-317.
- 68 Shrikhande, S.V., Barreto, S.G., Singh, S., Udwardia, T.E. and Agarwal, A.K. (2010) Cholelithiasis in gallbladder cancer: coincidence, cofactor, or cause! *Eur. J. Surg. Oncol.*, **36**, 514-519.

- 69 Kozuka, S., Tsubone, N., Yasui, A. and Hachisuka, K. (1982) Relation of adenoma to carcinoma in the gallbladder. *Cancer*, **50**, 2226-2234.
- 70 Harbison, J., Reynolds, J.V., Sheahan, K., Gibney, R.G. and Hyland, J.M. (1997) Evidence for the polyp-cancer sequence in gallbladder cancer. *Ir. Med. J.*, **90**, 98.
- 71 Wistuba, II, Miquel, J.F., Gazdar, A.F. and Albores-Saavedra, J. (1999) Gallbladder adenomas have molecular abnormalities different from those present in gallbladder carcinomas. *Hum. Pathol.*, **30**, 21-25.
- 72 Roa, I., de Aretxabala, X., Morgan, R., Molina, R., Araya, J.C., Roa, J. and Ibañeta, G. (2004) [Clinicopathological features of gallbladder polyps and adenomas]. *Rev. Med. Chil.*, **132**, 673-679.
- 73 Nomura, T., Shirai, Y., Sandoh, N., Nagakura, S. and Hatakeyama, K. (2002) Cholangiographic criteria for anomalous union of the pancreatic and biliary ducts. *Gastrointest. Endosc.*, **55**, 204-208.
- 74 Randi, G., Franceschi, S. and La Vecchia, C. (2006) Gallbladder cancer worldwide: geographical distribution and risk factors. *Int. J. Cancer*, **118**, 1591-1602.
- 75 Tanaka, K., Ikoma, A., Hamada, N., Nishida, S., Kadono, J. and Taira, A. (1998) Biliary Tract Cancer Accompanied by Anomalous Junction of Pancreaticobiliary Ductal System in Adults. *The American Journal of Surgery*, **175**, 218-220.
- 76 Adsay, V., Jang, K.T., Roa, J.C., Dursun, N., Ohike, N., Bagci, P., Basturk, O., Bandyopadhyay, S., Cheng, J.D., Sarmiento, J.M. *et al.* (2012) Intracholecystic papillary-tubular neoplasms (ICPN) of the gallbladder (neoplastic polyps, adenomas, and papillary neoplasms that are ≥ 1.0 cm): clinicopathologic and immunohistochemical analysis of 123 cases. *Am. J. Surg. Pathol.*, **36**, 1279-1301.
- 77 Yamamoto, M., Nakajo, S. and Tahara, E. (1988) Histological classification of epithelial polypoid lesions of the gallbladder. *Acta Pathol. Jpn.*, **38**, 181-192.
- 78 Sarcognato, S., Sacchi, D., Fassan, M., Fabris, L., Cadamuro, M., Zanusi, G., Cataldo, I., Covelli, C., Capelli, P., Furlanetto, A. *et al.* (2021) Benign biliary neoplasms and biliary tumor precursors. *Pathologica*, **113**, 147-157.
- 79 Albores-Saavedra, J., Tuck, M., McLaren, B.K., Carrick, K.S. and Henson, D.E. (2005) Papillary carcinomas of the gallbladder: analysis of noninvasive and invasive types. *Arch. Pathol. Lab. Med.*, **129**, 905-909.
- 80 Albores-Saavedra, J., Vardaman, C.J. and Vuitich, F. (1993) Non-neoplastic polypoid lesions and adenomas of the gallbladder. *Pathol. Annu.*, **28 Pt 1**, 145-177.
- 81 Kato, S. (1995) [Morphological analysis of the gallbladder elevated lesions--Macroscopic, stereoscopic, and histological study]. *Nihon Shokakibyō Gakkai Zasshi*, **92**, 1149-1160.
- 82 Laitio, M. (1983) Histogenesis of epithelial neoplasms of human gallbladder II. Classification of carcinoma on the basis of morphological features. *Pathol. Res. Pract.*, **178**, 57-66.
- 83 Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A. and Bray, F. (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.*, **71**, 209-249.
- 84 Stinton, L.M. and Shaffer, E.A. (2012) Epidemiology of gallbladder disease: cholelithiasis and cancer. *Gut Liver*, **6**, 172-187.
- 85 Nandakumar, A., Gupta, P.C., Gangadharan, P., Visweswara, R.N. and Parkin, D.M. (2005) Geographic pathology revisited: development of an atlas of cancer in India. *Int. J. Cancer*, **116**, 740-754.
- 86 Murthy NS, R.D., Gautham M, Shivraj N, Pruthvish S, George PS, Mathew. (2011) Trends in incidence of gallbladder cancer – Indian scenario. *Gastrointestinal Cancer: Targets and Therapy. Gastrointestinal Cancer: Targets and Therapy*, **1**, 1-9.

- 87 (2001) National Cancer Registry Programme (2001) Consolidated report of the Population Based Cancer Registries 1990–96. *Indian Council of Medical Research, New Delhi*, in press.
- 88 Das, A. (2016) Epidemiology of Gallbladder Cancer among North-Eastern State of India: A Review. *International Research Journal of Medical Sciences*, **4**, 11-15.
- 89 (2010) National Cancer Registry Programme (2010) Three year report of the Population Based Cancer Registries 2006–2008. *Indian Council of Medical Research, Bangalore India*, in press.
- 90 Huang, J., Patel, H.K., Boakye, D., Chandrasekar, V.T., Koulaouzidis, A., Lucero-Prisno Iii, D.E., Ngai, C.H., Pun, C.N., Bai, Y., Lok, V. *et al.* (2021) Worldwide distribution, associated factors, and trends of gallbladder cancer: A global country-level analysis. *Cancer Lett.*, **521**, 238-251.
- 91 Chen, C., Geng, Z., Shen, H., Song, H., Zhao, Y., Zhang, G., Li, W., Ma, L. and Wang, L. (2016) Long-Term Outcomes and Prognostic Factors in Advanced Gallbladder Cancer: Focus on the Advanced T Stage. *PLoS One*, **11**, e0166361.
- 92 (2020) Erratum: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*, **70**, 313.
- 93 Chhabra, D., Oda, K., Jagannath, P., Utsunomiya, H., Takekoshi, S. and Nimura, Y. (2012) Chronic heavy metal exposure and gallbladder cancer risk in India, a comparative study with Japan. *Asian Pac. J. Cancer Prev.*, **13**, 187-190.
- 94 Darby, S.C., Whitley, E., Howe, G.R., Hutchings, S.J., Kusiak, R.A., Lubin, J.H., Morrison, H.I., Tirmarche, M., Tomásek, L., Radford, E.P. *et al.* (1995) Radon and cancers other than lung cancer in underground miners: a collaborative analysis of 11 studies. *J. Natl. Cancer Inst.*, **87**, 378-384.
- 95 Malhotra, R.K., Manoharan, N., Shukla, N.K. and Rath, G.K. (2017) Gallbladder cancer incidence in Delhi urban: A 25-year trend analysis. *Indian J. Cancer*, **54**, 673-677.
- 96 Pilgrim, C.H., Groeschl, R.T., Christians, K.K. and Gamblin, T.C. (2013) Modern perspectives on factors predisposing to the development of gallbladder cancer. *HPB (Oxford)*, **15**, 839-844.
- 97 Lazcano-Ponce, E.C., Miquel, J.F., Muñoz, N., Herrero, R., Ferrecio, C., Wistuba, II, Alonso de Ruiz, P., Aristi Urista, G. and Nervi, F. (2001) Epidemiology and molecular pathology of gallbladder cancer. *CA Cancer J. Clin.*, **51**, 349-364.
- 98 Konstantinidis, I.T., Deshpande, V., Genevay, M., Berger, D., Fernandez-del Castillo, C., Tanabe, K.K., Zheng, H., Lauwers, G.Y. and Ferrone, C.R. (2009) Trends in Presentation and Survival for Gallbladder Cancer During a Period of More Than 4 Decades: A Single-Institution Experience. *Arch. Surg.*, **144**, 441-447.
- 99 Duffy, A., Capanu, M., Abou-Alfa, G.K., Huitzil, D., Jarnagin, W., Fong, Y., D'Angelica, M., Dematteo, R.P., Blumgart, L.H. and O'Reilly, E.M. (2008) Gallbladder cancer (GBC): 10-year experience at Memorial Sloan-Kettering Cancer Centre (MSKCC). *J. Surg. Oncol.*, **98**, 485-489.
- 100 J, V., Mishra, D., Meher, D., Dash, S., Besra, K., Pattnaik, N., Singh, S.P. and Dixit, M. (2021) Genetic association of MMP14 promoter variants and their functional significance in gallbladder cancer pathogenesis. *J. Hum. Genet.*, **66**, 947-956.
- 101 Rustagi, T. and Dasanu, C.A. (2012) Risk Factors for Gallbladder Cancer and Cholangiocarcinoma: Similarities, Differences and Updates. *J. Gastrointest. Cancer*, **43**, 137-147.
- 102 Chow, W.H., Johansen, C., Gridley, G., Mellemkjaer, L., Olsen, J.H. and Fraumeni, J.F., Jr. (1999) Gallstones, cholecystectomy and risk of cancers of the liver, biliary tract and pancreas. *Br. J. Cancer*, **79**, 640-644.
- 103 Paraskevopoulos, J.A., Dennison, A.R., Ross, B. and Johnson, A.G. (1992) Primary carcinoma of the gallbladder: a 10-year experience. *Ann. R. Coll. Surg. Engl.*, **74**, 222-224.
- 104 Zatonski, W.A., Lowenfels, A.B., Boyle, P., Maisonneuve, P., Bueno de Mesquita, H.B., Ghadirian, P., Jain, M., Przewozniak, K., Baghurst, P., Moerman, C.J. *et al.* (1997) Epidemiologic

- aspects of gallbladder cancer: a case-control study of the SEARCH Program of the International Agency for Research on Cancer. *J. Natl. Cancer Inst.*, **89**, 1132-1138.
- 105 Hsing, A.W., Gao, Y.T., Han, T.Q., Rashid, A., Sakoda, L.C., Wang, B.S., Shen, M.C., Zhang, B.H., Niwa, S., Chen, J. *et al.* (2007) Gallstones and the risk of biliary tract cancer: a population-based study in China. *Br. J. Cancer*, **97**, 1577-1582.
- 106 Serra, I., Yamamoto, M., Calvo, A., Cavada, G., Báez, S., Endoh, K., Watanabe, H. and Tajima, K. (2002) Association of chili pepper consumption, low socioeconomic status and longstanding gallstones with gallbladder cancer in a Chilean population. *Int. J. Cancer*, **102**, 407-411.
- 107 Diehl, A.K. (1983) Gallstone size and the risk of gallbladder cancer. *JAMA*, **250**, 2323-2326.
- 108 Rawla, P., Sunkara, T., Thandra, K.C. and Barsouk, A. (2019) Epidemiology of gallbladder cancer. *Clinical and experimental hepatology*, **5**, 93-102.
- 109 Muszynska, C., Lundgren, L., Lindell, G., Andersson, R., Nilsson, J., Sandström, P. and Andersson, B. (2017) Predictors of incidental gallbladder cancer in patients undergoing cholecystectomy for benign gallbladder disease: Results from a population-based gallstone surgery registry. *Surgery*, **162**, 256-263.
- 110 Koppatz, H., Nordin, A., Scheinin, T. and Sallinen, V. (2018) The risk of incidental gallbladder cancer is negligible in macroscopically normal cholecystectomy specimens. *HPB (Oxford)*, **20**, 456-461.
- 111 Engeland, A., Tretli, S., Austad, G. and Bjørge, T. (2005) Height and body mass index in relation to colorectal and gallbladder cancer in two million Norwegian men and women. *Cancer Causes Control*, **16**, 987-996.
- 112 Jain, K., Sreenivas, V., Velpandian, T., Kapil, U. and Garg, P.K. (2013) Risk factors for gallbladder cancer: a case-control study. *Int. J. Cancer*, **132**, 1660-1666.
- 113 Berk, R.N., Armbuster, T.G. and Saltzstein, S.L. (1973) Carcinoma in the porcelain gallbladder. *Radiology*, **106**, 29-31.
- 114 Stephen, A.E. and Berger, D.L. (2001) Carcinoma in the porcelain gallbladder: a relationship revisited. *Surgery*, **129**, 699-703.
- 115 Kim, J.H., Kim, W.H., Yoo, B.M., Kim, J.H. and Kim, M.W. (2009) Should we perform surgical management in all patients with suspected porcelain gallbladder? *Hepatogastroenterology*, **56**, 943-945.
- 116 Morimoto, M., Matsuo, T. and Mori, N. (2021) Management of Porcelain Gallbladder, Its Risk Factors, and Complications: A Review. *Diagnostics (Basel, Switzerland)*, **11**.
- 117 Bonatti, M., Vezzali, N., Lombardo, F., Ferro, F., Zamboni, G., Tauber, M. and Bonatti, G. (2017) Gallbladder adenomyomatosis: imaging findings, tricks and pitfalls. *Insights into imaging*, **8**, 243-253.
- 118 Maurer, K.J., Ihrig, M.M., Rogers, A.B., Ng, V., Bouchard, G., Leonard, M.R., Carey, M.C. and Fox, J.G. (2005) Identification of cholelithogenic enterohepatic helicobacter species and their role in murine cholesterol gallstone formation. *Gastroenterology*, **128**, 1023-1033.
- 119 Csendes, A., Burdiles, P., Maluenda, F., Diaz, J.C., Csendes, P. and Mitru, N. (1996) Simultaneous bacteriologic assessment of bile from gallbladder and common bile duct in control subjects and patients with gallstones and common duct stones. *Arch. Surg.*, **131**, 389-394.
- 120 Leung, J.W., Sung, J.Y. and Costerton, J.W. (1989) Bacteriological and electron microscopy examination of brown pigment stones. *J. Clin. Microbiol.*, **27**, 915-921.
- 121 Hazrah, P., Oahn, K.T., Tewari, M., Pandey, A.K., Kumar, K., Mohapatra, T.M. and Shukla, H.S. (2004) The frequency of live bacteria in gallstones. *HPB (Oxford)*, **6**, 28-32.
- 122 Yucebiligili, K., Mehmetoglu, T., Gucin, Z. and Salih, B.A. (2009) Helicobacter pylori DNA in gallbladder tissue of patients with cholelithiasis and cholecystitis. *J Infect Dev Ctries*, **3**, 856-859.

- 123 Vaishnavi, C., Singh, S., Kochhar, R., Bhasin, D., Singh, G. and Singh, K. (2005) Prevalence of *Salmonella enterica* serovar typhi in bile and stool of patients with biliary diseases and those requiring biliary drainage for other purposes. *Jpn. J. Infect. Dis.*, **58**, 363-365.
- 124 Kawai, M., Iwahashi, M., Uchiyama, K., Ochiai, M., Tanimura, H. and Yamaue, H. (2002) Gram-positive cocci are associated with the formation of completely pure cholesterol stones. *Am. J. Gastroenterol.*, **97**, 83-88.
- 125 Capoor, M.R., Nair, D., Rajni, Khanna, G., Krishna, S.V., Chintamani, M.S. and Aggarwal, P. (2008) Microflora of bile aspirates in patients with acute cholecystitis with or without cholelithiasis: a tropical experience. *Braz. J. Infect. Dis.*, **12**, 222-225.
- 126 Tewari, M., Mishra, R.R. and Shukla, H.S. (2010) *Salmonella typhi* and gallbladder cancer: report from an endemic region. *Hepatobiliary Pancreat. Dis. Int.*, **9**, 524-530.
- 127 Matsukura, N., Yokomuro, S., Yamada, S., Tajiri, T., Sundo, T., Hadama, T., Kamiya, S., Naito, Z. and Fox, J.G. (2002) Association between *Helicobacter bilis* in bile and biliary tract malignancies: *H. bilis* in bile from Japanese and Thai patients with benign and malignant diseases in the biliary tract. *Jpn. J. Cancer Res.*, **93**, 842-847.
- 128 Lewis, J.T., Talwalkar, J.A., Rosen, C.B., Smyrk, T.C. and Abraham, S.C. (2007) Prevalence and risk factors for gallbladder neoplasia in patients with primary sclerosing cholangitis: evidence for a metaplasia-dysplasia-carcinoma sequence. *Am. J. Surg. Pathol.*, **31**, 907-913.
- 129 Razumilava, N., Gores, G.J. and Lindor, K.D. (2011) Cancer surveillance in patients with primary sclerosing cholangitis. *Hepatology*, **54**, 1842-1852.
- 130 Chapman, R., Fevery, J., Kalloo, A., Nagorney, D.M., Boberg, K.M., Shneider, B. and Gores, G.J. (2010) Diagnosis and management of primary sclerosing cholangitis. *Hepatology*, **51**, 660-678.
- 131 Deshpande, V., Nduaguba, A., Zimmerman, S.M., Kehoe, S.M., MacConaill, L.E., Lauwers, G.Y., Ferrone, C., Bardeesy, N., Zhu, A.X. and Hezel, A.F. (2011) Mutational profiling reveals PIK3CA mutations in gallbladder carcinoma. *BMC Cancer*, **11**, 60.
- 132 Giraldo, N.A., Drill, E., Satravada, B.A., Dika, I.E., Brannon, A.R., Dermawan, J., Mohanty, A., Ozcan, K., Chakravarty, D., Benayed, R. *et al.* (2022) Comprehensive Molecular Characterization of Gallbladder Carcinoma and Potential Targets for Intervention. *Clin. Cancer. Res.*, **28**, 5359-5367.
- 133 Mishra, S., Kumari, S., Srivastava, P., Pandey, A., Shukla, S. and Husain, N. (2022) Genomic profiling of gallbladder carcinoma: Targetable mutations and pathways involved. *Pathol. Res. Pract.*, **232**, 153806.
- 134 Lin, J., Peng, X., Dong, K., Long, J., Guo, X., Li, H., Bai, Y., Yang, X., Wang, D., Lu, X. *et al.* (2021) Genomic characterization of co-existing neoplasia and carcinoma lesions reveals distinct evolutionary paths of gallbladder cancer. *Nature communications*, **12**, 4753.
- 135 Tada, M., Yokosuka, O., Omata, M., Ohto, M. and Isono, K. (1990) Analysis of ras gene mutations in biliary and pancreatic tumors by polymerase chain reaction and direct sequencing. *Cancer*, **66**, 930-935.
- 136 Wistuba, II, Sugio, K., Hung, J., Kishimoto, Y., Virmani, A.K., Roa, I., Albores-Saavedra, J. and Gazdar, A.F. (1995) Allele-specific mutations involved in the pathogenesis of endemic gallbladder carcinoma in Chile. *Cancer Res.*, **55**, 2511-2515.
- 137 Almoguera, C., Shibata, D., Forrester, K., Martin, J., Arnheim, N. and Perucho, M. (1988) Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. *Cell*, **53**, 549-554.
- 138 Hanada, K., Tsuchida, A., Iwao, T., Eguchi, N., Sasaki, T., Morinaka, K., Matsubara, K., Kawasaki, Y., Yamamoto, S. and Kajiyama, G. (1999) Gene mutations of K-ras in gallbladder mucosae and gallbladder carcinoma with an anomalous junction of the pancreaticobiliary duct. *Am. J. Gastroenterol.*, **94**, 1638-1642.

- 139 Masuhara, S., Kasuya, K., Aoki, T., Yoshimatsu, A., Tsuchida, A. and Koyanagi, Y. (2000) Relation between K-ras codon 12 mutation and p53 protein overexpression in gallbladder cancer and biliary ductal epithelia in patients with pancreaticobiliary maljunction. *J. Hepatobiliary. Pancreat. Surg.*, **7**, 198-205.
- 140 Matsubara, T., Sakurai, Y., Sasayama, Y., Hori, H., Ochiai, M., Funabiki, T., Matsumoto, K. and Hirono, I. (1996) K-ras point mutations in cancerous and noncancerous biliary epithelium in patients with pancreaticobiliary maljunction. *Cancer*, **77**, 1752-1757.
- 141 Iwase, T., Nakazawa, S., Yamao, K., Yoshino, J., Inui, K., Yamachika, H., Kanemaki, N., Fujimoto, M., Okushima, K., Miyoshi, H. *et al.* (1997) Ras gene point mutations in gallbladder lesions associated with anomalous connection of pancreatobiliary ducts. *Hepatogastroenterology*, **44**, 1457-1462.
- 142 Hanada, K., Itoh, M., Fujii, K., Tsuchida, A., Ooishi, H. and Kajiyama, G. (1996) K-ras and p53 mutations in stage I gallbladder carcinoma with an anomalous junction of the pancreaticobiliary duct. *Cancer*, **77**, 452-458.
- 143 Yokoyama, N., Hitomi, J., Watanabe, H., Ajioka, Y., Pruyas, M., Serra, I., Shirai, Y. and Hatakeyama, K. (1998) Mutations of p53 in gallbladder carcinomas in high-incidence areas of Japan and Chile. *Cancer Epidemiol. Biomarkers Prev.*, **7**, 297-301.
- 144 Bustos, B.I., Perez-Palma, E., Buch, S., Azocar, L., Riveras, E., Ugarte, G.D., Toliat, M., Nurnberg, P., Lieb, W., Franke, A. *et al.* (2019) Variants in ABCG8 and TRAF3 genes confer risk for gallstone disease in admixed Latinos with Mapuche Native American ancestry. *Sci. Rep.*, **9**, 772.
- 145 Mhatre, S., Wang, Z., Nagrani, R., Badwe, R., Chiplunkar, S., Mittal, B., Yadav, S., Zhang, H., Chung, C.C., Patil, P. *et al.* (2017) Common genetic variation and risk of gallbladder cancer in India: a case-control genome-wide association study. *Lancet Oncol.*, **18**, 535-544.
- 146 Cha, P.-C., Zembutsu, H., Takahashi, A., Kubo, M., Kamatani, N. and Nakamura, Y. (2012) A genome-wide association study identifies SNP in DCC is associated with gallbladder cancer in the Japanese population. *J. Hum. Genet.*, **57**, 235-237.
- 147 Cha, P.C., Zembutsu, H., Takahashi, A., Kubo, M., Kamatani, N. and Nakamura, Y. (2012) A genome-wide association study identifies SNP in DCC is associated with gallbladder cancer in the Japanese population. *J. Hum. Genet.*, **57**, 235-237.
- 148 Srivastava, K., Srivastava, A. and Mittal, B. (2010) Polymorphisms in ERCC2, MSH2, and OGG1 DNA repair genes and gallbladder cancer risk in a population of Northern India. *Cancer*, **116**, 3160-3169.
- 149 Castro, F.A., Koshiol, J., Hsing, A.W., Gao, Y.T., Rashid, A., Chu, L.W., Shen, M.C., Wang, B.S., Han, T.Q., Zhang, B.H. *et al.* (2012) Inflammatory gene variants and the risk of biliary tract cancers and stones: a population-based study in China. *BMC Cancer*, **12**, 468.
- 150 Hsing, A.W., Sakoda, L.C., Rashid, A., Andreotti, G., Chen, J., Wang, B.S., Shen, M.C., Chen, B.E., Rosenberg, P.S., Zhang, M. *et al.* (2008) Variants in inflammation genes and the risk of biliary tract cancers and stones: a population-based study in China. *Cancer Res.*, **68**, 6442-6452.
- 151 Vishnoi, M., Pandey, S.N., Choudhuri, G. and Mittal, B. (2008) IL-1 gene polymorphisms and genetic susceptibility of gallbladder cancer in a north Indian population. *Cancer Genet Cytogenet.*, **186**, 63-68.
- 152 Vishnoi, M., Pandey, S.N., Choudhury, G., Kumar, A., Modi, D.R. and Mittal, B. (2007) Do TNFA -308 G/A and IL6 -174 G/C gene polymorphisms modulate risk of gallbladder cancer in the north Indian population? *Asian Pac J Cancer Prev*, **8**, 567-572.
- 153 Vishnoi, M., Pandey, S.N., Modi, D.R., Kumar, A. and Mittal, B. (2008) Genetic susceptibility of epidermal growth factor +61A>G and transforming growth factor beta1 -509C>T gene polymorphisms with gallbladder cancer. *Hum Immunol*, **69**, 360-367.

- 154 Pandey, S.N., Srivastava, A., Dixit, M., Choudhuri, G. and Mittal, B. (2007) Haplotype analysis of signal peptide (insertion/deletion) and XbaI polymorphisms of the APOB gene in gallbladder cancer. *Liver international : official journal of the International Association for the Study of the Liver*, **27**, 1008-1015.
- 155 Pandey, S.N., Dixit, M., Choudhuri, G. and Mittal, B. (2006) Lipoprotein receptor associated protein (LRPAP1) insertion/deletion polymorphism: association with gallbladder cancer susceptibility. *Int. J. Gastrointest. Cancer*, **37**, 124-128.
- 156 Xu, H.L., Cheng, J.R., Andreotti, G., Gao, Y.T., Rashid, A., Wang, B.S., Shen, M.C., Chu, L.W., Yu, K. and Hsing, A.W. (2011) Cholesterol metabolism gene polymorphisms and the risk of biliary tract cancers and stones: a population-based case-control study in Shanghai, China. *Carcinogenesis*, **32**, 58-62.
- 157 Shukla, V.K., das, P.C., Dixit, R., Bhartiya, S.K., Basu, S. and Raman, M.J. (2012) Study of AP endonuclease (APEX1/REF1), a DNA repair enzyme, in gallbladder carcinoma. *Anticancer Res*, **32**, 1489-1492.
- 158 Mahananda, B., Vinay, J., Palo, A., Singh, A., Sahu, S.K., Singh, S.P. and Dixit, M. (2021) SERPINB5 Genetic Variants rs2289519 and rs2289521 are Significantly Associated with Gallbladder Cancer Risk. *DNA Cell Biol.*, **40**, 706-712.
- 159 Sinha, K.K., Vinay, J., Parida, S., Singh, S.P. and Dixit, M. (2022) Association and functional significance of genetic variants present in regulatory elements of SERPINB5 gene in gallbladder cancer. *Gene*, **808**, 145989.
- 160 Jiao, X., Ren, J., Chen, H., Ma, J., Rao, S., Huang, K., Wu, S., Fu, J., Su, X., Luo, C. *et al.* (2011) Ala499Val (C>T) and Lys939Gln (A>C) polymorphisms of the XPC gene: their correlation with the risk of primary gallbladder adenocarcinoma--a case-control study in China. *Carcinogenesis*, **32**, 496-501.
- 161 Tsuchiya, Y., Baez, S., Calvo, A., Pruyas, M., Nakamura, K., Kiyohara, C., Oyama, M., Ikegami, K. and Yamamoto, M. (2010) Evidence that genetic variants of metabolic detoxication and cell cycle control are not related to gallbladder cancer risk in Chilean women. *Int. J. Biol. Markers*, **25**, 75-78.
- 162 Kimura, A., Tsuchiya, Y., Lang, I., Zoltan, S., Nakadaira, H., Ajioka, Y., Kiyohara, C., Oyama, M. and Nakamura, K. (2008) Effect of genetic predisposition on the risk of gallbladder cancer in Hungary. *Asian Pac. J. Cancer Prev.*, **9**, 391-396.
- 163 Tsuchiya, Y., Kiyohara, C., Sato, T., Nakamura, K., Kimura, A. and Yamamoto, M. (2007) Polymorphisms of cytochrome P450 1A1, glutathione S-transferase class mu, and tumour protein p53 genes and the risk of developing gallbladder cancer in Japanese. *Clin. Biochem.*, **40**, 881-886.
- 164 Huang, W.Y., Gao, Y.T., Rashid, A., Sakoda, L.C., Deng, J., Shen, M.C., Wang, B.S., Han, T.Q., Zhang, B.H., Chen, B.E. *et al.* (2008) Selected base excision repair gene polymorphisms and susceptibility to biliary tract cancer and biliary stones: a population-based case-control study in China. *Carcinogenesis*, **29**, 100-105.
- 165 Srivastava, A., Pandey, S.N., Choudhuri, G. and Mittal, B. (2008) CCR5 Delta32 polymorphism: associated with gallbladder cancer susceptibility. *Scand. J. Immunol.*, **67**, 516-522.
- 166 Zhang, M., Huang, W.Y., Andreotti, G., Gao, Y.T., Rashid, A., Chen, J., Sakoda, L.C., Shen, M.C., Wang, B.S., Chanock, S. *et al.* (2008) Variants of DNA repair genes and the risk of biliary tract cancers and stones: a population-based study in China. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 2123-2127.
- 167 Jiao, X., Wu, Y., Zhou, L., He, J., Yang, C., Zhang, P., Hu, R., Luo, C., Du, J., Fu, J. *et al.* (2015) Variants and haplotypes in Flap endonuclease 1 and risk of gallbladder cancer and gallstones: a population-based study in China. *Sci. Rep.*, **5**, 18160.

- 168 Srivastava, A., Pandey, S.N., Dixit, M., Choudhuri, G. and Mittal, B. (2008) Cholecystokinin receptor A gene polymorphism in gallstone disease and gallbladder cancer. *J. Gastroenterol. Hepatol.*, **23**, 970-975.
- 169 Xu, H.L., Hsing, A.W., Vogtmann, E., Chu, L.W., Cheng, J.R., Gao, J., Tan, Y.T., Wang, B.S., Shen, M.C. and Gao, Y.T. (2013) Variants in CCK and CCKAR genes to susceptibility to biliary tract cancers and stones: a population-based study in Shanghai, China. *J. Gastroenterol. Hepatol.*, **28**, 1476-1481.
- 170 Srivastava, A., Sharma, K.L., Srivastava, N., Misra, S. and Mittal, B. (2012) Significant role of estrogen and progesterone receptor sequence variants in gallbladder cancer predisposition: a multi-analytical strategy. *PLoS One*, **7**, e40162.
- 171 Meyer, T.E., O'Brien, T.G., Andreotti, G., Yu, K., Li, Q., Gao, Y.T., Rashid, A., Shen, M.C., Wang, B.S., Han, T.Q. *et al.* (2010) Androgen receptor CAG repeat length and risk of biliary tract cancer and stones. *Cancer Epidemiol. Biomarkers Prev.*, **19**, 787-793.
- 172 Park, S.K., Andreotti, G., Sakoda, L.C., Gao, Y.T., Rashid, A., Chen, J., Chen, B.E., Rosenberg, P.S., Shen, M.C., Wang, B.S. *et al.* (2009) Variants in hormone-related genes and the risk of biliary tract cancers and stones: a population-based study in China. *Carcinogenesis*, **30**, 606-614.
- 173 Chang, S.C., Rashid, A., Gao, Y.T., Andreotti, G., Shen, M.C., Wang, B.S., Han, T.Q., Zhang, B.H., Sakoda, L.C., Leitzmann, M.F. *et al.* (2008) Polymorphism of genes related to insulin sensitivity and the risk of biliary tract cancer and biliary stone: a population-based case-control study in Shanghai, China. *Carcinogenesis*, **29**, 944-948.
- 174 Srivastava, A. and Mittal, B. (2009) Complement receptor 1 (A3650G RsaI and intron 27 HindIII) polymorphisms and risk of gallbladder cancer in north Indian population. *Scand. J. Immunol.*, **70**, 614-620.
- 175 Srivastava, K., Srivastava, A., Pandey, S.N., Kumar, A. and Mittal, B. (2009) Functional polymorphisms of the cyclooxygenase (PTGS2) gene and risk for gallbladder cancer in a North Indian population. *J. Gastroenterol.*, **44**, 774-780.
- 176 Sakoda, L.C., Gao, Y.T., Chen, B.E., Chen, J., Rosenberg, P.S., Rashid, A., Deng, J., Shen, M.C., Wang, B.S., Han, T.Q. *et al.* (2006) Prostaglandin-endoperoxide synthase 2 (PTGS2) gene polymorphisms and risk of biliary tract cancer and gallstones: a population-based study in Shanghai, China. *Carcinogenesis*, **27**, 1251-1256.
- 177 Srivastava, A., Pandey, S.N., Pandey, P., Choudhuri, G. and Mittal, B. (2008) No association of Methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism in susceptibility to gallbladder cancer. *DNA Cell Biol.*, **27**, 127-132.
- 178 Pandey, S.N., Modi, D.R., Choudhuri, G. and Mittal, B. (2007) Slow acetylator genotype of N-acetyl transferase2 (NAT2) is associated with increased susceptibility to gallbladder cancer: the cancer risk not modulated by gallstone disease. *Cancer Biol. Ther.*, **6**, 91-96.
- 179 Pandey, S.N., Jain, M., Nigam, P., Choudhuri, G. and Mittal, B. (2006) Genetic polymorphisms in GSTM1, GSTT1, GSTP1, GSTM3 and the susceptibility to gallbladder cancer in North India. *Biomarkers*, **11**, 250-261.
- 180 Hou, L., Xu, J., Gao, Y.T., Rashid, A., Zheng, S.L., Sakoda, L.C., Shen, M.C., Wang, B.S., Deng, J., Han, T.Q. *et al.* (2006) CYP17 MspA1 polymorphism and risk of biliary tract cancers and gallstones: a population-based study in Shanghai, China. *Int. J. Cancer*, **118**, 2847-2853.
- 181 Rai, R., Sharma, K.L., Misra, S., Kumar, A. and Mittal, B. (2014) CYP17 polymorphism (rs743572) is associated with increased risk of gallbladder cancer in tobacco users. *Tumour Biol.*, **35**, 6531-6537.
- 182 Pandey, S.N., Choudhuri, G. and Mittal, B. (2008) Association of CYP1A1 MspI polymorphism with tobacco-related risk of gallbladder cancer in a north Indian population. *Eur. J. Cancer Prev.*, **17**, 77-81.

- 183 Andreotti, G., Chen, J., Gao, Y.T., Rashid, A., Chen, B.E., Rosenberg, P., Sakoda, L.C., Deng, J., Shen, M.C., Wang, B.S. *et al.* (2008) Polymorphisms of genes in the lipid metabolism pathway and risk of biliary tract cancers and stones: a population-based case-control study in Shanghai, China. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 525-534.
- 184 Báez, S., Tsuchiya, Y., Calvo, A., Pruyas, M., Nakamura, K., Kiyohara, C., Oyama, M. and Yamamoto, M. (2010) Genetic variants involved in gallstone formation and capsaicin metabolism, and the risk of gallbladder cancer in Chilean women. *World J. Gastroenterol.*, **16**, 372-378.
- 185 Srivastava, A., Choudhuri, G. and Mittal, B. (2010) CYP7A1 (-204 A>C; rs3808607 and -469 T>C; rs3824260) promoter polymorphisms and risk of gallbladder cancer in North Indian population. *Metabolism*, **59**, 767-773.
- 186 Isomura, Y., Yamaji, Y., Ohta, M., Seto, M., Asaoka, Y., Tanaka, Y., Sasaki, T., Nakai, Y., Sasahira, N., Isayama, H. *et al.* (2010) A genetic polymorphism of CYP2C19 is associated with susceptibility to biliary tract cancer. *J. Gastroenterol.*, **45**, 1045-1052.
- 187 Rai, R., Kim, J.J., Misra, S., Kumar, A. and Mittal, B. (2015) A Multiple Interaction Analysis Reveals ADRB3 as a Potential Candidate for Gallbladder Cancer Predisposition via a Complex Interaction with Other Candidate Gene Variations. *Int. J. Mol. Sci.*, **16**, 28038-28049.
- 188 Srivastava, K., Srivastava, A. and Mittal, B. (2010) Caspase-8 polymorphisms and risk of gallbladder cancer in a northern Indian population. *Mol. Carcinog.*, **49**, 684-692.
- 189 Sharma, K.L., Misra, S., Kumar, A. and Mittal, B. (2013) Association of liver X receptors (LXRs) genetic variants to gallbladder cancer susceptibility. *Tumour Biol.*, **34**, 3959-3966.
- 190 Sharma, K.L., Yadav, A., Gupta, A., Tulsayan, S., Kumar, V., Misra, S., Kumar, A. and Mittal, B. (2014) Association of genetic variants of cancer stem cell gene CD44 haplotypes with gallbladder cancer susceptibility in North Indian population. *Tumour Biol.*, **35**, 2583-2589.
- 191 Yadav, A., Gupta, A., Rastogi, N., Agrawal, S., Kumar, A., Kumar, V. and Mittal, B. (2016) Association of cancer stem cell markers genetic variants with gallbladder cancer susceptibility, prognosis, and survival. *Tumour Biol.*, **37**, 1835-1844.
- 192 Ono, H., Chihara, D., Chiwaki, F., Yanagihara, K., Sasaki, H., Sakamoto, H., Tanaka, H., Yoshida, T., Saeki, N. and Matsuo, K. (2013) Missense allele of a single nucleotide polymorphism rs2294008 attenuated antitumor effects of prostate stem cell antigen in gallbladder cancer cells. *J. Carcinog.*, **12**, 4.
- 193 Rai, R., Sharma, K.L., Misra, S., Kumar, A. and Mittal, B. (2013) PSCA gene variants (rs2294008 and rs2978974) confer increased susceptibility of gallbladder carcinoma in females. *Gene*, **530**, 172-177.
- 194 Srivastava, K., Srivastava, A. and Mittal, B. (2010) Common genetic variants in pre-microRNAs and risk of gallbladder cancer in North Indian population. *J. Hum. Genet.*, **55**, 495-499.
- 195 Gupta, A., Sharma, A., Yadav, A., Rastogi, N., Agrawal, S., Kumar, A., Kumar, V., Misra, S. and Mittal, B. (2015) Evaluation of miR-27a, miR-181a, and miR-570 genetic variants with gallbladder cancer susceptibility and treatment outcome in a North Indian population. *Mol. Diagn. Ther.*, **19**, 317-327.
- 196 Rai, R., Sharma, K.L., Tiwari, S., Misra, S., Kumar, A. and Mittal, B. (2013) DCC (deleted in colorectal carcinoma) gene variants confer increased susceptibility to gallbladder cancer (Ref. No.: Gene-D-12-01446). *Gene*, **518**, 303-309.
- 197 Yadav, A., Gupta, A., Yadav, S., Rastogi, N., Agrawal, S., Kumar, A., Kumar, V., Misra, S. and Mittal, B. (2016) Association of Wnt signaling pathway genetic variants in gallbladder cancer susceptibility and survival. *Tumour Biol.*, **37**, 8083-8095.
- 198 Pramanik, V., Sarkar, B.N., Kar, M., Das, G., Malay, B.K., Sufia, K.K., Lakkakula, B.V. and Vadlamudi, R.R. (2011) A novel polymorphism in codon 25 of the KRAS gene associated with

- gallbladder carcinoma patients of the eastern part of India. *Genet. Test. Mol. Biomarkers*, **15**, 431-434.
- 199 Srivastava, K., Srivastava, A. and Mittal, B. (2010) Angiotensin I-converting enzyme insertion/deletion polymorphism and increased risk of gall bladder cancer in women. *DNA Cell Biol.*, **29**, 417-422.
- 200 Srivastava, K., Srivastava, A. and Mittal, B. (2010) DNMT3B -579 G>T promoter polymorphism and risk of gallbladder carcinoma in North Indian population. *J. Gastrointest. Cancer*, **41**, 248-253.
- 201 Srivastava, K., Srivastava, A., Kumar, A. and Mittal, B. (2010) Significant association between toll-like receptor gene polymorphisms and gallbladder cancer. *Liver international : official journal of the International Association for the Study of the Liver*, **30**, 1067-1072.
- 202 Rai, R., Sharma, K.L., Misra, S., Kumar, A. and Mittal, B. (2014) Association of adrenergic receptor gene polymorphisms in gallbladder cancer susceptibility in a North Indian population. *J. Cancer Res. Clin. Oncol.*, **140**, 725-735.
- 203 Sharma, K.L., Umar, M., Pandey, M., Misra, S., Kumar, A., Kumar, V. and Mittal, B. (2013) Association of potentially functional genetic variants of PLCE1 with gallbladder cancer susceptibility in north Indian population. *J. Gastrointest. Cancer*, **44**, 436-443.
- 204 Li, Z., Yuan, W.T., Ning, S.J. and Zhang, S.J. (2014) Vitamin D receptor genetic variants are associated with susceptibility of gallbladder adenocarcinoma in a Chinese cohort. *Genet. Mol. Res.*, **13**, 5387-5394.
- 205 Hay, N. and Sonenberg, N. (2004) Upstream and downstream of mTOR. *Genes Dev.*, **18**, 1926-1945.
- 206 Saxton, R.A. and Sabatini, D.M. (2017) mTOR Signaling in Growth, Metabolism, and Disease. *Cell*, **169**, 361-371.
- 207 Lien, E.C., Dibble, C.C. and Toker, A. (2017) PI3K signaling in cancer: beyond AKT. *Curr. Opin. Cell Biol.*, **45**, 62-71.
- 208 Hassan, B., Akcakanat, A., Holder, A.M. and Meric-Bernstam, F. (2013) Targeting the PI3-kinase/Akt/mTOR signaling pathway. *Surg Oncol Clin N Am*, **22**, 641-664.
- 209 Liu, P., Cheng, H., Roberts, T.M. and Zhao, J.J. (2009) Targeting the phosphoinositide 3-kinase pathway in cancer. *Nat Rev Drug Discov*, **8**, 627-644.
- 210 Jhanwar-Uniyal, M., Wainwright, J.V., Mohan, A.L., Tobias, M.E., Murali, R., Gandhi, C.D. and Schmidt, M.H. (2019) Diverse signaling mechanisms of mTOR complexes: mTORC1 and mTORC2 in forming a formidable relationship. *Adv Biol Regul*, **72**, 51-62.
- 211 Kim, J. and Guan, K.L. (2019) mTOR as a central hub of nutrient signalling and cell growth. *Nat Cell Biol*, **21**, 63-71.
- 212 Karimi Roshan, M., Soltani, A., Soleimani, A., Rezaie Kahkhaie, K., Afshari, A.R. and Soukhtanloo, M. (2019) Role of AKT and mTOR signaling pathways in the induction of epithelial-mesenchymal transition (EMT) process. *Biochimie*, **165**, 229-234.
- 213 Wu, C.E., Chen, M.H. and Yeh, C.N. (2019) mTOR Inhibitors in Advanced Biliary Tract Cancers. *Int J Mol Sci*, **20**.
- 214 Leal, P., Garcia, P., Sandoval, A., Buchegger, K., Weber, H., Tapia, O. and Roa, J.C. (2013) AKT/mTOR substrate P70S6K is frequently phosphorylated in gallbladder cancer tissue and cell lines. *Onco Targets Ther*, **6**, 1373-1384.
- 215 Leal, P., Garcia, P., Sandoval, A., Letelier, P., Brebi, P., Ili, C., Alvarez, H., Tapia, O. and Roa, J.C. (2013) Immunohistochemical expression of phospho-mTOR is associated with poor prognosis in patients with gallbladder adenocarcinoma. *Arch Pathol Lab Med*, **137**, 552-557.

- 216 Roa, I., Garcia, H., Game, A., de Toro, G., de Aretxabala, X. and Javle, M. (2016) Somatic Mutations of PI3K in Early and Advanced Gallbladder Cancer: Additional Options for an Orphan Cancer. *J Mol Diagn*, **18**, 388-394.
- 217 Deshpande, V., Nduaguba, A., Zimmerman, S.M., Kehoe, S.M., Macconail, L.E., Lauwers, G.Y., Ferrone, C., Bardeesy, N., Zhu, A.X. and Hezel, A.F. (2011) Mutational profiling reveals PIK3CA mutations in gallbladder carcinoma. *BMC Cancer*, **11**, 60.
- 218 Yang, D., Chen, T., Zhan, M., Xu, S., Yin, X., Liu, Q., Chen, W., Zhang, Y., Liu, D., Yan, J. *et al.* (2021) Modulation of mTOR and epigenetic pathways as therapeutics in gallbladder cancer. *Mol Ther Oncolytics*, **20**, 59-70.
- 219 Chong, H., Vikis, H.G. and Guan, K.L. (2003) Mechanisms of regulating the Raf kinase family. *Cell Signal*, **15**, 463-469.
- 220 Cargnello, M. and Roux, P.P. (2011) Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol Mol Biol Rev*, **75**, 50-83.
- 221 Buchegger, K., Silva, R., Lopez, J., Ili, C., Araya, J.C., Leal, P., Brebi, P., Riquelme, I. and Roa, J.C. (2017) The ERK/MAPK pathway is overexpressed and activated in gallbladder cancer. *Pathol Res Pract*, **213**, 476-482.
- 222 Li, Q. and Yang, Z. (2009) Expression of phospho-ERK1/2 and PI3-K in benign and malignant gallbladder lesions and its clinical and pathological correlations. *J Exp Clin Cancer Res*, **28**, 65.
- 223 Du, P., Liang, H., Fu, X., Wu, P., Wang, C., Chen, H., Zheng, B., Zhang, J., Hu, S., Zeng, R. *et al.* (2019) SLC25A22 promotes proliferation and metastasis by activating MAPK/ERK pathway in gallbladder cancer. *Cancer Cell Int*, **19**, 33.
- 224 Cao, Y., Liang, H., Zhang, F., Luan, Z., Zhao, S., Wang, X.A., Liu, S., Bao, R., Shu, Y., Ma, Q. *et al.* (2016) Prohibitin overexpression predicts poor prognosis and promotes cell proliferation and invasion through ERK pathway activation in gallbladder cancer. *J Exp Clin Cancer Res*, **35**, 68.
- 225 Hong, H., He, C., Zhu, S., Zhang, Y., Wang, X., She, F. and Chen, Y. (2016) CCR7 mediates the TNF-alpha-induced lymphatic metastasis of gallbladder cancer through the "ERK1/2 - AP-1" and "JNK - AP-1" pathways. *J Exp Clin Cancer Res*, **35**, 51.
- 226 Hong, H., Jiang, L., Lin, Y., He, C., Zhu, G., Du, Q., Wang, X., She, F. and Chen, Y. (2016) TNF-alpha promotes lymphangiogenesis and lymphatic metastasis of gallbladder cancer through the ERK1/2/AP-1/VEGF-D pathway. *BMC Cancer*, **16**, 240.
- 227 Li, M., Chen, L., Qu, Y., Sui, F., Yang, Q., Ji, M., Shi, B., Chen, M. and Hou, P. (2017) Identification of MAP kinase pathways as therapeutic targets in gallbladder carcinoma using targeted parallel sequencing. *Oncotarget*, **8**, 36319-36330.
- 228 Dhillon, A.S., Hagan, S., Rath, O. and Kolch, W. (2007) MAP kinase signalling pathways in cancer. *Oncogene*, **26**, 3279-3290.
- 229 Safa, A., Abak, A., Shoorei, H., Taheri, M. and Ghafouri-Fard, S. (2020) MicroRNAs as regulators of ERK/MAPK pathway: A comprehensive review. *Biomed Pharmacother*, **132**, 110853.
- 230 Wu, X.S., Wang, X.A., Wu, W.G., Hu, Y.P., Li, M.L., Ding, Q., Weng, H., Shu, Y.J., Liu, T.Y., Jiang, L. *et al.* (2014) MALAT1 promotes the proliferation and metastasis of gallbladder cancer cells by activating the ERK/MAPK pathway. *Cancer Biol Ther*, **15**, 806-814.
- 231 Sun, K.K., Hu, P.P. and Xu, F. (2018) Prognostic significance of long non-coding RNA MALAT1 for predicting the recurrence and metastasis of gallbladder cancer and its effect on cell proliferation, migration, invasion, and apoptosis. *J Cell Biochem*, **119**, 3099-3110.
- 232 Wang, S.H., Zhang, W.J., Wu, X.C., Weng, M.Z., Zhang, M.D., Cai, Q., Zhou, D., Wang, J.D. and Quan, Z.W. (2016) The lncRNA MALAT1 functions as a competing endogenous RNA to

- regulate MCL-1 expression by sponging miR-363-3p in gallbladder cancer. *J Cell Mol Med*, **20**, 2299-2308.
- 233 Lin, N., Yao, Z., Xu, M., Chen, J., Lu, Y., Yuan, L., Zhou, S., Zou, X. and Xu, R. (2019) Long noncoding RNA MALAT1 potentiates growth and inhibits senescence by antagonizing ABI3BP in gallbladder cancer cells. *J Exp Clin Cancer Res*, **38**, 244.
- 234 Bao, R.F., Shu, Y.J., Hu, Y.P., Wang, X.A., Zhang, F., Liang, H.B., Ye, Y.Y., Li, H.F., Xiang, S.S., Weng, H. *et al.* (2016) miR-101 targeting ZFX suppresses tumor proliferation and metastasis by regulating the MAPK/Erk and Smad pathways in gallbladder carcinoma. *Oncotarget*, **7**, 22339-22354.
- 235 Shu, Y.J., Bao, R.F., Jiang, L., Wang, Z., Wang, X.A., Zhang, F., Liang, H.B., Li, H.F., Ye, Y.Y., Xiang, S.S. *et al.* (2017) MicroRNA-29c-5p suppresses gallbladder carcinoma progression by directly targeting CPEB4 and inhibiting the MAPK pathway. *Cell Death Differ*, **24**, 445-457.
- 236 Du, Z. and Lovly, C.M. (2018) Mechanisms of receptor tyrosine kinase activation in cancer. *Mol Cancer*, **17**, 58.
- 237 Sigismund, S., Avanzato, D. and Lanzetti, L. (2018) Emerging functions of the EGFR in cancer. *Mol Oncol*, **12**, 3-20.
- 238 Tomas, A., Futter, C.E. and Eden, E.R. (2014) EGF receptor trafficking: consequences for signaling and cancer. *Trends Cell Biol*, **24**, 26-34.
- 239 Yarden, Y. (2001) The EGFR family and its ligands in human cancer. signalling mechanisms and therapeutic opportunities. *Eur J Cancer*, **37 Suppl 4**, S3-8.
- 240 Nicholson, R.I., Gee, J.M. and Harper, M.E. (2001) EGFR and cancer prognosis. *Eur J Cancer*, **37 Suppl 4**, S9-15.
- 241 Li, M., Zhang, Z., Li, X., Ye, J., Wu, X., Tan, Z., Liu, C., Shen, B., Wang, X.A., Wu, W. *et al.* (2014) Whole-exome and targeted gene sequencing of gallbladder carcinoma identifies recurrent mutations in the ErbB pathway. *Nat Genet*, **46**, 872-876.
- 242 Gomes, R.V. (2014) The ERBB pathway is recurrently mutated in gallbladder carcinoma. *Cancer Discov*, **4**, OF12.
- 243 Hadi, R., Pant, M.C., Husain, N., Singhal, A., Khurana, R., Agarwal, G.R., Masood, S. and Awashthi, N.P. (2016) EGFR and HER-2/neu Expression in Gallbladder Carcinoma: An Institutional Experience. *Gulf J Oncolog*, **1**, 12-19.
- 244 Doval, D.C., Azam, S., Sinha, R., Batra, U. and Mehta, A. (2014) Expression of epidermal growth factor receptor, p53, Bcl2, vascular endothelial growth factor, cyclooxygenase-2, cyclin D1, human epidermal receptor-2 and Ki-67: Association with clinicopathological profiles and outcomes in gallbladder carcinoma. *J Carcinog*, **13**, 10.
- 245 Shen, H., He, M., Lin, R., Zhan, M., Xu, S., Huang, X., Xu, C., Chen, W., Yao, Y., Mohan, M. *et al.* (2019) PLEK2 promotes gallbladder cancer invasion and metastasis through EGFR/CCL2 pathway. *J. Exp. Clin. Cancer Res.*, **38**, 247.
- 246 Huang, S., Armstrong, E.A., Benavente, S., Chinnaiyan, P. and Harari, P.M. (2004) Dual-agent molecular targeting of the epidermal growth factor receptor (EGFR): combining anti-EGFR antibody with tyrosine kinase inhibitor. *Cancer Res*, **64**, 5355-5362.
- 247 Matar, P., Rojo, F., Cassia, R., Moreno-Bueno, G., Di Cosimo, S., Tabernero, J., Guzman, M., Rodriguez, S., Arribas, J., Palacios, J. *et al.* (2004) Combined epidermal growth factor receptor targeting with the tyrosine kinase inhibitor gefitinib (ZD1839) and the monoclonal antibody cetuximab (IMC-C225): superiority over single-agent receptor targeting. *Clin Cancer Res*, **10**, 6487-6501.
- 248 Behari, A., Sikora, S.S., Waghlikar, G.D., Kumar, A., Saxena, R. and Kapoor, V.K. (2003) Longterm survival after extended resections in patients with gallbladder cancer. *J. Am. Coll. Surg.*, **196**, 82-88.

- 249 Agarwal, A.K., Kalayarasan, R., Javed, A., Gupta, N. and Nag, H.H. (2013) The role of staging laparoscopy in primary gall bladder cancer--an analysis of 409 patients: a prospective study to evaluate the role of staging laparoscopy in the management of gallbladder cancer. *Ann. Surg.*, **258**, 318-323.
- 250 Shukla, H.S., Sirohi, B., Behari, A., Sharma, A., Majumdar, J., Ganguly, M., Tewari, M., Kumar, S., Saini, S., Sahni, P. *et al.* (2015) Indian Council of Medical Research consensus document for the management of gall bladder cancer. *Indian J. Med. Paediatr. Oncol.*, **36**, 79-84.
- 251 Takada, T., Amano, H., Yasuda, H., Nimura, Y., Matsushiro, T., Kato, H., Nagakawa, T. and Nakayama, T. (2002) Is postoperative adjuvant chemotherapy useful for gallbladder carcinoma? A phase III multicenter prospective randomized controlled trial in patients with resected pancreaticobiliary carcinoma. *Cancer*, **95**, 1685-1695.
- 252 Valle, J., Wasan, H., Palmer, D.H., Cunningham, D., Anthoney, A., Maraveyas, A., Madhusudan, S., Iveson, T., Hughes, S., Pereira, S.P. *et al.* (2010) Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer. *N. Engl. J. Med.*, **362**, 1273-1281.
- 253 Lee, J., Hong, T.H., Lee, I.S., You, Y.K. and Lee, M.A. (2015) Comparison of the Efficacy between Gemcitabine-Cisplatin and Capecitabine-Cisplatin Combination Chemotherapy for Advanced Biliary Tract Cancer. *Cancer Res. Treat.*, **47**, 259-265.
- 254 Kanthan, R., Senger, J.L., Ahmed, S. and Kanthan, S.C. (2015) Gallbladder Cancer in the 21st Century. *J. Oncol.*, **2015**, 967472.
- 255 Sachs, T.E., Akintorin, O. and Tseng, J. (2018) How Should Gallbladder Cancer Be Managed? *Adv. Surg.*, **52**, 89-100.
- 256 Zhong, Y., Wu, X., Li, Q., Ge, X., Wang, F., Wu, P., Deng, X. and Miao, L. (2019) Long noncoding RNAs as potential biomarkers and therapeutic targets in gallbladder cancer: a systematic review and meta-analysis. *Cancer Cell Int.*, **19**, 169.
- 257 Chandra, V., Kim, J.J., Mittal, B. and Rai, R. (2016) MicroRNA aberrations: An emerging field for gallbladder cancer management. *World J. Gastroenterol.*, **22**, 1787-1799.
- 258 Margonis, G.A., Gani, F., Buettner, S., Amini, N., Sasaki, K., Andreatos, N., Ethun, C.G., Poultsides, G., Tran, T., Idrees, K. *et al.* (2016) Rates and patterns of recurrence after curative intent resection for gallbladder cancer: a multi-institution analysis from the US Extra-hepatic Biliary Malignancy Consortium. *HPB (Oxford)*, **18**, 872-878.
- 259 Kim, W.S., Choi, D.W., You, D.D., Ho, C.Y., Heo, J.S. and Choi, S.H. (2010) Risk factors influencing recurrence, patterns of recurrence, and the efficacy of adjuvant therapy after radical resection for gallbladder carcinoma. *J. Gastrointest. Surg.*, **14**, 679-687.
- 260 Dixit, M., J. V. and Choudhury, S. (2023) Kumar Shukla, V., Pandey, M. and Dixit, R. (eds.), In *Gallbladder Cancer: Current Treatment Options*. Springer Nature Singapore, Singapore, in press., pp. 291-316.
- 261 Gross, J. and Lapiere, C.M. (1962) Collagenolytic activity in amphibian tissues: a tissue culture assay. *Proc. Natl. Acad. Sci. U. S. A.*, **48**, 1014-1022.
- 262 Klein, T. and Bischoff, R. (2011) Physiology and pathophysiology of matrix metalloproteinases. *Amino Acids*, **41**, 271-290.
- 263 Murphy, G. and Nagase, H. (2008) Progress in matrix metalloproteinase research. *Mol. Aspects Med.*, **29**, 290-308.
- 264 Vu, T.H. and Werb, Z. (2000) Matrix metalloproteinases: effectors of development and normal physiology. *Genes Dev.*, **14**, 2123-2133.
- 265 Cui, N., Hu, M. and Khalil, R.A. (2017) Biochemical and Biological Attributes of Matrix Metalloproteinases. *Prog. Mol. Biol. Transl. Sci.*, **147**, 1-73.
- 266 Ra, H.J. and Parks, W.C. (2007) Control of matrix metalloproteinase catalytic activity. *Matrix Biol.*, **26**, 587-596.

- 267 Moinfar, F., Man, Y.G., Arnould, L., Bratthauer, G.L., Ratschek, M. and Tavassoli, F.A. (2000) Concurrent and independent genetic alterations in the stromal and epithelial cells of mammary carcinoma: implications for tumorigenesis. *Cancer Res.*, **60**, 2562-2566.
- 268 Willenbacher, R.F., Aust, D.E., Chang, C.G., Zelman, S.J., Ferrell, L.D., Moore, D.H., 2nd and Waldman, F.M. (1999) Genomic instability is an early event during the progression pathway of ulcerative-colitis-related neoplasia. *Am. J. Pathol.*, **154**, 1825-1830.
- 269 Sternlicht, M.D., Coussens, L.M., Vu, T.H. and Werb, Z. (2001) Clendeninn, N.J. and Appelt, K. (eds.), In *Matrix Metalloproteinase Inhibitors in Cancer Therapy*. Humana Press, Totowa, NJ, in press., pp. 1-37.
- 270 Deryugina, E.I. and Quigley, J.P. (2015) Tumor angiogenesis: MMP-mediated induction of intravasation- and metastasis-sustaining neovasculature. *Matrix Biol.*, **44-46**, 94-112.
- 271 Wilson, C.L., Heppner, K.J., Labosky, P.A., Hogan, B.L. and Matrisian, L.M. (1997) Intestinal tumorigenesis is suppressed in mice lacking the metalloproteinase matrilysin. *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 1402-1407.
- 272 Colandrea, T.D., D'Armiento, J., Kesari, K.V. and Chada, K.K. (2000) Collagenase induction promotes mouse tumorigenesis by two independent pathways. *Mol. Carcinog.*, **29**, 8-16.
- 273 Coussens, L.M., Tinkle, C.L., Hanahan, D. and Werb, Z. (2000) MMP-9 supplied by bone marrow-derived cells contributes to skin carcinogenesis. *Cell*, **103**, 481-490.
- 274 Gross, J. and Lapiere, C.M. (1962) Collagenolytic activity in amphibian tissues: a tissue culture assay. *Proc. Natl. Acad. Sci. U. S. A.*, **48**, 1014-1022.
- 275 Nelson, A.R., Fingleton, B., Rothenberg, M.L. and Matrisian, L.M. (2000) Matrix metalloproteinases: biologic activity and clinical implications. *J. Clin. Oncol.*, **18**, 1135-1149.
- 276 Quintero-Fabián, S., Arreola, R., Becerril-Villanueva, E., Torres-Romero, J.C., Arana-Argáez, V., Lara-Riegos, J., Ramírez-Camacho, M.A. and Alvarez-Sánchez, M.E. (2019) Role of Matrix Metalloproteinases in Angiogenesis and Cancer. *Front. Oncol.*, **9**, 1370.
- 277 Uzui, H., Lee, J.D., Shimizu, H., Tsutani, H. and Ueda, T. (2000) The role of protein-tyrosine phosphorylation and gelatinase production in the migration and proliferation of smooth muscle cells. *Atherosclerosis*, **149**, 51-59.
- 278 Kräling, B.M., Wiederschain, D.G., Boehm, T., Rehn, M., Mulliken, J.B. and Moses, M.A. (1999) The role of matrix metalloproteinase activity in the maturation of human capillary endothelial cells in vitro. *J. Cell Sci.*, **112 (Pt 10)**, 1599-1609.
- 279 Boudreau, N., Werb, Z. and Bissell, M.J. (1996) Suppression of apoptosis by basement membrane requires three-dimensional tissue organization and withdrawal from the cell cycle. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 3509-3513.
- 280 Schedin, P., Strange, R., Mitrenga, T., Wolfe, P. and Kaeck, M. (2000) Fibronectin fragments induce MMP activity in mouse mammary epithelial cells: evidence for a role in mammary tissue remodeling. *J. Cell Sci.*, **113 (Pt 5)**, 795-806.
- 281 Cowan, K.N., Jones, P.L. and Rabinovitch, M. (2000) Elastase and matrix metalloproteinase inhibitors induce regression, and tenascin-C antisense prevents progression, of vascular disease. *J. Clin. Invest.*, **105**, 21-34.
- 282 Brown, L.M., Fox, H.L., Hazen, S.A., LaNoue, K.F., Rannels, S.R. and Lynch, C.J. (1997) Role of the matrixin MMP-2 in multicellular organization of adipocytes cultured in basement membrane components. *Am. J. Physiol.*, **272**, C937-949.
- 283 Miralles, F., Battelino, T., Czernichow, P. and Scharfmann, R. (1998) TGF-beta plays a key role in morphogenesis of the pancreatic islets of Langerhans by controlling the activity of the matrix metalloproteinase MMP-2. *J. Cell Biol.*, **143**, 827-836.

- 284 Schnaper, H.W., Grant, D.S., Stetler-Stevenson, W.G., Fridman, R., D'Orazi, G., Murphy, A.N., Bird, R.E., Hoythya, M., Fuerst, T.R., French, D.L. *et al.* (1993) Type IV collagenase(s) and TIMPs modulate endothelial cell morphogenesis in vitro. *J. Cell. Physiol.*, **156**, 235-246.
- 285 Fisher, C., Gilbertson-Beadling, S., Powers, E.A., Petzold, G., Poorman, R. and Mitchell, M.A. (1994) Interstitial collagenase is required for angiogenesis in vitro. *Dev. Biol.*, **162**, 499-510.
- 286 Pahwa, S., Stawikowski, M.J. and Fields, G.B. (2014) Monitoring and Inhibiting MT1-MMP during Cancer Initiation and Progression. *Cancers (Basel)*, **6**, 416-435.
- 287 Whitelock, J.M., Murdoch, A.D., Iozzo, R.V. and Underwood, P.A. (1996) The degradation of human endothelial cell-derived perlecan and release of bound basic fibroblast growth factor by stromelysin, collagenase, plasmin, and heparanases. *J. Biol. Chem.*, **271**, 10079-10086.
- 288 Imai, K., Hiramatsu, A., Fukushima, D., Pierschbacher, M.D. and Okada, Y. (1997) Degradation of decorin by matrix metalloproteinases: identification of the cleavage sites, kinetic analyses and transforming growth factor-beta1 release. *Biochem. J.*, **322 (Pt 3)**, 809-814.
- 289 Rajah, R., Katz, L., Nunn, S., Solberg, P., Beers, T. and Cohen, P. (1995) Insulin-like growth factor binding protein (IGFBP) proteases: functional regulators of cell growth. *Prog. Growth Factor Res.*, **6**, 273-284.
- 290 Thrailkill, K.M., Quarles, L.D., Nagase, H., Suzuki, K., Serra, D.M. and Fowlkes, J.L. (1995) Characterization of insulin-like growth factor-binding protein 5-degrading proteases produced throughout murine osteoblast differentiation. *Endocrinology*, **136**, 3527-3533.
- 291 Weng, C.J., Chen, M.K., Lin, C.W., Chung, T.T. and Yang, S.F. (2012) Single nucleotide polymorphisms and haplotypes of MMP-14 are associated with the risk and pathological development of oral cancer. *Ann. Surg. Oncol.*, **19 Suppl 3**, S319-327.
- 292 de Vos, I., Tao, E.Y., Ong, S.L.M., Goggi, J.L., Scerri, T., Wilson, G.R., Low, C.G.M., Wong, A.S.W., Grussu, D., Stegmann, A.P.A. *et al.* (2018) Functional analysis of a hypomorphic allele shows that MMP14 catalytic activity is the prime determinant of the Winchester syndrome phenotype. *Hum. Mol. Genet.*, **27**, 2775-2788.
- 293 Qi, Y., Wang, J., Sun, M., Ma, C., Jin, T., Liu, Y., Cao, Y. and Wang, J. (2019) MMP-14 single-nucleotide polymorphisms are related to steroid-induced osteonecrosis of the femoral head in the population of northern China. *Mol Genet Genomic Med*, **7**, e00519.
- 294 Zhang, J., Sun, X., Liu, J., Liu, J., Shen, B. and Nie, L. (2015) The role of matrix metalloproteinase 14 polymorphisms in susceptibility to intervertebral disc degeneration in the Chinese Han population. *Arch. Med. Sci.*, **11**, 801-806.
- 295 Munkert, A., Helmchen, U., Kemper, M.J., Bubenheim, M., Stahl, R.A. and Harendza, S. (2009) Characterization of the transcriptional regulation of the human MT1-MMP gene and association of risk reduction for focal-segmental glomerulosclerosis with two functional promoter SNPs. *Nephrol. Dial. Transplant.*, **24**, 735-742.
- 296 Zhou, Q.H., Huang, X.F., Wang, J.H., Lin, C.W., Yang, Y.Y., Huang, C.S., Wu, L.T. and Wu, Y.M. (2012) [Association of MMP14 gene polymorphisms and osteoporosis in Zhuang men from Baise region of Guangxi]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi*, **29**, 309-313.
- 297 Tee, Y.T., Liu, Y.F., Chang, J.T., Yang, S.F., Chen, S.C., Han, C.P., Wang, P.H. and Liao, C.L. (2012) Single-nucleotide polymorphisms and haplotypes of membrane type 1-matrix metalloproteinase in susceptibility and clinical significance of squamous cell neoplasia of uterine cervix in Taiwan women. *Reprod. Sci.*, **19**, 932-938.
- 298 Beeghly-Fadiel, A., Shu, X.O., Long, J., Li, C., Cai, Q., Cai, H., Gao, Y.T. and Zheng, W. (2009) Genetic polymorphisms in the MMP-7 gene and breast cancer survival. *Int. J. Cancer*, **124**, 208-214.
- 299 Kesh, K., Subramanian, L., Ghosh, N., Gupta, V., Gupta, A., Bhattacharya, S., Mahapatra, N.R. and Swarnakar, S. (2015) Association of MMP7 -181A-->G Promoter Polymorphism with

- Gastric Cancer Risk: INFLUENCE OF NICOTINE IN DIFFERENTIAL ALLELE-SPECIFIC TRANSCRIPTION VIA INCREASED PHOSPHORYLATION OF cAMP-RESPONSE ELEMENT-BINDING PROTEIN (CREB). *J. Biol. Chem.*, **290**, 14391-14406.
- 300 Langers, A.M., Verspaget, H.W., Hommes, D.W. and Sier, C.F. (2011) Single-nucleotide polymorphisms of matrix metalloproteinases and their inhibitors in gastrointestinal cancer. *World. J. Gastrointest. Oncol.*, **3**, 79-98.
- 301 Chen, G.L., Shen, T.C., Chang, W.S., Tsai, C.W., Li, H.T., Chuang, C.L., Lai, Y.L., Yueh, T.C., Hsia, T.C., Wang, S.C. *et al.* (2018) The Contribution of MMP-7 Promoter Polymorphisms to Taiwan Lung Cancer Susceptibility. *Anticancer Res.*, **38**, 5671-5677.
- 302 Jormsjo, S., Whatling, C., Walter, D.H., Zeiher, A.M., Hamsten, A. and Eriksson, P. (2001) Allele-specific regulation of matrix metalloproteinase-7 promoter activity is associated with coronary artery luminal dimensions among hypercholesterolemic patients. *Arterioscler. Thromb. Vasc. Biol.*, **21**, 1834-1839.
- 303 Richards, T.J., Park, C., Chen, Y., Gibson, K.F., Peter Di, Y., Pardo, A., Watkins, S.C., Choi, A.M., Selman, M., Pilewski, J. *et al.* (2012) Allele-specific transactivation of matrix metalloproteinase 7 by FOXA2 and correlation with plasma levels in idiopathic pulmonary fibrosis. *Am. J. Physiol. Lung Cell Mol. Physiol.*, **302**, L746-754.
- 304 Subramanian, L., Maghajothi, S., Singh, M., Kesh, K., Kalyani, A., Sharma, S., Khullar, M., Victor, S.M., Swarnakar, S., Asthana, S. *et al.* (2019) A Common Tag Nucleotide Variant in MMP7 Promoter Increases Risk for Hypertension via Enhanced Interactions With CREB (Cyclic AMP Response Element-Binding Protein) Transcription Factor. *Hypertension*, **74**, 1448-1459.
- 305 Zhang, J., Jin, X., Fang, S., Wang, R., Li, Y., Wang, N., Guo, W., Wang, Y., Wen, D., Wei, L. *et al.* (2005) The functional polymorphism in the matrix metalloproteinase-7 promoter increases susceptibility to esophageal squamous cell carcinoma, gastric cardiac adenocarcinoma and non-small cell lung carcinoma. *Carcinogenesis*, **26**, 1748-1753.
- 306 Tacheva, T., Chelenkova, P., Dimov, D., Petkova, R., Chakarov, S. and Vlaykova, T. (2015) Frequency of the common promoter polymorphism MMP2 -1306 C>T in a population from central Bulgaria. *Biotechnol Biotechnol Equip*, **29**, 351-356.
- 307 Yadav, S.S., Mandal, R.K., Singh, M.K., Usman, K. and Khattri, S. (2014) Genetic variants of matrix metalloproteinase (MMP2) gene influence metabolic syndrome susceptibility. *Genet. Test. Mol. Biomarkers*, **18**, 88-92.
- 308 Marson, B.P., Lacchini, R., Belo, V., Dickel, S., da Costa, B.P., Poli de Figueiredo, C.E. and Tanus-Santos, J.E. (2012) Matrix metalloproteinase (MMP)-2 genetic variants modify the circulating MMP-2 levels in end-stage kidney disease. *Am. J. Nephrol.*, **35**, 209-215.
- 309 Vasků, A., Goldbergová, M., Izakovicová Hollá, L., Sisková, L., Groch, L., Beránek, M., Tschöplová, S., Znojil, V. and Vácha, J. (2004) A haplotype constituted of four MMP-2 promoter polymorphisms (-1575G/A, -1306C/T, -790T/G and -735C/T) is associated with coronary triple-vessel disease. *Matrix Biol.*, **22**, 585-591.
- 310 Alg, V.S., Ke, X., Grieve, J., Bonner, S., Walsh, D.C., Bulters, D., Kitchen, N., Houlden, H. and Werring, D.J. (2018) Association of functional MMP-2 gene variant with intracranial aneurysms: case-control genetic association study and meta-analysis. *Br. J. Neurosurg.*, **32**, 255-259.
- 311 Durmanova, V., Javor, J., Parnicka, Z., Minarik, G., Ocenasova, A., Vaseckova, B., Reznakova, V., Kralova, M., Hromadka, T. and Shawkatova, I. (2021) Impact of MMP2 rs243865 and MMP3 rs3025058 Polymorphisms on Clinical Findings in Alzheimer's Disease Patients. *Mediators Inflamm.*, **2021**, 5573642.
- 312 Zhou, Y., Yu, C., Miao, X., Wang, Y., Tan, W., Sun, T., Zhang, X., Xiong, P. and Lin, D. (2005) Functional haplotypes in the promoter of matrix metalloproteinase-2 and lung cancer susceptibility. *Carcinogenesis*, **26**, 1117-1121.

- 313 Zhou, G., Zhai, Y., Cui, Y., Qiu, W., Yang, H., Zhang, X., Dong, X., He, Y., Yao, K., Zhang, H. *et al.* (2007) Functional polymorphisms and haplotypes in the promoter of the MMP2 gene are associated with risk of nasopharyngeal carcinoma. *Hum. Mutat.*, **28**, 1091-1097.
- 314 Yu, C., Zhou, Y., Miao, X., Xiong, P., Tan, W. and Lin, D. (2004) Functional haplotypes in the promoter of matrix metalloproteinase-2 predict risk of the occurrence and metastasis of esophageal cancer. *Cancer Res.*, **64**, 7622-7628.
- 315 Huang, C., Xu, S., Luo, Z., Li, D., Wang, R. and Wang, T. (2022) Epidemiological Evidence Between Variants in Matrix Metalloproteinases-2, -7, and -9 and Cancer Risk. *Front. Oncol.*, **12**, 856831.
- 316 Price, S.J., Greaves, D.R. and Watkins, H. (2001) Identification of novel, functional genetic variants in the human matrix metalloproteinase-2 gene: role of Sp1 in allele-specific transcriptional regulation. *J. Biol. Chem.*, **276**, 7549-7558.
- 317 Sambrook, J. and Russell, D.W. (2006) The inoue method for preparation and transformation of competent e. Coli: "ultra-competent" cells. *CSH protocols*, **2006**.
- 318 Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671-675.
- 319 Fedchenko, N. and Reifenrath, J. (2014) Different approaches for interpretation and reporting of immunohistochemistry analysis results in the bone tissue - a review. *Diagn. Pathol.*, **9**, 221.
- 320 Sambrook, J. and Russell, D.W. (2006) SDS-Polyacrylamide Gel Electrophoresis of Proteins. *CSH protocols*, **2006**.
- 321 Turunen, S.P., Tatti-Bugaeva, O. and Lehti, K. (2017) Membrane-type matrix metalloproteases as diverse effectors of cancer progression. *Biochim Biophys Acta Mol Cell Res*, **1864**, 1974-1988.
- 322 Gonzalez-Avila, G., Sommer, B., Mendoza-Posada, D.A., Ramos, C., Garcia-Hernandez, A.A. and Falfan-Valencia, R. (2019) Matrix metalloproteinases participation in the metastatic process and their diagnostic and therapeutic applications in cancer. *Crit. Rev. Oncol. Hematol.*, **137**, 57-83.
- 323 Ren, Z.H., Wu, K., Yang, R., Liu, Z.Q. and Cao, W. (2020) Differential expression of matrix metalloproteinases and miRNAs in the metastasis of oral squamous cell carcinoma. *BMC Oral Health*, **20**, 24.
- 324 Neuhaus, J., Schiffer, E., Mannello, F., Horn, L.C., Ganzer, R. and Stolzenburg, J.U. (2017) Protease Expression Levels in Prostate Cancer Tissue Can Explain Prostate Cancer-Associated Seminal Biomarkers-An Explorative Concept Study. *Int J Mol Sci*, **18**.
- 325 Gonzalez-Molina, J., Gramolelli, S., Liao, Z., Carlson, J.W., Ojala, P.M. and Lehti, K. (2019) MMP14 in Sarcoma: A Regulator of Tumor Microenvironment Communication in Connective Tissues. *Cells*, **8**.
- 326 Ota, I., Li, X.Y., Hu, Y. and Weiss, S.J. (2009) Induction of a MT1-MMP and MT2-MMP-dependent basement membrane transmigration program in cancer cells by Snail1. *Proc Natl Acad Sci U S A*, **106**, 20318-20323.
- 327 Sabeh, F., Li, X.Y., Saunders, T.L., Rowe, R.G. and Weiss, S.J. (2009) Secreted versus membrane-anchored collagenases: relative roles in fibroblast-dependent collagenolysis and invasion. *J Biol Chem*, **284**, 23001-23011.
- 328 Yusenko, M.V., Trentmann, A., Andersson, M.K., Ghani, L.A., Jakobs, A., Arteaga Paz, M.F., Mikesch, J.H., Peter von Kries, J., Stenman, G. and Klempnauer, K.H. (2020) Monensin, a novel potent MYB inhibitor, suppresses proliferation of acute myeloid leukemia and adenoid cystic carcinoma cells. *Cancer Lett.*, **479**, 61-70.
- 329 Information, N.C.f.B. (2022) PubChem Substance Record for SID 24896992, Monensin sodium salt, 90-95% (TLC), Source: Sigma-Aldrich. Retrieved July 30, 2022. *PubChem*, in press.

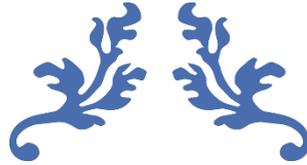
- 330 Han, K.Y., Dugas-Ford, J., Seiki, M., Chang, J.H. and Azar, D.T. (2015) Evidence for the Involvement of MMP14 in MMP2 Processing and Recruitment in Exosomes of Corneal Fibroblasts. *Invest. Ophthalmol. Vis. Sci.*, **56**, 5323-5329.
- 331 Sun, J. (2010) Matrix metalloproteinases and tissue inhibitor of metalloproteinases are essential for the inflammatory response in cancer cells. *J Signal Transduct*, **2010**, 985132.
- 332 Nyalendo, C., Sartelet, H., Gingras, D. and Beliveau, R. (2010) Inhibition of membrane-type 1 matrix metalloproteinase tyrosine phosphorylation blocks tumor progression in mice. *Anticancer Res.*, **30**, 1887-1895.
- 333 Qiang, L., Cao, H., Chen, J., Weller, S.G., Krueger, E.W., Zhang, L., Razidlo, G.L. and McNiven, M.A. (2019) Pancreatic tumor cell metastasis is restricted by MT1-MMP binding protein MTCBP-1. *J. Cell Biol.*, **218**, 317-332.
- 334 Coussens, L.M. and Werb, Z. (2002) Inflammation and cancer. *Nature*, **420**, 860-867.
- 335 Dufour, A. and Overall, C.M. (2013) Missing the target: matrix metalloproteinase antitargets in inflammation and cancer. *Trends Pharmacol. Sci.*, **34**, 233-242.
- 336 Kasurinen, A., Gramolelli, S., Hagstrom, J., Laitinen, A., Kokkola, A., Miki, Y., Lehti, K., Yashiro, M., Ojala, P.M., Bockelman, C. *et al.* (2019) High tissue MMP14 expression predicts worse survival in gastric cancer, particularly with a low PROX1. *Cancer Med*, **8**, 6995-7005.
- 337 Cui, G., Cai, F., Ding, Z. and Gao, L. (2019) MMP14 predicts a poor prognosis in patients with colorectal cancer. *Hum. Pathol.*, **83**, 36-42.
- 338 Kirimlioglu, H., Turkmen, I., Bassullu, N., Dirican, A., Karadag, N. and Kirimlioglu, V. (2009) The expression of matrix metalloproteinases in intrahepatic cholangiocarcinoma, hilar (Klatskin tumor), middle and distal extrahepatic cholangiocarcinoma, gallbladder cancer, and ampullary carcinoma: role of matrix metalloproteinases in tumor progression and prognosis. *Turk. J. Gastroenterol.*, **20**, 41-47.
- 339 Ridley, A.J. (2011) Life at the leading edge. *Cell*, **145**, 1012-1022.
- 340 Ciciro, Y. and Sala, A. (2021) MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. *Oncogenesis*, **10**, 19.
- 341 Liu, X., Xu, Y., Han, L. and Yi, Y. (2018) Reassessing the Potential of Myb-targeted Anti-cancer Therapy. *J Cancer*, **9**, 1259-1266.
- 342 Ramsay, R.G. and Gonda, T.J. (2008) MYB function in normal and cancer cells. *Nat Rev Cancer*, **8**, 523-534.
- 343 Rihova, K., Ducka, M., Zambo, I.S., Vymetalova, L., Sramek, M., Trcka, F., Verner, J., Drapela, S., Fedr, R., Suchankova, T. *et al.* (2022) Transcription factor c-Myb: novel prognostic factor in osteosarcoma. *Clin Exp Metastasis*, in press.
- 344 Geetha, A. (2002) Evidence for oxidative stress in the gall bladder mucosa of gall stone patients. *J Biochem Mol Biol Biophys*, **6**, 427-432.
- 345 Bies, J., Sramko, M. and Wolff, L. (2013) Stress-induced phosphorylation of Thr486 in c-Myb by p38 mitogen-activated protein kinases attenuates conjugation of SUMO-2/3. *J Biol Chem*, **288**, 36983-36993.
- 346 Pekarcikova, L., Knopfova, L., Benes, P. and Smarda, J. (2016) c-Myb regulates NOX1/p38 to control survival of colorectal carcinoma cells. *Cell Signal*, **28**, 924-936.
- 347 Kriegsmann, K., Flechtenmacher, C., Heil, J., Kriegsmann, J., Mechtersheimer, G., Aulmann, S., Weichert, W., Sinn, H.P. and Kriegsmann, M. (2020) Immunohistological Expression of SOX-10 in Triple-Negative Breast Cancer: A Descriptive Analysis of 113 Samples. *Int J Mol Sci*, **21**.
- 348 Yin, H., Qin, C., Zhao, Y., Du, Y., Sheng, Z., Wang, Q., Song, Q., Chen, L., Liu, C. and Xu, T. (2017) SOX10 is over-expressed in bladder cancer and contributes to the malignant bladder cancer cell behaviors. *Clin Transl Oncol*, **19**, 1035-1044.

- 349 Zhao, Y., Liu, Z.G., Tang, J., Zou, R.F., Chen, X.Y., Jiang, G.M., Qiu, Y.F. and Wang, H. (2016) High expression of Sox10 correlates with tumor aggressiveness and poor prognosis in human nasopharyngeal carcinoma. *Onco Targets Ther*, **9**, 1671-1677.
- 350 Mollaaghbababa, R. and Pavan, W.J. (2003) The importance of having your SOX on: role of SOX10 in the development of neural crest-derived melanocytes and glia. *Oncogene*, **22**, 3024-3034.
- 351 Garmon, T., Wittling, M. and Nie, S. (2018) MMP14 Regulates Cranial Neural Crest Epithelial-to-Mesenchymal Transition and Migration. *Dev Dyn*, **247**, 1083-1092.
- 352 Andrieu, C., Montigny, A., Bibonne, A., Despin-Guitard, E., Alfandari, D. and Theveneau, E. (2020) MMP14 is required for delamination of chick neural crest cells independently of its catalytic activity. *Development*, **147**.
- 353 Zhou, D., Bai, F., Zhang, X., Hu, M., Zhao, G., Zhao, Z. and Liu, R. (2014) SOX10 is a novel oncogene in hepatocellular carcinoma through Wnt/beta-catenin/TCF4 cascade. *Tumour Biol*, **35**, 9935-9940.
- 354 Kormish, J.D., Sinner, D. and Zorn, A.M. (2010) Interactions between SOX factors and Wnt/beta-catenin signaling in development and disease. *Dev Dyn*, **239**, 56-68.
- 355 Uka, R., Britschgi, C., Krattli, A., Matter, C., Mihic, D., Okoniewski, M.J., Gualandi, M., Stupp, R., Cinelli, P., Dummer, R. *et al.* (2020) Temporal activation of WNT/beta-catenin signaling is sufficient to inhibit SOX10 expression and block melanoma growth. *Oncogene*, **39**, 4132-4154.
- 356 Liu, P., Yang, J., Pei, J., Pei, D. and Wilson, M.J. (2010) Regulation of MT1-MMP activity by beta-catenin in MDCK non-cancer and HT1080 cancer cells. *J Cell Physiol*, **225**, 810-821.
- 357 Li, Y., Jin, K., van Pelt, G.W., van Dam, H., Yu, X., Mesker, W.E., Ten Dijke, P., Zhou, F. and Zhang, L. (2016) c-Myb Enhances Breast Cancer Invasion and Metastasis through the Wnt/beta-Catenin/Axin2 Pathway. *Cancer Res*, **76**, 3364-3375.
- 358 Ciznadija, D., Tothill, R., Waterman, M.L., Zhao, L., Huynh, D., Yu, R.M., Ernst, M., Ishii, S., Mantamadiotis, T., Gonda, T.J. *et al.* (2009) Intestinal adenoma formation and MYC activation are regulated by cooperation between MYB and Wnt signaling. *Cell Death Differ*, **16**, 1530-1538.
- 359 Saarialho-Kere, U.K., Crouch, E.C. and Parks, W.C. (1995) Matrix Metalloproteinase Matrilysin Is Constitutively Expressed in Adult Human Exocrine Epithelium. *J. Invest. Dermatol.*, **105**, 190-196.
- 360 Yamamoto, H., Adachi, Y., Itoh, F., Iku, S., Matsuno, K., Kusano, M., Arimura, Y., Endo, T., Hinoda, Y., Hosokawa, M. *et al.* (1999) Association of matrilysin expression with recurrence and poor prognosis in human esophageal squamous cell carcinoma. *Cancer Res.*, **59**, 3313-3316.
- 361 Yamamoto, H., Iku, S., Adachi, Y., Imsumran, A., Taniguchi, H., Nosho, K., Min, Y., Horiuchi, S., Yoshida, M., Itoh, F. *et al.* (2003) Association of trypsin expression with tumour progression and matrilysin expression in human colorectal cancer. *J. Pathol.*, **199**, 176-184.
- 362 Yamamoto, H., Itoh, F., Iku, S., Adachi, Y., Fukushima, H., Sasaki, S., Mukaiya, M., Hirata, K. and Imai, K. (2001) Expression of matrix metalloproteinases and tissue inhibitors of metalloproteinases in human pancreatic adenocarcinomas: clinicopathologic and prognostic significance of matrilysin expression. *J. Clin. Oncol.*, **19**, 1118-1127.
- 363 Jones, L.E., Humphreys, M.J., Campbell, F., Neoptolemos, J.P. and Boyd, M.T. (2004) Comprehensive analysis of matrix metalloproteinase and tissue inhibitor expression in pancreatic cancer: increased expression of matrix metalloproteinase-7 predicts poor survival. *Clin. Cancer Res.*, **10**, 2832-2845.
- 364 Cai, X.-Y., Guo, J.-G., Shen, X.-M., Dong, J., Hu, W.-H., Guo, G.-F., He, X.-Z., Xia, L.-P., Hu, P.-L., Qiu, H.-J. *et al.* (2016) The MMP-7-181 A/G gene polymorphism is a prognostic indicator in patients with gastric cancer. *2016*, **5**, 748-755.

- 365 Moskalenko, M., Ponomarenko, I., Reshetnikov, E., Dvornyk, V. and Churnosov, M. (2021) Polymorphisms of the matrix metalloproteinase genes are associated with essential hypertension in a Caucasian population of Central Russia. *Sci. Rep.*, **11**, 5224.
- 366 Fan, Y.M., Hernesniemi, J., Oksala, N., Levula, M., Raitoharju, E., Collings, A., Hutri-Kähönen, N., Juonala, M., Marniemi, J., Lyytikäinen, L.P. *et al.* (2014) Upstream Transcription Factor 1 (USF1) allelic variants regulate lipoprotein metabolism in women and USF1 expression in atherosclerotic plaque. *Sci. Rep.*, **4**, 4650.
- 367 Gao, Y., Nan, X., Shi, X., Mu, X., Liu, B., Zhu, H., Yao, B., Liu, X., Yang, T., Hu, Y. *et al.* (2019) SREBP1 promotes the invasion of colorectal cancer accompanied upregulation of MMP7 expression and NF- κ B pathway activation. *BMC Cancer*, **19**, 685.
- 368 Hofmann, U.B., Westphal, J.R., Zendman, A.J., Becker, J.C., Ruiter, D.J. and van Muijen, G.N. (2000) Expression and activation of matrix metalloproteinase-2 (MMP-2) and its co-localization with membrane-type 1 matrix metalloproteinase (MT1-MMP) correlate with melanoma progression. *J. Pathol.*, **191**, 245-256.
- 369 Fan, W.H. and Karnovsky, M.J. (2002) Increased MMP-2 expression in connective tissue growth factor over-expression vascular smooth muscle cells. *J. Biol. Chem.*, **277**, 9800-9805.
- 370 Qiao, Z.K., Li, Y.L., Lu, H.T., Wang, K.L. and Xu, W.H. (2013) Expression of tissue levels of matrix metalloproteinases and tissue inhibitors of metalloproteinases in renal cell carcinoma. *World J. Surg. Oncol.*, **11**, 1.
- 371 Chen, L.H., Chiu, K.L., Hsia, T.C., Lee, Y.H., Shen, T.C., Li, C.H., Shen, Y.C., Chang, W.S., Tsai, C.W. and Bau, D.T. (2020) Significant Association of MMP2 Promoter Genotypes to Asthma Susceptibility in Taiwan. *In Vivo*, **34**, 3181-3186.
- 372 Srivastava, P., Lone, T.A., Kapoor, R. and Mittal, R.D. (2012) Association of promoter polymorphisms in MMP2 and TIMP2 with prostate cancer susceptibility in North India. *Arch. Med. Res.*, **43**, 117-124.
- 373 Srivastava, P., Kapoor, R. and Mittal, R.D. (2013) Association of single nucleotide polymorphisms in promoter of matrix metalloproteinase-2, 8 genes with bladder cancer risk in Northern India. *Urol. Oncol.*, **31**, 247-254.
- 374 Yete, S., Pradhan, S. and Saranath, D. (2017) Single nucleotide polymorphisms in an Indian cohort and association of CNTN4, MMP2 and SNTB1 variants with oral cancer. *Cancer Genet.*, **214-215**, 16-25.
- 375 Peng, B., Cao, L., Ma, X., Wang, W., Wang, D. and Yu, L. (2010) Meta-analysis of association between matrix metalloproteinases 2, 7 and 9 promoter polymorphisms and cancer risk. *Mutagenesis*, **25**, 371-379.
- 376 Zhang, K., Chen, X., Zhou, J., Yang, C., Zhang, M., Chao, M., Zhang, L. and Liang, C. (2017) Association between MMP2-1306 C/T polymorphism and prostate cancer susceptibility: a meta-analysis based on 3906 subjects. *Oncotarget*, **8**, 45020-45029.
- 377 Song, J., Nabeel-Shah, S., Pu, S., Lee, H., Braunschweig, U., Ni, Z., Ahmed, N., Marcon, E., Zhong, G., Ray, D. *et al.* (2022) Regulation of alternative polyadenylation by the C2H2-zinc-finger protein Sp1. *Mol. Cell*, **82**, 3135-3150.e3139.
- 378 Xu, S., Zhan, M. and Wang, J. (2017) Epithelial-to-mesenchymal transition in gallbladder cancer: from clinical evidence to cellular regulatory networks. *Cell Death Discovery*, **3**, 17069.
- 379 Xu, S., Jiang, C., Lin, R., Wang, X., Hu, X., Chen, W., Chen, X. and Chen, T. (2021) Epigenetic activation of the elongator complex sensitizes gallbladder cancer to gemcitabine therapy. *J. Exp. Clin. Cancer Res.*, **40**, 373.
- 380 Cabral-Pacheco, G.A., Garza-Veloz, I., Castruita-De la Rosa, C., Ramirez-Acuña, J.M., Perez-Romero, B.A., Guerrero-Rodriguez, J.F., Martinez-Avila, N. and Martinez-Fierro, M.L. (2020)

- The Roles of Matrix Metalloproteinases and Their Inhibitors in Human Diseases. *Int. J. Mol. Sci.*, **21**.
- 381 Gobin, E., Bagwell, K., Wagner, J., Mysona, D., Sandirasegarane, S., Smith, N., Bai, S., Sharma, A., Schleifer, R. and She, J.-X. (2019) A pan-cancer perspective of matrix metalloproteinases (MMP) gene expression profile and their diagnostic/prognostic potential. *BMC Cancer*, **19**, 581.
- 382 Walker, C., Mojares, E. and Del Río Hernández, A. (2018) Role of Extracellular Matrix in Development and Cancer Progression. **19**, 3028.
- 383 Hu, Z.I. and Lim, K.-H. (2022) Evolving Paradigms in the Systemic Treatment of Advanced Gallbladder Cancer: Updates in Year 2022. **14**, 1249.
- 384 Nemunaitis, J.M., Brown-Glabeman, U., Soares, H., Belmonte, J., Liem, B., Nir, I., Phuoc, V. and Gullapalli, R.R. (2018) Gallbladder cancer: review of a rare orphan gastrointestinal cancer with a focus on populations of New Mexico. *BMC Cancer*, **18**, 665.
- 385 Gialeli, C., Theocharis, A.D. and Karamanos, N.K. (2011) Roles of matrix metalloproteinases in cancer progression and their pharmacological targeting. *FEBS J.*, **278**, 16-27.
- 386 Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
- 387 Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. *et al.* (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. **102**, 15545-15550.
- 388 Merico, D., Isserlin, R., Stueker, O., Emili, A. and Bader, G.D. (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One*, **5**, e13984.
- 389 Li, Y., Ge, D. and Lu, C. (2019) The SMART App: an interactive web application for comprehensive DNA methylation analysis and visualization. *Epigenetics & Chromatin*, **12**, 71.
- 390 Martino, F., Perestrelo, A.R., Vinarský, V., Pagliari, S. and Forte, G. (2018) Cellular Mechanotransduction: From Tension to Function. **9**.
- 391 Harjunpää, H., Lloret Asens, M., Guenther, C. and Fagerholm, S.C. (2019) Cell Adhesion Molecules and Their Roles and Regulation in the Immune and Tumor Microenvironment. **10**.
- 392 Guan, X. (2015) Cancer metastases: challenges and opportunities. *Acta pharmaceutica Sinica. B*, **5**, 402-418.
- 393 Bonnans, C., Chou, J. and Werb, Z. (2014) Remodelling the extracellular matrix in development and disease. *Nat. Rev. Mol. Cell Biol.*, **15**, 786-801.
- 394 Mohan, V., Das, A. and Sagi, I. (2020) Emerging roles of ECM remodeling processes in cancer. *Semin. Cancer Biol.*, **62**, 192-200.
- 395 Campbell, K.M., Thaker, M., Medina, E., Kalbasi, A., Singh, A., Ribas, A. and Nowicki, T.S. (2022) Spatial profiling reveals association between WNT pathway activation and T-cell exclusion in acquired resistance of synovial sarcoma to NY-ESO-1 transgenic T-cell therapy. **10**, e004190.
- 396 Khokha, R., Murthy, A. and Weiss, A. (2013) Metalloproteinases and their natural inhibitors in inflammation and immunity. *Nat Rev Immunol*, **13**, 649-665.
- 397 Amar, S. and Fields, G.B. (2015) Potential clinical implications of recent matrix metalloproteinase inhibitor design strategies. *Expert Rev Proteomics*, **12**, 445-447.
- 398 Sternlicht, M.D. and Werb, Z. (2001) How matrix metalloproteinases regulate cell behavior. *Annu. Rev. Cell. Dev. Biol.*, **17**, 463-516.
- 399 Xie, B., Zhang, Z., Wang, H., Chen, Z., Wang, Y., Liang, H., Yang, G., Yang, X. and Zhang, H. (2016) Genetic polymorphisms in MMP 2, 3, 7, and 9 genes and the susceptibility and clinical outcome of cervical cancer in a Chinese Han population. *Tumour Biol.*, **37**, 4883-4888.

-
- 400 Kim, Y.R., Jeon, Y.J., Kim, H.S., Kim, J.O., Moon, M.J., Ahn, E.H., Lee, W.S. and Kim, N.K. (2015) Association study of five functional polymorphisms in matrix metalloproteinase-2, -3, and -9 genes with risk of primary ovarian insufficiency in Korean women. *Maturitas*, **80**, 192-197.
- 401 Wieczorek, E., Reszka, E., Jablonowski, Z., Jablonska, E., Krol, M.B., Grzegorzczak, A., Gromadzinska, J., Sosnowski, M. and Wasowicz, W. (2013) Genetic polymorphisms in matrix metalloproteinases (MMPs) and tissue inhibitors of MPs (TIMPs), and bladder cancer susceptibility. *BJU Int.*, **112**, 1207-1214.
- 402 Beeghly-Fadiel, A., Lu, W., Long, J.R., Shu, X.O., Zheng, Y., Cai, Q., Gao, Y.T. and Zheng, W. (2009) Matrix metalloproteinase-2 polymorphisms and breast cancer susceptibility. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 1770-1776.
- 403 Puerta, D.T. and Cohen, S.M. (2004) A bioinorganic perspective on matrix metalloproteinase inhibition. *Curr Top Med Chem*, **4**, 1551-1573.
- 404 Ingvarsen, S., Porse, A., Erpicum, C., Maertens, L., Jurgensen, H.J., Madsen, D.H., Melander, M.C., Gardsvoll, H., Hoyer-Hansen, G., Noel, A. *et al.* (2013) Targeting a single function of the multifunctional matrix metalloprotease MT1-MMP: impact on lymphangiogenesis. *J Biol Chem*, **288**, 10195-10204.
- 405 Lyu, Y., Xiao, Q., Yin, L., Yang, L. and He, W. (2019) Potent delivery of an MMP inhibitor to the tumor microenvironment with thermosensitive liposomes for the suppression of metastasis and angiogenesis. *Signal Transduct Target Ther*, **4**, 26.



APPENDIX



Vinay J
NISER, Bhubaneswar

8.2 APPENDIX I: VECTORS

Vector name	Description	Bacterial resistance	Selection Marker	Source
pLKO.1 TRC Cloning Vector	Empty vector	Ampicillin	Puromycin	Sigma
pLKO.1_Scrambled	Empty vector	Ampicillin	Puromycin	Sigma
pcDNA 3.1	Empty vector	Ampicillin	Neomycin	Invitrogen
pLPCX2-MT1-MMP-eGFP	MMP14 expression vector	Ampicillin	Puromycin	Addgene
pmiRA1-MYB	MYB expression vector	Ampicillin	GFP	Addgene
FUW-Sox10	SOX10 expression vector	Ampicillin	doxycycline	Addgene
pGL4.23[luc2/minP]	Encodes luc2 gene minimal promoter	Ampicillin	NA	Addgene
pGL4.74[hRluc/TK]	Encodes hRluc gene with HSV-TK promoter	Ampicillin	NA	Addgene
pSpCas9(BB)-2A-Puro (PX459)	CRISPR-Cas9 System	Ampicillin	Puromycin	Addgene
pLKO.1 TRC_shSOX10	SOX10 knockdown vector (TRCN0000018985)	Ampicillin	Puromycin	Cloned in Lab
pLKO.1 TRC_shMYB	SOX10 knockdown vector (TRCN0000288659)	Ampicillin	Puromycin	Cloned in Lab

8.3 APPENDIX II: shRNAs and sgRNAs

shRNAs		
S. No.	Oligonucleotide name	Primer Sequence 5' to 3'
1	MYBshRNA_F (TRCN0000288659)	CCGGGCTCCTAATGTCAACCGAGAACTC GAGTTCTCGGTTGACATTAGGAGCTTTT TG
2	MYBshRNA_R (TRCN0000288659)	AATTCAAAAAGCTCCTAATGTCAACCGA GAACTCGAGTTCTCGGTTGACATTAGGA GC
3	SOX10shRNA_F (TRCN0000018985)	CCGGCCACGAGGTAATGTCCAACATCTC GAGATGTTGGACATTACCTCGTGGTTTT TG
4	SOX10shRNA_F (TRCN0000018985)	AATTCAAAAACCACGAGGTAATGTCCA ACATCTCGAGATGTTGGACATTACCTCG TGG
sgRNAs		
S. No.	Oligonucleotide name	Primer Sequence 5' to 3'

1	M14sgRNA set1 sense	CACCGCAGCCCCCTGCTGTCCATCG
2	M14sgRNA set1 antisense	AAACCGATGGACAGCAGGGGGCTGC
3	M14sgRNA set2 sense	CACCGTCCTTTCCTGGTTGGGGACG
4	M14sgRNA set2 antisense	AAACCGTCCCAACCAGGAAAGGAC

8.4 APPENDIX III: ANTIBODIES

Antibody name	Dilution	Cat. No.	Brand	Host
MMP14	WB: 1:2000	Ab51074	Abcam	Rabbit
MMP2	WB: 1:2000	AF0577	Affinity Bioscience	Rabbit
MMP7	WB: 1:3000	AF0218	Affinity Bioscience	Rabbit
MYB	WB: 1:1000	12319	CST	Rabbit
SOX10	WB: 1:1000	89356	CST	Rabbit
Rabbit IgG	ChIP: 1:5000	11-212	AbGenex	Rabbit
GAPDH	WB: 1:2000	10-10011	AbGenex	Mouse
Sec-Anti-Rabbit	WB: 1:80000	Ampicillin	GFP	Addgene
Sec-Anti- Mouse	WB: 1:20000	Ampicillin	doxycycline	Addgene

8.5 APPENDIX IV: PRIMER SEQUENCES

Genotyping primers			
S. No.	Oligo Name	Primer Sequence 5' to 3'	Annealing temperature
1	MMP2_F_SET1	GGAGTTCATCACAGCTTA	54°C
2	MMP2_R_SET1	GCCTCGTATAGTGCGAGATG	
3	MMP2_F_SET2	CCCAAGCCGCAGAGACTTTT	54°C
4	MMP2_R_SET2	GCCTGACTTCAGCCCCTAAAC	
5	MMP7_F_SET1	CCTGAATGATACCTATGAGAGCA GT	54°C
6	MMP7_R_SET1	CATAGCTGCCGTCCAGAGAC	
7	MMP7_F_SET2	CACCCAATTTGTGGCTTGTGTG	53°C
8	MMP7_R_SET2	CATGGTAATTGAGCACTGTGAGC	
9	MMP14_F_SET1	CAGAGGAATCAAGCCACTCAGA	54°C
10	MMP14_F_SET1	TCCTCTCCGAATAGAGGCTGT	
11	MMP14_F_SET2	GCTGACTGGCTTTGTGCTTAAAT	53°C

12	MMP14_R_SET2	CAAAGTTCCCGTCACAGATGTTG	
Common vector sequencing primers			
1	RV Forword primer	TAGCAAATAGGCTGTCCCC	
2	RV reverse primer	TTACCAACAGTACCGGATTG	
3	hU6 F	GAGGGCCTATTTCCCATGATT	
4	LKO.1 Reverse primer	AAAGTGGATCTCTGCTGTCC	
Oligonucleotides used in luciferase assay constructs			
1	Luc_rs30MC_rs49MG_F	CATCCGGTACCGAGGTGTTTTTTTCTTTTTTCCTTC CAGTTCTTGGTTGTAACCTCGAGGGTAT	
2	Luc_rs30MC_rs49MG_R	ATACCCTCGAGTTACAACCAAGAACTGGAAGGAA AAAAGAAAAAACACCTCGGTACCGGATG	
3	Luc_rs30WtT_rs49WtT_F	CATCCGGTACCGAGGTGTTTTTTTTTTTTTTCCTTC CATTTCTTGGTTGTAACCTCGAGGGTAT	
4	Luc_rs30WtT_rs49WtT_R	ATACCCTCGAGTTACAACCAAGAAATGGAAGGAA AAAAAAAAAAAAACACCTCGGTACCGGATG	
5	Luc_rs30MC_rs49WtT_F	CATCCGGTACCGAGGTGTTTTTTTCTTTTTTCCTTC CATTTCTTGGTTGTAACCTCGAGGGTAT	
6	Luc_rs30MC_rs49WtT_R	ATACCCTCGAGTTACAACCAAGAAATGGAAGGAA AAAAGAAAAAACACCTCGGTACCGGATG	
7	Luc_rs30WtT_rs49MG_F	CATCCGGTACCGAGGTGTTTTTTTTTTTTTTCCTTC CAGTTCTTGGTTGTAACCTCGAGGGTAT	
8	Luc_rs30WtT_rs49MG_R	ATACCCTCGAGTTACAACCAAGAACTGGAAGGAA AAAAAAAAAAAAACACCTCGGTACCGGATG	
9	LucM7_rs18WtG_F	CATCCCTCGAGGTCAGAGTTTGACATGTGATAAG GTGCACCAGGTACCGGTAT	
10	LucM7_rs18WtG_R	ATACCGGTACCTGGTGCACCTTATCACATGTCAA ACTCTGACCTCGAGGGATG	
11	LucM7_rs18MA_F	CATCCCTCGAGGTCAGAGTTTGACATATGATAAG GTGCACCAGGTACCGGTAT	
12	LucM7_rs18MA_R	ATACCGGTACCTGGTGCACCTTATCATATGTCAA ACTCTGACCTCGAGGGATG	
13	LucM7_rs71WtA_F	CATCCCTCGAGGTACGGTCACAGTTTAACTAGAG TAATTGGTGGTACCGGTAT	
14	LucM7_rs71WtA_R	ATACCGGTACCACCAATTACTCTAGTTAACTGT GACCGTACCTCGAGGGATG	
15	LucM7_rs71MC_F	CATCCCTCGAGGTACGGTCACAGTTTCACTAGAG TAATTGGTGGTACCGGTAT	
16	LucM7_rs71MC_R	ATACCGGTACCACCAATTACTCTAGTGAACTGT GACCGTACCTCGAGGGATG	
17	LucM2Wt_Set3_F	CATCCGGTACCACCAGCACTCCACCTCTTTAGCT CTTCCTCGAGGGTAT	
18	LucM2Wt_Set3_R	ATACCCTCGAGGAAGAGCTAAAGAGGTGGAGTG CTGGGTGGTACCGGATG	
19	LucM2Mut_Set3_F	CATCCGGTACCACCAGCACTCTACCTCTTTAGCT CTTCCTCGAGGGTAT	
20	LucM2Mut_Set3_R	ATACCCTCGAGGAAGAGCTAAAGAGGTAGAGTG CTGGGTGGTACCGGATG	
Probes used in electrophoretic mobility shift assay			

1	EMSArs30MC49MGF5 5'Ĕ	AGGTGTTTTTTTTCTTTTTTCCTTCCAGTTCTTGGTT GTA
2	EMSArs30MC49MGR 5'Ĕ	TACAACCAAGAAGCTGGAAGGAAAAAAGAAAAAA ACACCT
3	EMSArs30MC49MGF	AGGTGTTTTTTTTCTTTTTTCCTTCCAGTTCTTGGTT GTA
4	EMSArs30MC49MGR	TACAACCAAGAAGCTGGAAGGAAAAAAGAAAAAA ACACCT
5	EMSArs30WtT49WtTF 5'Ĕ	AGGTGTTTTTTTTTTTTTTTTTCCTTCCATTTCTTGGTT GTA
6	EMSArs30WtT49WtT R5'Ĕ	TACAACCAAGAAATGGAAGGAAAAAAAAAAAAAA ACACCT
7	EMSArs30WtT49WtTF	AGGTGTTTTTTTTTTTTTTTTTCCTTCCATTTCTTGGTT GTA
8	EMSArs30WtT49WtT R	TACAACCAAGAAATGGAAGGAAAAAAAAAAAAAA ACACCT
9	EMSArs30MC49WtTF 5'Ĕ	AGGTGTTTTTTTTCTTTTTTCCTTCCATTTCTTGGTT GTA
10	EMSArs30MC49WtTR 5'Ĕ	TACAACCAAGAAATGGAAGGAAAAAAGAAAAAA ACACCT
11	EMSArs30MC49WtTF	AGGTGTTTTTTTTCTTTTTTCCTTCCATTTCTTGGTT GTA
12	EMSArs30MC49WtTR	TACAACCAAGAAATGGAAGGAAAAAAGAAAAAA ACACCT
13	EMSArs30WtT49MGF 5'Ĕ	AGGTGTTTTTTTTTTTTTTTTTCCTTCCAGTTCTTGGTT GTA
14	EMSArs30WtT49MGR 5'Ĕ	TACAACCAAGAAGCTGGAAGGAAAAAAGAAAAAA ACACCT
15	EMSArs30WtT49MGF	AGGTGTTTTTTTTTTTTTTTTTCCTTCCAGTTCTTGGTT GTA
16	EMSArs30WtT49MGR	TACAACCAAGAAGCTGGAAGGAAAAAAGAAAAAA ACACCT
ChIP-qPCR primers		
1	MMP14 ChIP primer_F	CCTGCACCACAAAAGGCAA
2	MMP14 ChIP primer_R	GGGACGTGGTTGTTTTAGCC

8.6 APPENDIX V: KIT PROTOCOLS/ ASSAY PROCEDURES

Agarose gel extraction of DNA using QIA-quick PCR & Gel Clean-up Kit

1. The DNA fragment was carefully removed from the agarose gel using a sterile scalpel.
2. The gel slice was weighed in a colorless tube, and 3 volumes of Buffer QG were added to 1 volume of gel (approximately 100 mg gel ~100 μ l).

3. The tube was incubated at 50°C for 10 minutes while being vortexed every 2-3 minutes to aid gel dissolution.
4. Isopropanol (1 gel volume) was added to the sample and mixed.
5. A QIAquick spin column was placed in a provided 2 ml collection tube. The sample was applied to the QIAquick column to bind DNA, followed by centrifugation for 1 minute. The flow-through was discarded, and the QIAquick column was returned to the same tube.
6. The flow-through was discarded, and the QIAquick column was placed back into the same tube.
7. 750 µl of Buffer PE was added to the QIAquick column, and it was centrifuged for 1 minute. The flow-through was discarded, and the QIAquick column was placed back into the same tube.
8. The QIAquick column was placed in the provided 2 ml collection tube and centrifuged for 1 minute to remove any residual wash buffer.
9. The QIAquick column was transferred to a clean 1.5 ml microcentrifuge tube.
10. To elute the DNA, 50 µl of Buffer EB was added to the center of the QIAquick membrane, and the column was centrifuged for 1 minute.

Agarose gel DNA extraction using NucleoSpin Gel and PCR Clean-up kit

1. To extract a DNA fragment from an agarose gel, use a clean scalpel to excise the fragment and remove any excess agarose.
2. Weigh the gel slice (less than 2% agarose) in a colorless microcentrifuge tube and add volumes of Buffer NTI to achieve a 1:1 ratio of gel volume (approximately 100 mg gel ~100 µl).

3. Incubate the tube at 50 °C for 5-10 minutes. Briefly vortex the sample every 2-3 minutes until the gel slice is completely dissolved.
4. To bind the DNA, place a NucleoSpin Gel and PCR Clean-up Column in a Collection Tube (2 mL) and load up to 700 µL of the sample. Centrifuge for 30 seconds at 11,000 x g. Discard the flow-through and return the column to the collection tube. Repeat the centrifugation step if necessary to load the remaining sample.
5. Wash the silica membrane by adding 700 µL of Buffer NT3 to the NucleoSpin Gel and PCR Clean-up Column. Centrifuge for 30 seconds at 11,000 x g. Discard the flow-through and return the column to the collection tube.
6. Dry the silica membrane by centrifuging for 1 minute at 11,000 x g to remove Buffer NT3 completely. Ensure that the spin column does not come into contact with the flow-through while removing it from the centrifuge and collection tube.
7. Elute the DNA by placing the NucleoSpin Gel and PCR Clean-up Column into a new 1.5 mL microcentrifuge tube. Add 15-30 µL of Buffer NE and incubate at room temperature (15-25 °C) for 1 minute. Centrifuge for 1 minute at 11,000 x g.

Extraction of plasmids using the Qiagen Plasmid Mini kit:

1. The bacterial cells were collected and resuspended in 250 µl of Buffer.
2. 250 µl of Buffer P2 was added and the tube was gently inverted 4-6 times to ensure thorough mixing.
3. 350 µl of Buffer N3 was added and the tube was inverted 4-6 times for proper mixing.
4. The tube was centrifuged at 13,000 rpm (~17,900 x g) for 10 minutes using a table-top microcentrifuge.

5. A compact white pellet was formed, and 800 μ l of the supernatant from step 4 was transferred to the QIAprep 2.0 spin column using a pipette.
6. The column was centrifuged for 30-60 seconds, and the resulting flow-through was discarded.
7. The QIAprep 2.0 spin column was washed by adding 0.75 ml of Buffer PE and centrifuging for 30-60 seconds.
8. The flow-through was discarded, and an additional centrifugation at full speed for 1 minute was performed to remove residual wash buffer.
9. The QIAprep 2.0 column was placed in a clean 1.5 ml microcentrifuge tube. DNA was eluted by adding 40-50 μ l of Buffer EB (10 mM Tris·Cl, pH 8.5) to the center of each QIAprep 2.0 spin column, followed by centrifugation for one minute.

Extraction of plasmids using the Qiagen Plasmid Midi kit:

1. The bacterial culture (100 ml) was grown overnight and harvested by centrifugation at 6000 x g for 15 minutes at 4°C.
2. The bacterial pellet was resuspended in 4 ml of Buffer P1.
3. 4 ml of Buffer P2 was added, and the mixture was thoroughly mixed by inverting 4–6 times. It was then incubated at room temperature for 5 minutes.
4. 4 ml of prechilled Buffer P3 was added and mixed thoroughly by inverting 4–6 times. The mixture was then incubated on ice for 15 minutes.
5. Centrifugation was performed at $\geq 20,000$ x g for 30 minutes at 4°C. The supernatant was then recentrifuged at $\geq 20,000$ x g for 15 minutes at 4°C.
6. QIAGEN-tip 100 was equilibrated by applying 4 ml of Buffer QBT.

7. The supernatant was poured into the QIAGEN-tip and allowed to enter the resin by gravity flow.
8. The QIAGEN-tip was washed twice with 10 ml of QC buffer.
9. DNA was eluted with 5 ml of Buffer QF into a 15 ml tube.
10. To precipitate the DNA, 3.5 ml of room-temperature isopropanol was added to the eluted DNA and mixed. Centrifugation was performed at $\geq 15,000 \times g$ for 30 minutes at 4°C. The resulting supernatant was carefully decanted.
11. The pellet was washed twice with 70% ethanol and centrifuged at $\geq 15,000 \times g$ for 10 minutes. The supernatant was carefully decanted.
12. The final pellet was air-dried for 5–10 minutes, and the DNA was then redissolved in a suitable volume of TE buffer.

APPENDIX VI: BUFFERS
SDS-PAGE and Western Blot

Components	Weight/Volume
30% Acrylamide (100 ml)	
Bis-acrylamide	1 g
Acrylamide	29 g
dH ₂ O	to 100 ml
0.5 M Tris-Cl, pH 6.8 (100 ml)	
Tris-base	6 g
pH 8.8 with 6 N HCl	Qs
dH ₂ O	to 100 ml
1.5 M Tris-Cl, pH-8.8 (100 ml)	
Tris-base	18.15 g
dH ₂ O	80 ml
pH 8.8 with 6 N HCl	Qs
dH ₂ O	to 100 ml
10% (w/v) APS (10 ml)	
Ammonium persulfate	1 g
dH ₂ O	to 10 ml
10% (w/v) SDS (100 ml)	
SDS	10 g
dH ₂ O	to 100 ml
Water-Saturated n-Butanol (55 ml)	
n-Butanol	50 ml
dH ₂ O	5 ml
2× SDS-PAGE (Laemmli, 30 ml)	
50% Glycerol	15 ml
1.0% Bromophenol blue	0.3 ml
10% SDS	6 ml

Agarose gel electrophoresis

Components	Components
	6x DNA loading buffer (100 ml)
Bromophenol blue	Bromophenol blue
0.5 M EDTA, pH 8	0.5 M EDTA, pH 8
Glycerol	Glycerol
dH ₂ O	to 100 ml
	50x TAE (1 L)
Tris-base	Tris-base
dH ₂ O	dH ₂ O
Glacial acetic acid	Glacial acetic acid
0.5 M EDTA, pH 8	0.5 M EDTA, pH 8
dH ₂ O	to 1 L
	Other buffers and solutions
	10x PBS (1 L)
Na ₂ HPO ₄ ·7H ₂ O	40 g
NaCl	80 g
KCL	2 g
KH ₂ PO ₄	2 g
dH ₂ O	to 1 L
	Crystal violet stain (100 ml)
Crystal violet	2 g
Ethyl alcohol	20 ml
Ammonium citrate monohydrate	0.8 ml
dH ₂ O	80 ml
	4% Paraformaldehyde (1 L)
1x PBS	800 ml
Paraformaldehyde powder	40 g
1x PBS	to 1 L