Abstract

Gallbladder cancer (GBC) is an aggressive malignancy originating from the epithelial cells of the mucous membrane of the gallbladder. It stands as the most prevalent cancer within the biliary tract system and ranks as the fifth most common form of gastrointestinal cancer. This cancer is often diagnosed at advanced stages due to its asymptomatic nature in the early phases and therefore presents significant challenges in treatment and prognosis. GBC is characterized by rapid progression, a delayed prognosis, and high mortality rates, with an overall 5-year survival rate of merely 5%. The incidence of GBC significantly varies across regions and ethnicities, indicating diverse etiological factors among populations. These regional and ethnic diversity in GBC incidence ratios suggest variations in the etiological factors contributing to GBC across different populations. The presence of gallstones and chronic inflammation is associated with increased GBC risk, where the progression usually follows the metaplasia-dysplasia-carcinoma pathway sequence. Nevertheless, the exact molecular mechanisms driving GBC progression through a gallstone-independent pathway are still unclear.

The frequency of GBC cases is highly prevalent among women in India, particularly in the North and North Eastern (NE) regions. The challenge lies in its asymptomatic nature, leading to late-stage diagnoses and limited therapeutic responses, often resulting in palliative care. The 2020 GLOBOCAN report predicts that by 2025, the GBC burden in India, especially in the NE region, will represent 10% of the global population. Unlike other cancers, there is a scarcity of publicly available OMICs scale data on GBC. Comprehensive transcriptome-level molecular data from the highly affected NE Indian region, encompassing GBC and gallstone disease (GSD) patients, is lacking. Therefore, it is of utmost importance to take the initiative for molecular-level understanding of pathophysiological mechanisms responsible for GBC development. Recent advancements in Next Generation Sequencing (NGS) technology, specifically transcriptomics have emerged as a pivotal tool to identify the gene expression patterns in GBC and unravel crucial molecular signatures associated with GBC pathogenesis. The studies envisaged in this thesis have unraveled complex and interacting networks of regulatory molecules involved in GBC pathogenesis.

This thesis addresses the urgent need for a molecular-level understanding of GBC pathogenesis, particularly in the high-incidence NE region of India. By utilizing cutting-edge NGS technology and integrating coding and noncoding regulatory interactions, the study

uncovers distinct gene expression patterns, biological processes, and pathways in GBC, shedding light on the complex molecular landscape of this aggressive malignancy. The thesis makes significant strides in filling the knowledge gap by generating a unique transcriptome dataset from GBC patients in Assam, revealing novel insights into the systemic aspects of GBC pathogenesis. By identifying key molecular signatures, including mRNAs, lncRNAs, and transcription factors, the work lays the foundation for potential diagnostic and therapeutic biomarkers specific to GBC patients. This research not only contributes to the understanding of GBC at the systems level but also holds promising applications for future medical practices, addressing a critical need in a region where GBC poses a substantial health burden.

The entire work in this thesis has been divided into seven chapters:

Chapter I gives a brief introduction to the pathophysiology, etiological factors, and clinical presentations associated with GBC pathogenesis and development. This chapter also describes the hypothesis and key objectives of the work carried out in this thesis.

Chapter II describes the striking epidemiology and molecular pathogenesis of GBC, reviewing the role of noncoding and transcriptional regulation in GBC. This chapter also gives a brief description of the role of cancer systems biology in unraveling the complex molecular interactions in complex diseases such as cancer. It also reviews the unavailability of GBC-specific biomarkers as the molecular-level understanding of GBC pathogenesis remains unclear.

Chapter III provides the materials and methodologies employed to identify potential molecular signatures associated with GBC pathogenesis.

Chapter IV explores the GBC transcriptome for the identification of potential genes and pathological pathways associated with the pathogenesis of GBC. Employing a comprehensive systems-level approach, this chapter unravels gene expression patterns, biological functions, and pathways in GBC. The analysis encompasses both publicly available and in-house generated GBC transcriptomic datasets, with the goal of identifying differentially expressed genes and biological pathways associated with GBC development. The findings revealed that the gallstones-associated (GBC+GS) and gallstones-independent (GBC) show distinct

differential gene expression patterns and biological processes and pathways. This implies that GBC+GS and GBC carcinogenesis progress through distinct pathological pathways.

Chapter V focused on examining the differential lncRNA expression to understand the lncRNA-mediated regulation in GBC pathogenesis and development. This study includes the identification of novel and annotated lncRNA from in-house generated GBC transcriptomic datasets through an integrative network biology-based approach. This chapter presents the identification of key lncRNA signatures and their regulatory components involved in GBC which could provide further avenues to improve the diagnosis and therapeutic strategies of GBC patients.

Chapter VI illustrates the rewiring patterns of transcriptional regulatory networks in GBC pathogenesis and development. This chapter identified crucial regulatory transcription factors associated with GBC development and pathogenesis.

Chapter VII presents the overall outcome and conclusion, offering a concise overview of the future prospects associated with the study findings.