scientific reports

Check for updates

OPEN

Identification of differentially expressed genes and associated immune cell types in South African gallbladder cancer patients

John Devar[1,2], Pavan Baichan[1], Phelelani T. Mpangase[3], Nnenna Elebo[1], Geoffrey Patrick Candy[1], Martin Smith[1] & Ekene Emmanuel Nweke[1,4]✉

Gallbladder cancer (GBC) is a highly aggressive malignancy with limited therapeutic options, particularly in underrepresented populations, including South Africa. Understanding the molecular landscape of GBC may provide novel insights into its pathogenesis and potential therapeutic targets. Molecular changes are known to be associated with GBC, however, there is a paucity of this information especially in African populations. Furthermore, within the tumour microenvironment, different immune cells contribute to GBC progression. We investigated gene expression patterns in GBC tumours and their association with different immune cells in a cohort of South African patients. RNA sequencing was conducted on 2 normal and 8 gallbladder cancer tissues from South African patients to identify differentially expressed genes. Bioinformatics tools were used for pathway analysis, while immune cell quantification was performed using the quanTIseq software and presented as median [IQR]. Verification studies were further carried out using real-time PCR on an independent cohort comprising 7 gallstone samples and 26 gallbladder tumour samples. A total of 65 genes were found to be significantly differentially expressed between the gallbladder tumours and gallstone controls. We also identified 37 upregulated and 28 downregulated genes in this cohort. Among the most upregulated genes, *MUC16* was confirmed to be significantly overexpressed in tumours. Normal tissues exhibited a significantly higher proportion of dysregulated genes associated with B cells (17.132 [14.866–18.483], $p < 0.0001$) and M1 macrophages (18.943 [1.097–36.790], $p < 0.0001$) compared to tumours. In contrast, tumours showed a greater association with dysregulated genes linked to regulatory T cells (Tregs) (14.373 [9.696–20.162]) relative to normal tissues. Pathway analysis further revealed the upregulation of defective GALNT12, defective GALNT3, defective C1GALT1C1 and termination of O-glycan biosynthesis, highlighting key mechanisms potentially involved in tumour progression. The study has shown the dysregulation of key genes in South African gallbladder cancer patients. Specifically, *MUC16* was verified to be significantly elevated in tumour samples. Furthermore, the association of these dysregulated genes with key immune cells in this patient group may further highlight their roles in dysfunctional immune processes linked with tumourigenesis.

Keywords Gallbladder cancer, MUC16, Transcriptomics, Immune response, Dysregulated pathways

Gallbladder cancer (GBC) is an aggressive malignancy with poor prognosis and limited therapeutic options[1]. The disease burden varies globally, with a particularly high incidence in certain regions, including parts of Asia and South America[2,3]. However, GBC's epidemiology, molecular characteristics, and immune landscape in African populations, particularly in South Africa, remain largely understudied[4]. One study found a higher incidence of GBC in South Africa than was reported[5]. Gallstone disease, also known as cholelithiasis, is one of the most strongly associated risk factors for GBC, with 70–90% having a history of cholelithiasis[6]. Understanding the

[1]Department of Surgery, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg 2193, South Africa. [2]Hepatopancreatobiliary Unit, Department of Surgery, Chris Hani-Baragwanath Academic Hospital, Soweto, Johannesburg 1864, South Africa. [3]Sydney Brenner Institute for Molecular Bioscience, University of the Witwatersrand, Johannesburg, South Africa. [4]Department of Life and Consumer Sciences, College of Agriculture and Environmental Sciences, University of South Africa, Florida, Roodepoort 1709, South Africa. ✉email: ekene.nweke@wits.ac.za

molecular mechanisms and immune architecture of GBC in our population is crucial for developing effective diagnostic and therapeutic strategies[7].

Recent advances in transcriptomic profiling have enabled the identification of differentially expressed genes (DEGs) in various cancers, providing insights into disease pathogenesis and potential biomarkers for prognosis and treatment[8]. Several studies have utilized transcriptomic techniques to distinguish GBC from control samples, such as gallstones, to identify distinct molecular signatures[9–11]. The findings support the biological link between chronic gallstone disease and GBC development as well as key differences in gene expression related to inflammation, immune regulation, and tumour progression[6,12]. Moreover, they highlight potential biomarkers that could help distinguish malignant from benign gallbladder conditions[11].

Studies have also implicated the mucin (MUC) family in cancer progression due to its involvement in cell signaling, immune evasion, and metastasis[13]. For instance, alterations in *MUC* have been associated with aggressive tumour phenotypes and poor clinical outcomes in various malignancies, including GBC[14].

The tumour microenvironment (TME), particularly the composition of immune cell types, plays a critical role in cancer progression and response to therapy[15]. GBC is characterized by an imbalance between immune activation and suppression, contributing to tumour progression and therapy resistance. The immune landscape of GBC is complex, with significant infiltration of immune cells such as tumour-associated macrophages, T cells, and dendritic cells, which can either promote or inhibit tumour growth[16]. Key immune pathways, including the PD-1/PD-L1 axis, JAK-STAT signaling, and TGF-β signaling, are involved in modulating immune responses within the GBC microenvironment, affecting tumour progression and therapeutic resistance[17]. The immune microenvironment in investigating the interplay between gene expression changes and immune cell infiltration in South African GBC patients can offer novel perspectives on disease biology and potential therapeutic targets.

Despite the global efforts to characterize GBC at the molecular level, African populations remain underrepresented in these studies. Given the genetic diversity of different populations and potential environmental and lifestyle factors influencing cancer development, it is imperative to conduct population-specific studies. Therefore, this study aimed to identify DEGs and associated immune cell types in South African gallbladder cancer patients by conducting transcriptomic analyses.

Results

Differentially expressed genes in gallbladder tumours using RNA sequencing

RNA-seq was performed on 2 normal Gallbladder tissues and 12 Gallbladder tumours. However, 4 of the 12 tumours failed quality control. The characteristics of the RNA-seq data used for differential gene expression analysis were assessed using principal component analysis (PCA) plots (Figure S1). The PCA plot shows a clear distinction, or clustering, of cases from the controls. The first principal component can explain 61% of the variance, and the second principal component can explain 10% of the variance in the RNA-seq data used. Using an FDR of 0.05 (5%) and an absolute $\log_2$FC threshold of 2, 65 genes were found to be significantly differentially expressed between the cases and normal tissue controls (Table S1). Of the 65 genes, 37 were up-regulated and 28 were down-regulated. Figure 1 shows the heatmap of how the 65 genes cluster due to their expression levels.

Verification of expression of MUC 16 and MUC 19 using real-time PCR

Furthermore, the MUC gene family members were significantly overexpressed in tumours. Due to this and their potential clinical relevance, we sought to verify their expression in additional tissue samples. Seven gallstones (GS) and 26 gallbladder tumour (GBT) patients were further recruited to the study (Table S2). Although not statistically significant, real-time PCR demonstrated that both *MUC16* (831.02 ± 2260.15) and *MUC19* (10.64 ± 30.80) genes have higher values in GBT than in GS patients: *MUC16* (2.45 ± 2.42) and *MUC19* (2.77 ± 3.80) genes. (Fig. 1).

Pathway analysis of dysregulated genes

The dysregulated genes were subjected to pathway analysis using Reactome. Pathways such as defective GALNT12, defective GALNT3, defective C1GALT1C1 and termination of O-glycan biosynthesis were enriched by the upregulated genes while pathways such as Smooth Muscle Contraction, ECM proteoglycans and Platelet degranulation were down-regulated (Table 1).

Quantification of immune cell types

The 10 different types of immune cells the quanTIseq pipeline characterizes were found at varying proportions in the control and tumour samples. Figure 2 summarizes the type of immune cells and their proportions in the RNA-seq samples used in this study. B cells were found at ~2.5% in the control samples ~13–18% in tumour samples. M1 macrophages were the most dominant cell type in the control samples (~17–20%) and less dominant in the tumour samples (< 9%). The M2 macrophages were found to be at similar proportions throughout all the samples (~4%), except for GB02 (a control sample), which had no M2 macrophage signatures identified. Monocytes were only identified in the tumour sample GBT08 at 3.5%. Neutrophils and NK cells were identified in all samples at similar proportions. Non-regulatory $CD4^+$ T cells were not identified in any of the samples, whilst $CD8^+$ T-cells and T_{reg} cells were found in much higher proportions in tumour samples compared to the control samples. DCs were only identified in the tumour samples. M1 macrophages were significantly higher in the normal tumours while Tregs and B.cells were significantly elevated in the GBC tumours as shown in Table 2.

Discussion

This study provides a comprehensive analysis of differentially expressed genes (DEGs) and their associated immune cell profiles in gallbladder cancer (GBC) patients from South Africa, offering novel insights into the

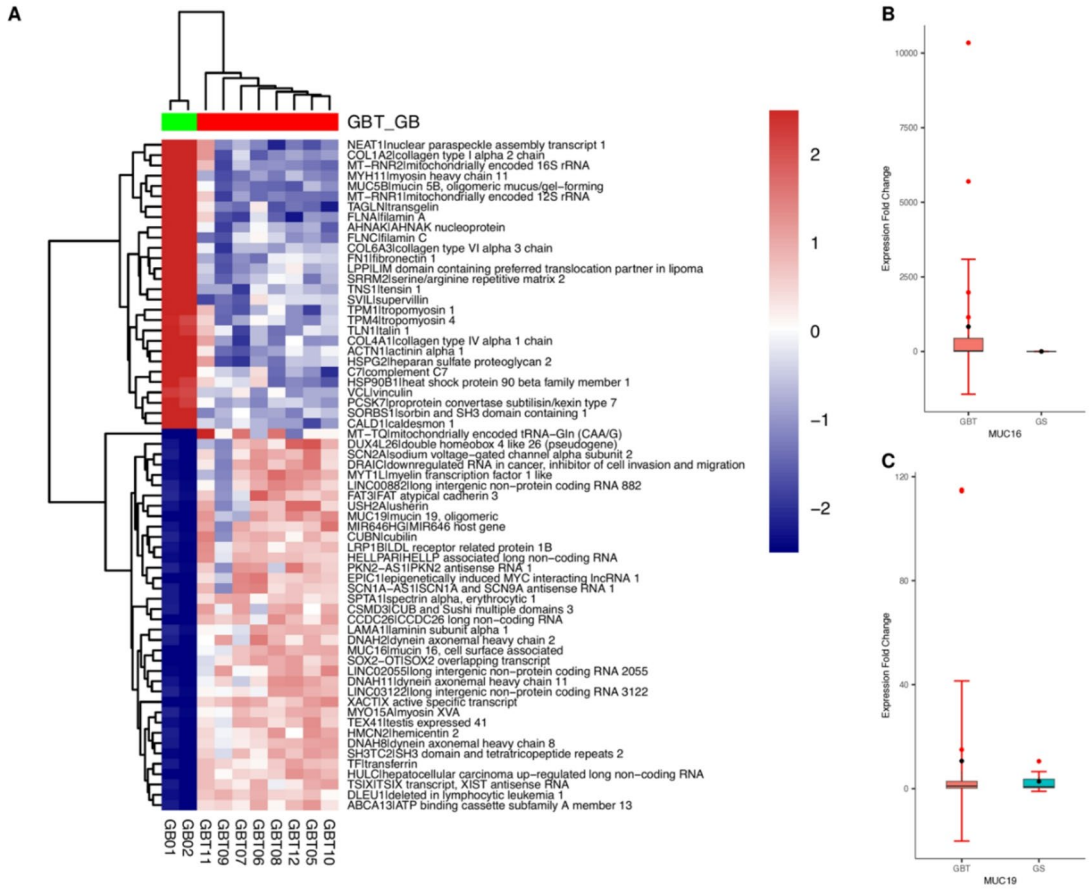

Fig. 1. Gene expression profiling of Gallbladder tumours. **A**: Heatmap showing the clustering and level of expression of the 65 significantly differentially expressed genes between the tumour (GBT05, GBT06, GBT07, GBT08, GBT09, GBT10, GBT11 and GBT12) and control (GB01 and GB02) samples. Most of these genes were downregulated in GB and upregulated in GBT. **B** and **C**: Fold change differences in both *MUC 16* ($p = 0.034$) and *MUC 19* ($p = 0.682$) were elevated in GBT when compared to GS patients, although not significant. GBT: Gallbladder tumours, GB: Normal gallbladder, GS: Gallstones *MUC 16*: Mucin-16, *MUC 19*: Mucin-19.

molecular landscape of this malignancy. Our findings highlight key molecular differences between gallbladder tumours and gallstones (controls), with 65 genes identified as differentially expressed. Among these, 37 genes were found to be upregulated, while 28 were downregulated, underscoring key alterations in gene expression profiles. Among the most notable upregulated genes, members of the mucin family genes, such as *MUC16* and *MUC19*, were further demonstrated to be consistently overexpressed in tumour tissues, suggesting a potential role in GBC pathogenesis. MUC-16, also known as CA-125, has been implicated in tumour progression, immune evasion, and metastasis in various cancers, including pancreatic and ovarian malignancies[18]. It is widely used as a tumour-associated marker for the diagnosis of ovarian cancer by promoting cancer cell migration and tumour metastasis by binding mesothelin to reduce Dickkopf-1(DKK1) expression, as well as activating serum/glucocorticoid-regulated kinase family member 3/forkhead box O3 (SGK3/FOXO3) pathway[19]. In GBC, it has a diagnostic potential for differentiating malignant from gallstone diseases[20]. In combination with other markers such as CA 19 – 9, MUC-16 is used in the diagnosis and prognosis of GBC[21]. Similarly, MUC-19, although less studied, has been associated with mucinous carcinomas[22]. They promote upregulation of α-fetoprotein (AFP) expression in gastric cancer cells and played a critical role in hepatoid adenocarcinoma of the stomach (HAS) development[23], hence may contribute to the aggressive nature of GBC. Their elevated expression in tumours suggests their involvement in gallbladder carcinogenesis and their potential utility as biomarkers or therapeutic targets.

Pathway analysis further elucidated the functional impact of these DEGs, showing a marked upregulation of pathways involved in cell adhesion and intracellular signalling, such as defective GALNT3/GALNT12, C1GALT1C1-related pathways and the termination of O-glycan biosynthesis pathways in tumours. The defective GALNT3 and GALNT12 pathways, which are associated with hereditary tumour syndromes, may indicate disruptions in glycosylation processes that drive oncogenesis[24,25]. Aberrant glycosylation is known to affect cellular interactions, signaling, and immune recognition, potentially contributing to tumour aggressiveness and metastatic potential[26]. In particular, C1GALT1 has been shown to possess dual regulatory roles, including both in tumour promotion and suppression[27]. Hence, it has been suggested as a potential biomarker and therapeutic target. Conversely, downregulated pathways were predominantly linked to extracellular matrix

Upregulated			Downregulated		
Pathway name	***p*-value**	**Proteins enriched**	**Pathway name**	***p*-value**	**Proteins enriched**
Defective GALNT12 causes CRCS1	2.6×10^{-2}	MUC16; MUC19	Smooth Muscle Contraction	2.11×10^{-5}	CALD1;TPM1;MYH11;SORBS1;TLN1
Defective GALNT3 causes HFTC	2.6×10^{-2}	MUC16; MUC19	ECM proteoglycans	4.50×10^{-6}	COL1A2;COL4A1;FN1;COL6A3;HSPG2
Defective C1GALT1C1 causes TNPS	2.6×10^{-2}	MUC16; MUC19	Non-integrin membrane-ECM interactions	4.50×10^{-6}	COL1A2;COL4A1;ACTN1;FN1;HSPG2
Termination of O-glycan biosynthesis	3.3×10^{-2}	MUC16; MUC19	Integrin cell surface interactions	4.50×10^{-6}	COL1A2;COL4A1;FN1;COL6A3;HSPG2
			Degradation of the extracellular matrix	6.27×10^{-5}	COL1A2;COL4A1;FN1;COL6A3;HSPG2
			Collagen chain trimerization	1.49×10^{-4}	COL1A2; COL4A1; COL6A3
			Cell-extracellular matrix interactions	3.86×10^{-4}	ACTN1; FLNA; FLNC
			Anchoring fibril formation	2.20×10^{-4}	COL1A2; COL4A1
			Extracellular matrix organization	4.26×10^{-4}	COL1A2; COL4A1; ACTN1; FN1; COL6A3; HSPG2
			Platelet degranulation	4.26×10^{-4}	TF; ACTN1; FN1; FLNA; TLN1

Table 1. Top dysregulated pathways in gallbladder Tumours. GALNT12: Polypeptide N-acetylgalactosaminyltransferase 12, CRCS1; colorectal cancer susceptibility 1, GALNT3: Polypeptide N-acetylgalactosaminyltransferase 3, HFTC; hyperphosphatemic familial tumoural calcinosis, C1GALT1C1: Glycoprotein-N-acetylgalactosamine 3-beta galactosyltransferase 1, TNPS: Tn polyagglutination syndrome, MUC16; mucin 16, MUC19: mucin 19, CALD1: calmodulin- and actin-binding protein 1, TPM1; Tropomyosin 1, MYH11; myosin heavy chain 11, SORBS1; sorbin and SH3 domain containing 1, TLN1; Talin 1, COL1A2; collagen type I alpha 2 chain, COL4A1; collagen type IV alpha 1 chain, COL6A3; collagen type VI alpha 3 chain, FN1; fibronectin 1, HSPG2; heparan sulfate proteoglycan 2, ACTN1;FN1; actinin alpha 1, FLNA; filamin A, FLNC; filamin C, TF; transferrin, ECM: Extracellular matrix.

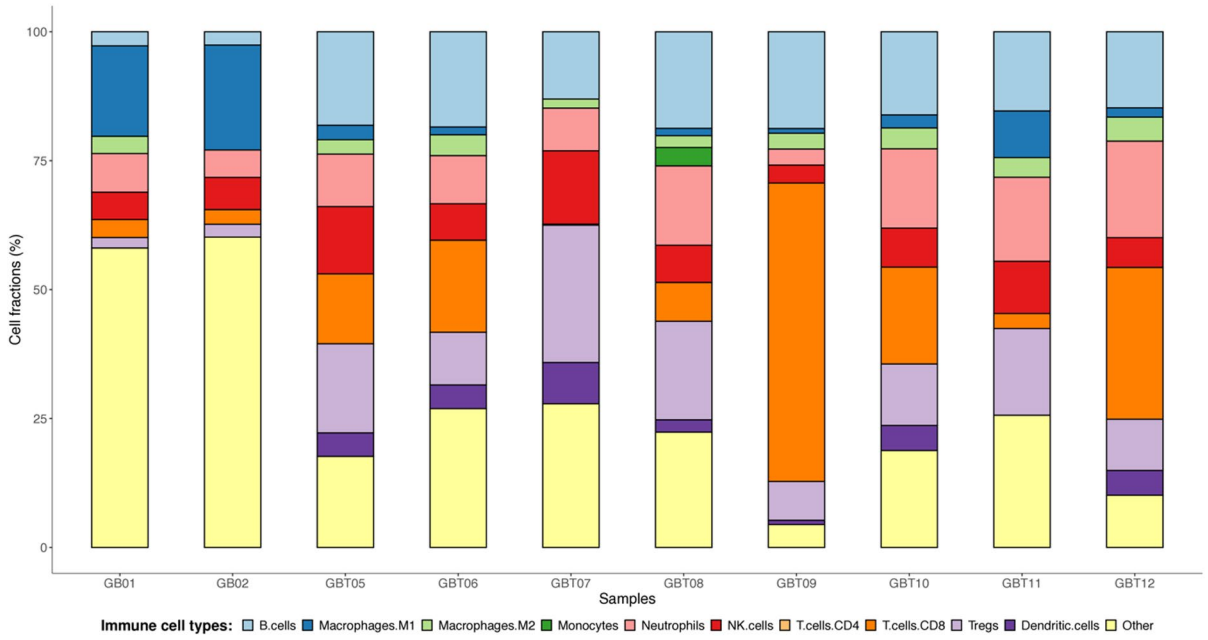

Fig. 2. Quantification of Immune Cell Types in Gallbladder Tumour. Proportions of the 10 different immune cell types relevant for tumour (GBT05, GBT06, GBT07, GBT08, GBT09, GBT10, GBT11 and GBT12) and control (GB01 and GB02) samples were identified in the RNA-seq data using quanTIseq, the pipeline. M1 macrophages were most common in the control samples, while B.cells, CD8 and Tregs were most common in the tumour samples.

(ECM) organisation and interactions, including smooth muscle contraction, ECM proteoglycans, integrin cell surface interactions, and degradation of the extracellular matrix. These findings suggest that tumour progression in GBC may be accompanied by a loss of normal ECM integrity, potentially facilitating tumour invasion and metastasis[28]. Reduced ECM organisation is a recognised feature of tumour progression, as it facilitates cellular detachment, invasion, and metastatic dissemination[28,29]. Collectively, these findings reinforce the molecular

Immune Cell	Mean [SD]	Median [IQR]	*P*-value
B Cells			< **0.0001**
Normal	2.667 [0.123]	2.667 [1.560–3.775]	
Tumour	16.674 [2.163]	17.132 [14.866 18.483]	
Macrophages (M1)			< **0.0001**
Normal	18.943 [1.986]	18.943 [1.097 36.790]	
Tumour	2.505 [2.780]	1.669 [0.180 4.829]	
Macrophages (M2)			0.133
Normal	1.664 [2.354]	1.664 [−19.484 22.813]	
Tumour	3.311 [0.992]	3.451 [2.481 4.141]	
Monocytes			0.645
Normal	0.000 [0.000]	0.000 [0.000 0,000]	
Tumour	0.448 [1.267]	0.000 [−0.611 1.507]	
Neutrophils			0.181
Normal	6.410 [1.549]	6.410 [−7.511 20.332]	
Tumour	12.081 [5.204]	12.774 [7.731 16.432]	
NK Cells			0.333
Normal	5.775 [0.674]	5.775 [−0.284 11.834]	
Tumour	8.562 [3.647]	7.396 [5.514 11.611]	
T-Cells CD4			-
Normal	0.000 [0.000]	0.000 [0.000 0.000]	
Tumour	0.000 [0.000]	0.000 [0.000 0.000]	
T-Cells CD8			0.294
Normal	3.166 [0.461]	3.166 [−0.976 7.308]	
Tumour	18.514 [18.486]	15.696 [3.060 33.968]	
Tregs			**0.026**
Normal	2.257 [0.327]	2.257 [−0.681 5.195]	
Tumour	14.929 [6.260]	14.373 [9.696 20.162]	
Dendritic Cells			0.084
Normal	0.000 [0.000]	0.000 [0.000 0.000]	
Tumour	3.751 [2.568]	4.575 1.605 5.898]	
Other			< **0.0001**
Normal	59.117 [1.500]	59.117 [45.642 72.592]	
Tumour	19.224 [8.357]	20.578 [12.238 26.210]	

Table 2. Comparison of percentage of immune cells between gallbladder cancer tumours and Controls. SD: Standard Deviation, IQR: Interquartile range.

complexity of GBC patients. The upregulation of cell adhesion and key intracellular signalling pathways suggests a critical role for tumour-stroma interactions in disease progression, while the downregulation of ECM-related pathways highlights key structural alterations that facilitate tumour invasion.

The immune microenvironment plays a crucial role in tumour progression and response to therapy. Our analysis of immune cell associations with dysregulated genes revealed a striking difference between normal and tumour tissues. Normal tissues exhibited a significantly higher percentage of gene associations with B cells and M1 macrophages ($P < 0.0001$). The presence of M1 macrophages suggests an active pro-inflammatory response in normal tissue, potentially acting as a barrier against tumorigenesis[30]. In contrast, tumours demonstrated an increased association with Tregs, known for their immunosuppressive functions[31]. Tregs are key mediators of immune evasion in cancers, inhibiting cytotoxic T-cell responses and promoting tumour immune tolerance[32]. Furthermore, the proportion of CD8 cells was shown to be higher in tumour samples compared to the controls. Recent studies showed that elevated levels of both CD4 and CD8 infiltrating lymphocytes suggest a favourable prognosis after surgical removal of GBC[33]. This shift in immune cell dynamics may provide insights into the immunosuppressive tumour microenvironment in GBC and highlight potential therapeutic targets.

Conclusion

This study highlights key molecular differences between gallbladder tumours and gallstone controls, identifying significant alterations in gene expression, immune response, and glycosylation pathways. The significant overexpression of *MUC16* in tumours corroborates its potential role as a biomarker. In addition, the enrichment of pathways associated with defective glycosylation suggests potential mechanisms driving tumour progression. Additionally, the distinct immune landscape observed in tumours compared to normal tissues emphasizes the role of immune evasion in gallbladder carcinogenesis. These findings contribute to a better understanding of

gallbladder cancer biology in our patient population and provide a foundation for future research aimed at identifying novel biomarkers.

Limitations

This study offers novel insights into the transcriptomic landscape of GBC in South African patients, highlighting significant alterations in mucin expression and glycosylation pathways. However, functional experiments were not conducted to directly establish causal roles for *MUC16* and *MUC19* in tumour progression. Future work integrating in vitro and in vivo validation models will be essential to confirm the mechanistic contributions of these pathways. Additionally, to confirm the expression of the genes in the tumour, techniques such as Immunohistochemistry would need to be performed.

Methods

Ethics statement

The Human Research Ethics Committee of the University of Witwatersrand, Johannesburg, South Africa, approved this study unconditionally (M160640). Written consent was obtained from all patients before sample collection. The research was conducted in line with the Declaration of Helsinki and other relevant guidelines and regulations.

Sample collection

A total of 43 tissues were collected for this study, consisting of 2 normal tissues (GB), 7 gallstones (GS) and 34 (advanced stage) gallbladder tumours (GBT). The two normal samples were obtained from samples scheduled for cadaveric liver transplantation. Tumour biopsies were collected by FNA (Fine Needle Aspiration) or Tru-cut biopsy in advanced disease or by post-surgical histopathology in early, resectable disease. All tumours obtained were histologically confirmed as cancerous. All samples were obtained from patients presenting at both Chris Hani Baragwanath Academic and Wits Donald Gordon Hospitals, Johannesburg, South Africa, between the period of September 2019 to January 2021. To preserve RNA integrity upon retrieval from patients, tissue samples were submerged in 700 µl RNAlater (Cat No: 76104, Qiagen Hilden, Germany), stored at 4 °C and then later transferred to −80 °C freezer until ready to be used.

RNA extraction

Biopsy samples were homogenized using a TissueRuptor (Cat No: 9002755 Qiagen Hilden, Germany). Total RNA was extracted by following the steps described in the RNeasy mini kit (Qiagen Cat No: 74104). Total RNA underwent DNASE I (Cat No: AMPD1-1KT St. Louis, Missouri, United States) digestion to eradicate genomic DNA contamination. The Nanodrop™ 2000 (Cat No: ND-2000 Thermofisher, Waltham, Massachusetts, United States) was used to check for quantity and quality of RNA. A A260/280 greater than 1.8 was observed in all samples.

RNA sequencing

Total RNA was extracted from 2 normal gallbladder tissues (GB) and 8 gallbladder tumours (GBT). The total RNA was initially stored at the − 80 °C freezer. Briefly, samples were processed on the Agilent Bioanalyzer 2100 (Cat No: G2939BA Agilent, Santa Clara, California, United States) to determine the integrity of the RNA (RIN number). Following this, the RNA underwent a ribosomal RNA (rRNA) depletion step using the NEBNext® rRNA Depletion Kit (Human/Mouse/Rat) (Cat No: E6310S NEB Ipswich, Massachusetts, United States) as per the manufacturer's protocol. The resulting rRNA-depleted RNA was converted to double-stranded cDNA, using the Ultra RNA Library Prep kit (NEB Cat No: 7530 L). A DNA chip (Agilent BioAnalyzer) was then used to determine the size distribution of the newly synthesised cDNA.

Following BioAnalyzer analysis, the DNA was enzymatically fragmented using the NEB Ultra II FS Kit (Cat No: 7805 L). Resulting fragments were end-repaired, and an Illumina-specific adapter sequence was ligated to all fragments. The adaptor-ligated DNA then underwent a clean-up using AMPure XP beads. The samples were then individually indexed, and a second size selection step was performed, using AMPure XP beads (Brea, California, United States). The samples were then loaded on the Bioanalyzer to determine the size distribution of the Illumina libraries. The samples were diluted, denatured and finally sequenced on Illumina's MiSeq platform, using a MiSeq v3 (600 cycle) kit, according to the manufacturer's protocol.

Differential gene expression

To perform differential expression analysis, the Nextflow nf-rnaSeqCount pipeline (https://github.com/phelelani/nf-rnaSeqCount) was used to obtain a matrix of raw read counts from the RNA-Seq data. The nf-rnaSeqCount pipeline offers a fast and efficient method for mapping raw RNA-seq reads to a reference genome and quantifying the abundance of genomic features for differential gene expression analysis[34]. The resulting read count matrix (N), where N_{ij} represents the number of reads assigned to gene i in sequencing experiment j[35], can be analyzed using tools such as DESeq2[36] and edgeR[37] for differential expression analysis. In this study, the read count matrix generated by the nf-rnaSeqCount pipeline was analyzed using DESeq2 (version 1.44.0) in R (version 4.4.0) to identify DEGs. A false discovery rate (FDR) of 0.05 (5%) and an absolute $\log_2$ fold-change ($\log_2$FC) threshold of 2 were applied to determine significant DEGs between cases and controls. Gene annotation was conducted using the biomaRt package (version 2.60.1)[38], while gene-set enrichment analysis was performed using the enrichR package (version 3.2)[39] in R.

Quantification of immune cell types

The quanTIseq pipeline[40] was used to quantify the different fractions of immune cell types from the RNA-seq data. quanTIseq is a computational tool that uses a novel deconvolution approach to quantify tumour immune contexture of 10 different immune cell types relevant for cancer immunology in human RNA-seq data. The ten different immune cell types include: B cells, M1 and M2 macrophages, monocytes, neutrophils, natural killer (NK) cells, non-regulatory $CD4^+$ T cells, $CD8^+$ T cells, T_{reg} cells, and myeloid dendritic cells (DC). A quanTIseq Docker container for performing the analysis was downloaded from DockerHub (https://hub.docker.com/r/icbi/quantiseq2) using Singularity, and analysis was performed on the University of the Witwatersrand Computer Cluster.

Quantitative real-time polymerase chain reaction for MUC-family genes

An additional 7 gallstones and 26 Gallbladder tumours were used to verify the expression of *MUC16* and *MUC19*. Before qPCR for gene expression assay, cDNA synthesis was performed using an equal amount of Total RNA across the samples. The Invitrogen SuperScript™ VILO™ cDNA Synthesis Kit (Cat No: 11754-050) was used to perform cDNA synthesis; this method is optimised for generating first-strand cDNA for the use of two-step qPCR. PCR tubes were prepared according to the manufacturer's instructions. The total RNA added per sample was calculated to be 600 ng (each sample RNA volume differed). A no-template control (no RNA) was also prepared. Once the tubes were prepared, they were gently mixed and incubated. Once the cycling was completed, the concentration of the samples was quantified using the Nanodrop ND-2000 spectrophotometer. Thereafter, the cDNA samples were diluted to stock concentrations of 1000 ng/µl. Primer assays for *MUC16* (Hs01065175_m1), *MUC19* (Hs00417048_m1) and *ACTB* (Hs01060665_g1) were used for this study, with *ACTB* being the reference gene. The real-time PCR cycle was 95 °C for 1 min, 95 °C for 15 s and 60 °C for 1 min. The MIQE guidelines[41] were adhered to, and the expression values were obtained delta-delta CT method[42].

Statistical and pathway analysis

The differential expression results were used to construct pathways using gage (version 2.54.0;[43]) and pathview (version 1.44.0;[44]) in R to identify those pathways affected by the up-and down-regulated genes. The differentially expressed genes were mapped to ENTREZ IDs, and their Log_2FC was extracted for the pathway analysis. Pathway analysis was conducted using the Reactome pathway browser v3.7[45]. A Student's t-test was used for expression in tumours vs. stone samples to compare. A p-value of less than 0.05 was considered significant.

Data availability

Sequence data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) repository with the link [GSE308395] (https:/www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE308395).

Received: 6 March 2025; Accepted: 3 October 2025
Published online: 11 November 2025

References

1. Bowah, C. S. et al. Management of advanced gallbladder adenocarcinoma: A case report and review of treatment strategies. *Int. J. Surg. Case Rep.* **127**, 110965 (2025).
2. Izarzugaza, M. I., Fernández, L., Forman, D. & Sierra, M. S. Burden of gallbladder cancer in central and South America. *Suppl. Cancer Cent. S Am.* **44**, S82–S89 (2016).
3. Afzal, A. et al. Epidemiology of gall bladder cancer and its prevalence worldwide: a meta-analysis. *Orphanet J. Rare Dis.* **20**, 143 (2025).
4. Baichan, P. et al. Proteomic analysis identifies dysregulated proteins and associated molecular pathways in a cohort of gallbladder cancer patients of African ancestry. *Clin. Proteom.* **20**, 8 (2023).
5. Khan, Z. A., Khan, M. U. & Brand, M. Gallbladder cancer in africa: A higher than expected rate in a low-risk population. *Surgery* **171**, 855–858 (2022).
6. Schmidt, M. A., Marcano-Bonilla, L. & Roberts, L. R. Gallbladder cancer: epidemiology and genetic risk associations. *Chin. Clin. Oncol. Vol 8 No 4 August 27* Clin. Oncol. Gallbladder Cancer (2019). (2019).
7. Nepal, C. et al. Integrative molecular characterisation of gallbladder cancer reveals micro-environment-associated subtypes. *J. Hepatol.* **74**, 1132–1144 (2021).
8. Samadi, P. et al. An integrative transcriptome analysis reveals potential predictive, prognostic biomarkers and therapeutic targets in colorectal cancer. *BMC Cancer.* **22**, 835 (2022).
9. Yang, S. et al. RNA sequencing revealed the multi-stage transcriptome transformations during the development of gallbladder cancer associated with chronic inflammation. *Res. Sq.* https://doi.org/10.21203/rs.3.rs-1516146/v2 (2022).
10. Yu, H. et al. Potential molecular markers for cholecystic Inflammation-Induced carcinogenesis based on RNA-Seq gene screening. *Comb. Chem. High. Throughput Screen.* **28**, 931–943 (2025).
11. Wang, J. et al. RNA sequencing revealed signals of evolution from gallbladder stone to gallbladder carcinoma. *Front. Oncol.* **10**, 823 (2020).
12. Giraldo, N. A. et al. Comprehensive molecular characterization of gallbladder carcinoma and potential targets for intervention. *Clin. Cancer Res.* **28**, 5359–5367 (2022).
13. Chen, X., Sandrine, I. K., Yang, M., Tu, J. & Yuan, X. MUC1 and MUC16: critical for immune modulation in cancer therapeutics. *Front. Immunol.* **15**, 1356913 (2024).
14. Barreto, S. G., Dutt, A. & Chaudhary, A. A genetic model for gallbladder carcinogenesis and its dissemination. *Ann. Oncol.* **25**, 1086–1097 (2014).
15. Wang, Q. et al. Role of tumor microenvironment in cancer progression and therapeutic strategy. *Cancer Med.* **12**, 11149–11165 (2023).
16. Espinoza, J. A. et al. The inflammatory inception of gallbladder cancer. *Biochim. Biophys. Acta.* **1865**, 245–254 (2016).
17. Wang, X. et al. Key oncogenic signaling pathways affecting tumor-infiltrating lymphocytes infiltration in hepatocellular carcinoma: basic principles and recent advances. *Front. Immunol.* **15**, 1354313 (2024).
18. Aithal, A. et al. MUC16 as a novel target for cancer therapy. *Expert Opin. Ther. Targets.* **22**, 675–686 (2018).

19. Huo, Q. et al. Free CA125 promotes ovarian cancer cell migration and tumor metastasis by binding mesothelin to reduce DKK1 expression and activate the SGK3/FOXO3 pathway. *Int. J. Biol. Sci.* **17**, 574–588 (2021).
20. Chaube, A., Tewari, M., Singh, U. & Shukla, H. CA 125: a potential tumor marker for gallbladder cancer. *J. Surg. Oncol.* **93**, 665–669 (2006).
21. Sinha, S. R., Prakash, P., Singh, R. K. & Sinha, D. K. Assessment of tumor markers CA 19 – 9, CEA, CA 125, and CA 242 for the early diagnosis and prognosis prediction of gallbladder cancer. *World J. Gastrointest. Surg.* **14**, 1272–1284 (2022).
22. Kufe, D. W. Mucins in cancer: function, prognosis and therapy. *Nat. Rev. Cancer*. **9**, 874–885 (2009).
23. Zhu, M. et al. Genomic profiling and the impact of MUC19 mutation in hepatoid adenocarcinoma of the stomach. *Cancer Commun.* **42**, 1032–1035 (2022).
24. Hussain, M. R. M. & Hoessli, D. C. Fang, M. N-acetylgalactosaminyltransferases in cancer. *Oncotarget* **7**, 54067–54081 (2016).
25. Yu, S. et al. GALNT12 promotes fibrosarcoma growth by accelerating YAP1 nuclear localization. *Oncol. Lett.* **26**, 543 (2023).
26. Tian, H. et al. Unraveling the role of C1GALT1 in abnormal glycosylation and colorectal cancer progression. *Front. Oncol.* **14**, 1389713 (2024).
27. Xia, T., Xiang, T. & Xie, H. Update on the role of C1GALT1 in cancer. *Oncol. Lett.* **23**, 97 (2022).
28. Elgundi, Z. et al. Cancer metastasis: the role of the extracellular matrix and the Heparan sulfate proteoglycan Perlecan. *Front. Oncol.* **9**, 1482 (2019).
29. Malik, R., Lelkes, P. I. & Cukierman, E. Biomechanical and biochemical remodeling of stromal extracellular matrix in cancer. *Trends Biotechnol.* **33**, 230–236 (2015).
30. Huang, R., Kang, T. & Chen, S. The role of tumor-associated macrophages in tumor immune evasion. *J. Cancer Res. Clin. Oncol.* **150**, 238 (2024).
31. Qin, D. et al. Targeting tumor-infiltrating Tregs for improved antitumor responses. *Front. Immunol.* **15**, 1325946 (2024).
32. Zhang, A. et al. Regulatory T cells in immune checkpoint Blockade antitumor therapy. *Mol. Cancer*. **23**, 251 (2024).
33. Nakakubo, Y. et al. Clinical significance of immune cell infiltration within gallbladder cancer. *Br. J. Cancer*. **89**, 1736–1742 (2003).
34. Mpangase, P. T., Frost, J., Tikly, M., Ramsay, M. & Hazelhurst, S. nf-rnaSeqCount: A nextflow pipeline for obtaining Raw read counts from RNA-seq data. *Afr. Comput. J.* **33**, 830 (2021).
35. Rapaport, F. et al. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol.* **14**, R95 (2013).
36. Love, M. I., Huber, W. & Anders, S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
37. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
38. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomart. *Nat. Protoc.* **4**, 1184–1191 (2009).
39. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
40. Finotello, F. et al. Molecular and Pharmacological modulators of the tumor immune contexture revealed by Deconvolution of RNA-seq data. *Genome Med.* **11**, 34 (2019).
41. Bustin, S. A. et al. The MIQE guidelines: minimum information for publication of quantitative Real-Time PCR experiments. *Clin. Chem.* **55**, 611–622 (2009).
42. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta delta C(T)) method. *Methods San Diego Calif.* **25**, 402–408 (2001).
43. Luo, W., Friedman, M. S., Shedden, K., Hankenson, K. D. & Woolf, P. J. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinform.* **10**, 161 (2009).
44. Luo, W., Brouwer, C. & Pathview An R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **29**, 1830–1831 (2013).
45. Haw, R., Hermjakob, H., D'Eustachio, P. & Stein, L. Reactome pathway analysis to enrich biological discovery in proteomics data sets. *Proteomics* **11**, 3598–3613 (2011).

Acknowledgements

The authors would like to acknowledge the clinical staff of the hepatopancreatobiliary unit, Department of Surgery, Chris Hani Baragwanath Academic Hospital, for their assistance with patient recruitment and sample collection.

Author contributions

J.D. collected the data/sample. J.D, P.B and E.E.N. performed the experiments. J.D, P.B, N.E, P.M and E.E.N analysed the results and generated relevant figures/tables. J.D., P.M, N.E and E.E.N wrote the main manuscript text. M.S and G.P.C obtained the funding. All authors reviewed the manuscript and provided critical feedback.

Funding

The study was partially funded by the South African Medical Research Council through a grant awarded to the Wits Common Epithelial Cancer Research Centre.

Declarations

Competing interests

The authors declare no competing interests.

Patient consent for publication

Before sample and data collection, all patients recruited for this study provided written consent.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-23056-1.

Correspondence and requests for materials should be addressed to E.E.N.