Oncology

Oncology and Informatics

Oncology 2026;104:956–969
DOI: 10.1159/000549897

Received: September 12, 2025
Accepted: November 28, 2025
Published online: December 22, 2025

Transcriptomic Characterization of North Queensland Hepatocellular Carcinoma

Rhys Gillman[a, b] Miriam Wankell[a] Eun Jin Sun[a, b] Matan Ben David[a, c] Rozemary Karamatic[a, d] Pranavan Palamuthusingam[e, f] Matt A. Field[a, b, g, h] Ulf Schmitz[a, b, i] Lionel Hebbard[a, b, j, k]

[a]Department of Biomedical Sciences and Molecular and Cell Biology, College of Medicine and Dentistry, College of Science and Engineering, James Cook University, Townsville, QLD, Australia; [b]Centre for Tropical Bioinformatics and Molecular Biology, James Cook University, Cairns, QLD, Australia; [c]Department of Surgery, Townsville Hospital and Health Service, Townsville, QLD, Australia; [d]Gastroenterology and Hepatology, Townsville University Hospital, Townsville, QLD, Australia; [e]Department of Surgery, Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia; [f]Department of Surgery, The Wesley Hospital, Brisbane, QLD, Australia; [g]Immunogenomics Lab, Garvan Institute of Medical Research, Darlinghurst, NSW, Australia; [h]Menzies School of Health Research, Charles Darwin University, Darwin, NT, Australia; [i]Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia; [j]Storr Liver Centre, Westmead Institute for Medical Research, Westmead Hospital and University of Sydney, Sydney, NSW, Australia; [k]Australian Institute for Tropical Health and Medicine, Townsville, QLD, Australia

Keywords
Hepatocellular carcinoma · Liver cancer · Rural area · Transcriptome · Driver genes

Abstract

Introduction: Hepatocellular carcinoma (HCC) is a growing burden, particularly in rural, regional, and remote areas, but samples from these communities are underrepresented in public cancer data repositories. It remains unclear whether the findings of large, commonly studied cohorts such as The Cancer Genome Atlas (TCGA) are applicable to these remote communities. ***Methods:*** We profiled paired tumour and adjacent non-tumour liver biopsies from 19 patients admitted to the Townsville University Hospital in rural Australia. We used RNA-seq to characterize transcriptomic and mutational features and compared these with the TCGA Liver Hepatocellular Carcinoma (LIHC) cohort. Furthermore, we used these data to test a transcriptome-only adaptation of our TARGET-SL pipeline for low-cost drug target prediction. ***Results:*** Differential expression analysis identified 923 genes altered in our cohort, of which 64% overlapped with TCGA-LIHC, and the cohort-mean gene expression correlated strongly (Spearman rho = 0.96). Somatic variant calling from RNA highlighted mutational heterogeneity, with CTNNB1 (47%) and TP53 (21%) the most frequently mutated genes, consistent with TCGA findings. Copy number inference detected recurrent deletions on 8p, 6q, and 17p, congruous with known HCC patterns. We ran TARGET-SL solely on RNA-seq to identify personalized driver genes in these patients and were able to identify a drug candidate in 63% of patients. ***Conclusion:*** Our results

karger@karger.com
www.karger.com/ocl

Correspondence to:
Lionel Hebbard, lionel.hebbard@jcu.edu.au

demonstrate that NQ HCC shares core molecular features with larger TCGA cohorts and that a transcriptome-based approach can feasibly support precision oncology in resource-limited regional settings.

Introduction

Primary liver cancer is the 6th most common cancer globally, yet it ranks as the 3rd highest in cancer-related mortality overall and the 2nd highest in males, largely due to its poor prognosis [1]. Alarmingly, although liver cancer incidence has historically been relatively low in Australia, incidence and mortality are growing faster than in any other country. Of note, liver cancer has shown the highest increase in mortality of any cancer over the past 4 decades, increasing by 204.3% between 1982 and 2019 [2]. Furthermore, there is a clear trend of increasing incidence and poorer survival in rural and regional and remote areas relative to major cities in Australia [3].

Primary liver cancer includes several distinct subtypes; however, hepatocellular carcinoma (HCC) is by far the most predominant form, accounting for approximately 75% of cases [4]. The risk factors of HCC are diverse, though historically it has been commonly associated with infection from hepatitis B (HBV) and C virus. However, due to improved management and treatment of viral hepatitis in Australia, there has been a temporal change towards other risk factors, namely, metabolic dysfunction-associated steatotic liver disease (MASLD) and steatohepatitis and alcohol-related liver disease [5].

HCC is a cancer with poor prognosis and limited treatment options that only improve life expectancy by several months [6]. Standardizing treatment for HCC is challenging as HCC is known to be extremely genetically heterogeneous both within and between patients [7], contributing to the low efficacy of available chemotherapies. For example, of the 6 most frequently recurring mutations in liver cancer, no single one is present in more than 31% of patients [8, 9]. As such, the Gastroenterological Society of Australia (GESA) consensus on HCC frames treatment decision-making around tumour size, number, liver function, and anatomical staging, rather than on genomic or molecular profiling of the tumour [10]. This is also reflected in international guidelines from the American Association for the Study of Liver Diseases (AASLD) [11] and the European Association for the Study of the Liver (EASL) [12].

This has made apparent the need for precision medicine implementation into HCC therapy, and the development of targeted therapies for rare mutations. This requires study of the molecular causes of the disease and methods to better understand the genetic diversity amongst patients. Most molecular research into HCC draws from major public databases of publicly available sequencing data, notably The Cancer Genome Atlas (TCGA) project, which is entirely sourced from the United States of America and has a high HBV burden [8]. However, it remains unclear whether the differing major disease aetiologies of geographically different cohorts greatly affect the molecular landscape of the disease. Moreover, this gap in knowledge is potentially highly impactful to clinical decision-makers in underrepresented communities, as new patient-centred targeted therapies become available. Therefore, here we aimed to assess the molecular landscape of liver cancer in remote, rural, and regional North Queensland (NQ), Australia, and to compare this cohort with the TCGA.

Although the need for precision medicine implementation into clinical practice is clear, it is yet to become a reality due to a variety of factors including cost, limited approved targeted therapies, and largely untested software prediction pipelines. To address this, we tested our recently developed computational pipeline TARGET-SL (Tumour-specific Algorithm for Ranking Genetic Targets via Synthetic Lethality) to predict novel drug targets in the NQ patients. By default, TARGET-SL requires transcriptomic data and genetic variant information, thus relying on both genomic and transcriptomic sequencing. However, with the development of improved methods to detect variants from transcriptomic data, the potential for low-cost drug-target prediction using only transcriptomic sequencing was explored. Importantly, the reduced cost of RNA sequencing (~USD 130 per sample) makes such an approach clinically viable in some health care systems. Therefore, this study also serves as a proof of concept for implementing TARGET-SL as a low-cost drug-prediction method using only transcriptomic sequencing.

Methods

Patients

Consenting patients with suspected HCC undergoing resection surgery at the Townsville University Hospital (TUH) were recruited between January 1, 2021, and December 20, 2024. The study was performed in concordance with the NHMRC National Statement on Ethical Conduct in Human Research (NHMRC 2007, updated 2018) and approved by the Townsville Health Service Human Research Ethics Committee (HREC/16/QTHS/245) and the James Cook University Human Ethics Committee

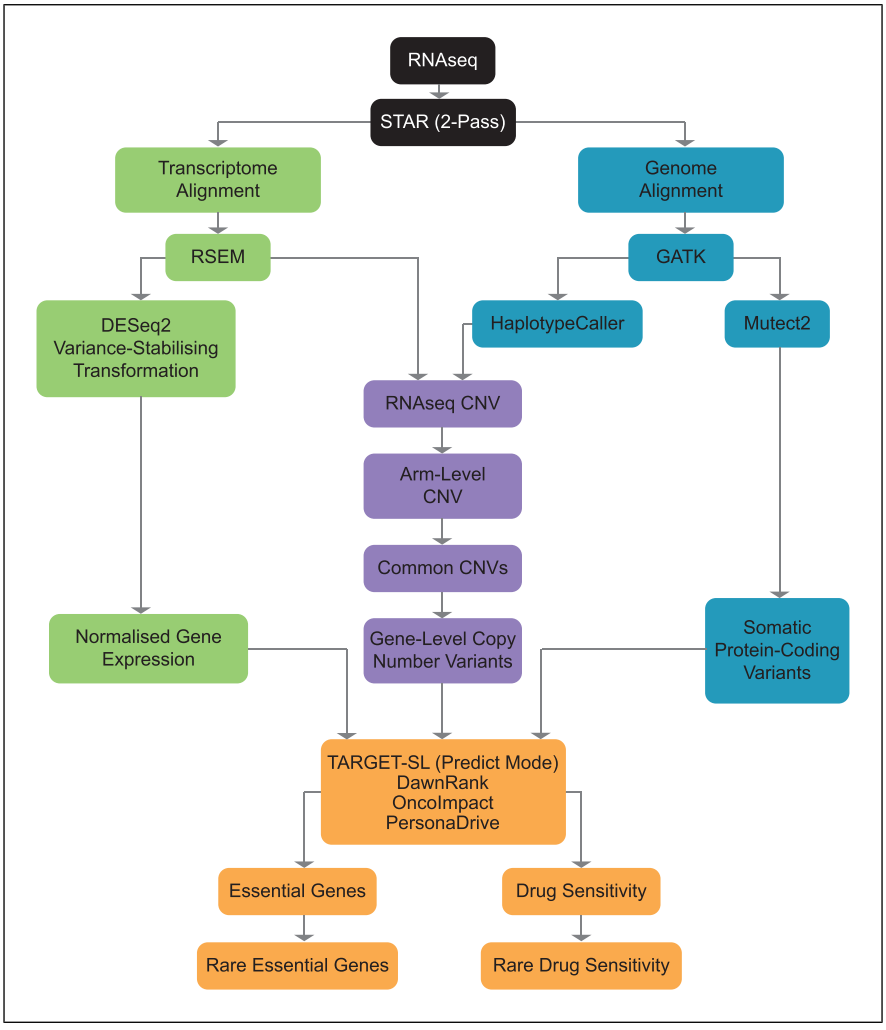


Fig. 1. Overview of RNA-seq-only approach to essential gene prediction using TARGET-SL.

(H7985). For publication, patients were de-identified, and their characteristics, including age, sex, cancer type, tumour type and stage, pathology, HBV and HBC status, alcohol and smoking history, body mass index, and serum markers, alpha foetal protein (AFP), alkaline phosphatase (ALP), alanine transaminase (ALT), aspartate transaminase (AST), gamma glutamyl transferase (GGT), albumin, and bilirubin, were recorded.

RNA Sequencing Analysis

Total RNA was extracted from snap-frozen biopsies of tumour core and adjacent non-tumour regions of resected tumours using the Isolate II RNA Mini Kit (BIO-52073, Meridian Bioscience), according to manufacturer guidelines. Illumina (NovaSeq X Plus) sequencing was performed by Novogene. Sequencing was performed to a depth of >50 million (100 million paired) reads per sample. Sequencing data were aligned using STAR (v2.7.11b) [13] with gene expression quantification using RSEM (v1.3.3) [14] and differential gene expression analysis using DESeq2 (v1.42.1) [15]. Variants were also called on the STAR-aligned RNA-seq data following GATK best practices and using GATK4 (v4.6.1.0) [16], with HaplotypeCaller for germline variants and Mutect2 for somatic variants. Mutational signature of somatic variants were detected using MutSignatures (v2.1.5) [17]. Copy number inference was performed using RNAseqCNV (v1.2.2) [18]. Finally, TARGET-SL [19] was used to predict essential genes and rare-essential genes with corresponding small-molecule inhibitors using differentially expressed genes, mutated genes, and copy number-altered genes as input. Finally, cancer-approved drugs were annotated based on their presence in CancerDrugsDB (https://www.anticancerfund.org/en/database-cancer-drugs). An overview of this pipeline is presented in Figure 1. Further details are available in the online supplementary Methods (for all online suppl. material, see https://doi.org/10.1159/000549897).

Table 1. Categorical patient characteristics

Feature	Value	Count	Percentage
Cancer type	HCC	16	84.21
	Intrahepatic cholangiocarcinoma and HCC	1	5.26
	Hepatic adenoma	1	5.26
	Bile duct adenoma	1	5.26
Gender	M	15	78.95
	F	4	21.05
Multifocal tumour	True	5	26.32
	False	11	57.89
	Unavailable	3	15.79
Tumour stage	pT1a	1	5.26
	pT1b	5	26.32
	pT2	7	36.84
	pT3	2	10.53
	Unavailable	4	21.05
Cirrhosis present	True	8	42.11
	False	7	36.84
	Unavailable	4	21.05
Steatosis present	True	10	52.63
	False	4	21.05
	Unavailable	5	26.32
Fibrosis present	True	11	57.89
	False	2	10.53
	Unavailable	6	31.58
History of hepatitis C	True	8	42.11
	False	11	57.89
History of hepatitis B	True	1	5.26
	False	18	94.74
History of alcohol consumption	True	17	89.47
	False	2	10.53
History of smoking	True	10	52.63
	False	7	36.84
	Unavailable	2	10.53
BMI classification	Healthy weight	6	31.58
	Overweight	8	42.11
	Obese	5	26.32

BMI, body mass index.

Results

Patient Clinical Characteristics

Tissue biopsies were acquired from 19 patients undergoing resection surgery for suspected HCC at the Townsville University Hospital (TUH). Samples were taken from the resected tumour and adjacent non-tumour tissues. Patients had a mean age of 68.89 (±7.89 SD) years. Based on available clinical data, 16 (84.21%) patients had confirmed HCC, and one (5.2%) had a combined intrahepatic cholangiocarcinoma and HCC. The remaining 2 patients were identified as having hepatic and bile duct adenomas, respectively. Patients were mostly male (78.95%) with single tumours (57.89%) and were mostly stage pT1b to pT2 (63.15% combined). In terms of surrounding liver condition, most patients presented with cirrhosis (42.11%), fibrosis (57.89%), and steatosis (52%). The most common HCC risk factors were alcohol consumption (89.46%), overweight or obese (68.42% combined), smoking (52.63%), and

Table 2. Numerical patient characteristics

Feature	Mean (±SD)	*n*
Age, years	68.89 (±7.89)	19
BMI, kg/m^2	27.47 (±4.42)	19
Tumour size	42.37 (±20.06)	19
AFP, ng/mL	1,918.27 (±3,197.34)	11
ALP, U/L	103.74 (±54.52)	19
ALT, U/L	54.11 (±53.99)	19
AST, U/L	54.11 (±50.08)	19
GGT, U/L	94.79 (±75.92)	19
Albumin, g/L	37.95 (±3.78)	19
Bilirubin, μmol/L	13.84 (±7.53)	19

BMI, body mass index.

hepatitis C virus (42%) (Tables 1, 2). The full characteristics of the patients are supplied in online supplementary Table 1.

Analysis of Differential Gene Expression in NQ Liver Cancer

In total, 923 significantly differentially expressed genes (DEGs) were identified, 653 of which were down-regulated and 270 upregulated relative to adjacent non-tumour tissue (shown in Fig. 2a). There was particularly strong down-regulation of two innate-immune gene families, the C-type lectins *CLEC4G*, *CLEC4M*, and *CLEC1B*, and ficolin genes *FCN1* and *FCN3*. We also noted strong up-regulation of genes in classic oncogenic pathways including the epidermal growth factor receptor kinase substrate *EPS8L3* and DNA replication and mitosis-related genes including *CENPF* involved in chromosomal segregation, the topoisomerase *TOP2A*, the microtubule-associated protein *DLGAP5*, cytokinesis protein *ANLN*, and the mitosis regulator *CDC20*. However, gene ontology analysis of all DEGs revealed that the most significant enrichment was in various liver metabolic pathways (shown in Fig. 2b). These results suggest a suppression of innate immune signalling and enhancement of cell proliferation and migration, along with metabolic reprogramming.

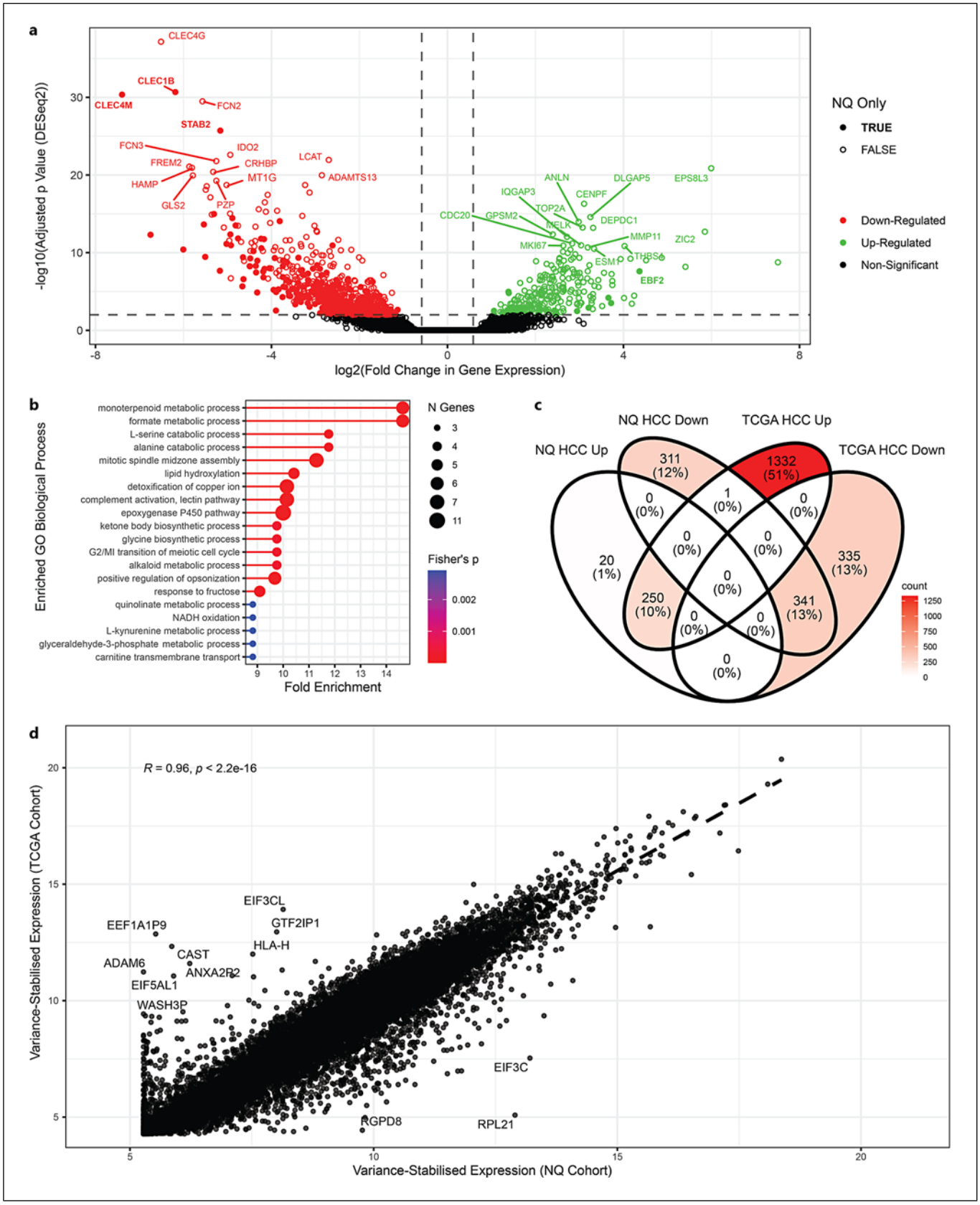
As a comparison, we performed an unpaired differential expression analysis on liver cancer samples from the TCGA LIHC cohort, including 373 tumour and 50 non-tumour samples. We found 2,259 significantly DEGs in the cohort, including 676 down-regulated and 1,583 up-regulated genes. We then compared the overlap of significantly up-regulated and down-regulated genes between the NQ and TCGA cohorts. In total, 591/923 (64%) of NQ DEGs overlapped with the TCGA DEGs. In addition, only a single DEG from either cohort showed the opposite trend in the other cohort (shown in Fig. 2c). Finally, we considered the global transcriptomic correlation between the two cohorts and found very strong and significant correlation (Spearman rho 0.96, p = 2.2e−16) with few outliers (shown in Fig. 2d). Outliers almost entirely comprised pseudogenes or poorly characterized genes, except for three translation-initiation factors (*EIF3C*, *EIF3CL*, and *EIF5AL1*), calpastain (*CAST*), and a component of the large 60S ribosomal subunit (*RPL21*).

Analysis of Somatic Variants in NQ Liver Cancer

Next, we identified 675 unique somatic variants across all sample transcriptomes. Using mutSignatures [17], we evaluated transcriptome-wide mutation patterns, showing a strong bias for T>C substitutions (shown in Fig. 3a). These variants were predominantly missense single-nucleotide variants (53.3%), followed by equal fractions of small insertion/deletions and synonymous single-nucleotide variants (21.2%) (shown in Fig. 3b). Individual patient mutation profiles were mapped to known COSMIC mutational signatures, with variants most frequently assigned to SBS26 (defective DNA mismatch repair), SBS40c (unknown aetiology), SBS54 (possible sequencing artefact) (shown in Fig. 3c).

Non-synonymous somatic variants were identified in 496 genes across the cohort. The number of mutated genes per patient varied from 5 to 42, with the lowest number of mutated genes identified in the patient with a bile duct adenoma. Notably, the patient with a hepatic adenoma harboured a high mutational burden (shown in Fig. 4a). Mutated genes were highly heterogeneous, with the most frequently mutated genes being *CTNNB1* (47.37%), *TP53* (21.05%), and *EIF3I* (21.05%). Interestingly, the patient with a combined intrahepatic

Fig. 2. Analysis of gene expression in North Queensland (NQ) liver cancer. **a** Volcano plot of DEGs in the NQ liver cancer cohort (DESeq2, alpha = 0.01, log2 fold change threshold = 0.58). Filled points indicate NQ differentially expressed genes (DEGs) that were not significant in The Cancer Genome Atlas (TCGA) cohort. **b** Gene Ontology (GO) biological process enrichment of significant DEGs in the NQ cohort. **c** Venn diagram indicating overlap of up- and down-regulated DEGs between the NQ and TCGA cohorts. **d** Correlation (Spearman) plot of variance-stabilized tumour gene expression between the NQ and TCGA cohorts.



(For legend see previous page.)

Fig. 3. Analysis of somatic mutational signature in North Queensland (NQ) liver cancer. **a** Tri-nucleotide context substitution frequencies in NQ somatic mutations. **b** Frequency of variant classifications in NQ somatic mutations. **c** Proportion of variant assignment to the Catalogue of Somatic Mutations in Cancer (COSMIC) SBS96 signatures.

cholangiocarcinoma and HCC shared no mutated genes within the top 20 most frequently mutated genes in the cohort, while still harbouring many mutated genes. Otherwise, we did not observe any obvious association between clinical features and gene mutations (shown in Fig. 4b, c).

We compared mutation frequency within 363 TCGA patients and found that *CTNNB1* mutations were more frequent in the NQ cohort, though *TP53* and *CTNNB1* were similarly the most frequently mutated genes. Interestingly, *EIF3I* mutations were not present in a single TCGA tumour. Furthermore, some highly mutated genes in the TCGA cohort were not detected in the NQ cohort, including *TTN*, *MUC16*, and *PCLO* (shown in Fig. 4d). Statistical comparison of the mutation frequencies was not applicable due to the low sample sizes.

Analysis of Copy Number Variants in NQ Liver Cancer

Subsequently, we used RNAseqCNV to infer arm-level copy number variants (shown in Fig. 5). High-evidence deletions were most common in chromosome

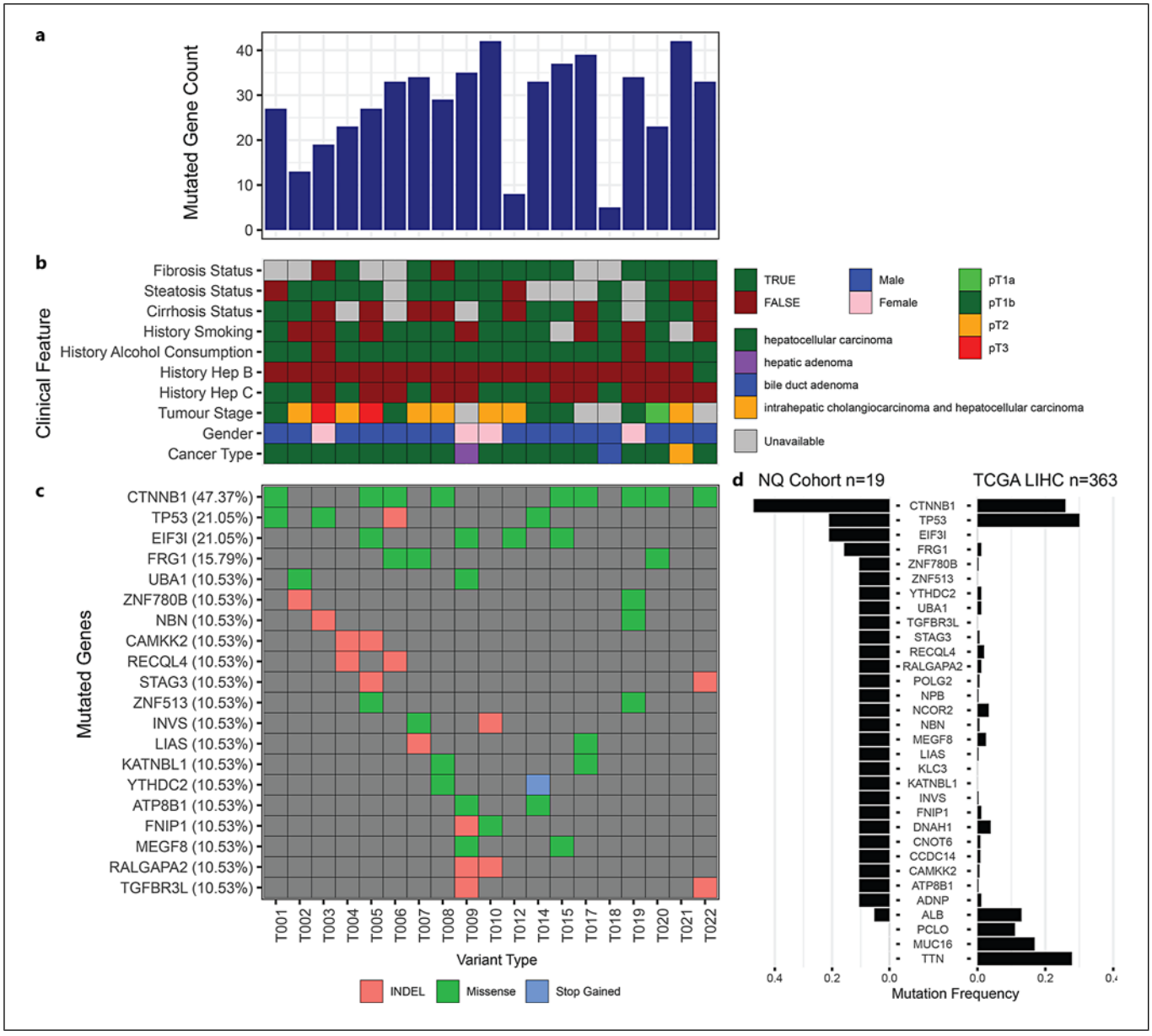


Fig. 4. Analysis of non-synonymous somatic mutated genes in North Queensland (NQ) liver cancer. **a** Numbers of mutated genes per patient. **b** Clinical features of NQ cancer patients. **c** Top 20 most-frequently mutated genes in the NQ cohort and heatmap of mutated genes per patient. **d** Comparison of the mutation frequency (proportion of patients harbouring a mutation in the gene) between the NQ and TCGA cohorts, showing all genes with >10% mutation frequency in either cohort.

arms 8p, 6q, and 17p. While no high-evidence amplifications were assigned, low-evidence amplifications were frequent in chromosome arms 8p and 1p. Several patients showed evidence of entire chromosomal deletions, including chromosomes 14, 16, and 21 occurring in 2 patients each, as well as deletions of chromosomes 18 and 22 in only a single patient each.

Personalized Driver Gene and Essential Gene Prediction

Given the varied mutational profile of all patients, we used TARGET-SL [19] to predict patient-specific driver genes, rare essential genes, and drugs for each patient. TARGET-SL was first used to find a consensus of personalized driver genes from the DawnRank,

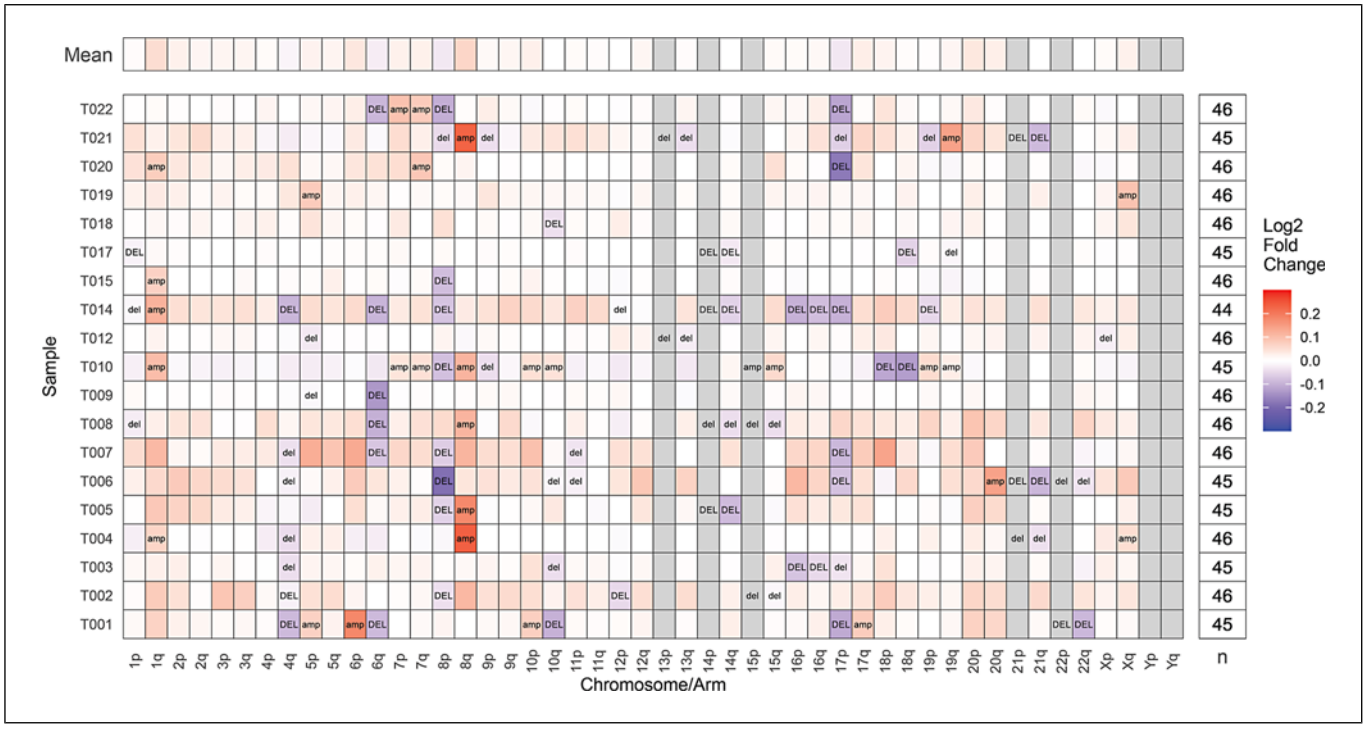


Fig. 5. Analysis of inferred chromosome arm-level copy numbers in North Queensland (NQ) liver cancer. Copy number was inferred from RNA-seq data using RNAseqCNV. Arm-level alterations with "high evidence" are indicated by capitalized annotations, while those with "low evidence" are indicated by lower-case annotations. Grey boxes had insufficient read depth to calculate log2 fold change. Top row indicates mean log2 fold change across all samples. Right-hand column indicates the predicted number of chromosomes per sample.

PersonaDrive, and OncoImpact algorithms. We considered the frequency at which drivers were ranked within the top 10 for each patient and their mean rank (shown in Fig. 6a). Many well-known driver genes listed in the Cancer Gene Census received high driver prediction rankings, including *CTNNB1*, *TP53*, *RB1*, *CREBBP*, and *TCF7L2*. *CTNNB1* was by far the most frequently predicted driver gene, occurring in 47.37% of patients, and was the rank one driver in every case. *TP53* was also predicted to be a driver in all patients (21.05%) harbouring a *TP53* mutation, with a mean rank of 1.25. However, several less well-characterized mutated genes also received high driver rankings in small numbers of patients, including *RBL2*, *APOB*, *DRD4*, and *MED1*, and *SLC11A2*. Driver predictions were also relatively common for *TIRAP* and *MAPT*, at the same frequency as *TP53*, but with much lower average rankings.

TARGET-SL can also use personalized driver predictions to infer rare essential genes based on variant effect prediction and known or predicted synthetically lethal relationships. Again, we only considered the top 10 rare essential genes per patient (shown in Fig. 6b). Four genes occurred with relatively high frequency of 15.8%, namely, *EGFR*, *MTOR*, *ATR*, *MUS81*, *CDK9*, *CDC7*, and *CHEK1*, with all except *MUS81* having known small-molecule inhibitors (shown in Fig. 6b). Specifically, in all patients with an EGFR variant, it was the top ranked essential gene, and the HCC-approved tyrosine kinase inhibitors can suppress this receptor [20]. Additionally, several essential genes with approved therapies were identified in only a single (5.26%) patient and included *PIK3CA*, *BLM*, *FGFR*3, and *NRAS* and an additional four essential genes, *EZH2*, *TYK2*, *CSNK2A1*, and *KDM6A*, that have known inhibitors and are not currently cancer approved. These genes could be subject to further research, not to mention several high-ranking essential genes with no known inhibitors.

Importantly, of the essential genes predicted for the 19 patients, 12 patients (63.2%) recorded at least one available targeting drug (shown in Fig. 6c). Furthermore, drug predictions included approved cancer drugs in 10 patients (52.6%) and approved HCC drugs in just a single patient (5.26%). No essential genes were predicted for patient 018 with bile duct adenoma due to the small number of mutations.

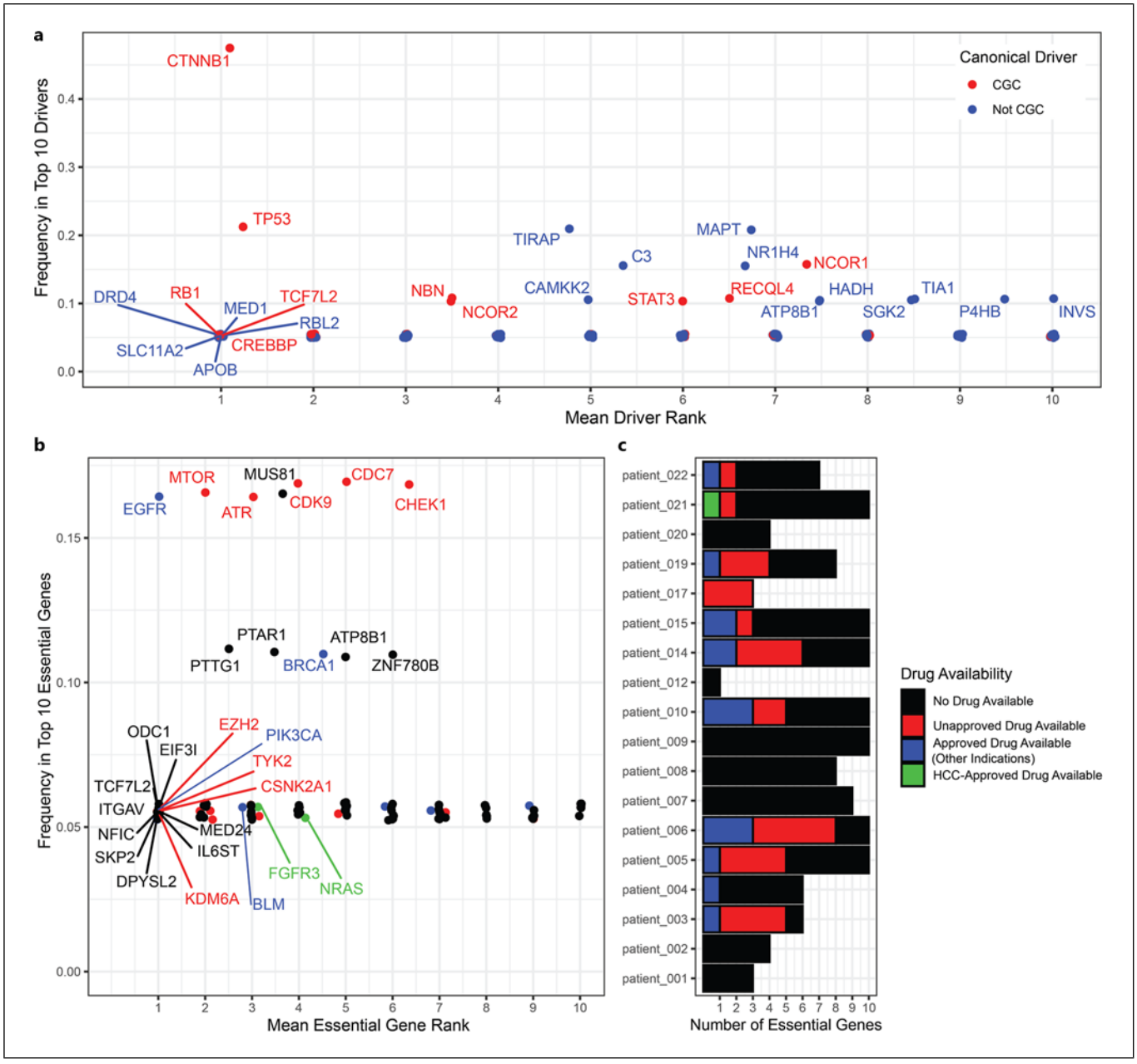


Fig. 6. Analysis of personalized driver genes and essential genes using TARGET-SL. We used TARGET-SL to generate a consensus list of personalized driver gene predictions for each patient based on DawnRank, OncoImpact, and PersonaDrive predictions and to generate rare essential gene and drug predictions for each patient. **a** We calculated the frequency at which genes were ranked in the top 10 drivers per patient and their mean rank. We also indicate whether predicted driver genes are canonical drivers present in the Cancer Gene Census (CGC). **b** We calculated the frequency at which genes were ranked in the top 10 rare essential genes per patient and their mean rank. Essential genes with available drug inhibitors are indicated and differentiated into whether they are approved cancer drugs according to CancerDrugsDB. **c** Number of predicted rare essential genes per patient and availability of drug inhibitors.

Discussion

The availability of large public cancer sequencing datasets has been invaluable for drug and biomarker discovery, shaping clinical decision-making around the world. Nonetheless, the relevance of these datasets to smaller communities remains unknown. To address this, we characterized the transcriptome of 19 liver cancer patients from NQ, Australia, and contrasted our results with the TCGA LIHC cohort. The disease aetiology of our NQ cohort differed most notably from the TCGA by HBV frequency, which was the most frequent risk factor (22%) in the TCGA cohort and represented by only a single patient (5.26%) in the NQ cohort. Furthermore, higher rates of steatosis and obesity in the NQ cohort suggest more MASLD-associated HCC in the NQ cohort, consistent with previous reports [5].

Gene expression was highly correlated between the cohorts, and most DEGs overlapped. DEGs were enriched in metabolic pathways, consistent with the known metabolic reprogramming, which occurs in HCC [21]. Interestingly, we observed very strong suppression of several pattern recognition receptors in the C-type lectin receptor (*CLR*) and ficolin gene families in the NQ cohort, consistent with findings in the TCGA cohort [22]. These genes play important roles in fungal defence and complement activation [23, 24]. A recent review hypothesized that low CLR expression can disrupt the gut-liver axis and contribute to HCC progression, suggesting that anti-fungal treatments could improve immunotherapy efficacy in such patients [24]. Also, recovering the suppressed expression of ficolins in HCC cell lines can reduce cell proliferation and migration [25]. Additionally, we observed the over-expression of many TCGA-overexpressed genes that have led to research into HCC treatments, such as a recent study demonstrating that TOP2A inhibition can reverse regorafenib resistance in vitro and in vivo [26], and the same has been found after suppressing *EPS8L3* [27]. Reassuringly, this suggests that research outputs derived from the TCGA are relevant to patients in NQ.

Somatic variant calling showed a strong bias for T>C transitions, followed by a much smaller proportion of C>T transitions. This is consistent with other cohorts, though more frequent C>T transitions and C>A transversions are often observed [28]. While the dominant assignment of the SBS26 mutational signature often suggests mismatch repair deficiency, we only observed a mismatch repair mutation (*MSH6)* in a single patient and no significant differential expression of mismatch repair enzymes. Indeed, mismatch repair deficiencies are relatively rare in HCC [29]. Nonetheless, T>C substitutions are a known hallmark of HCC with somewhat unknown aetiology [30].

We also inferred arm-level copy number variants. Chromosome amplifications were mostly limited to 1q and 8p, along with frequent deletions in 4q, 6q, 8p, 17p, reflecting an almost identical pattern to the TCGA [31]. The observed amplified regions most notably contain the *MCL1* (1p) and *MYC* (8p) oncogenes, while the deleted regions included the *LATS1* (6q) and *TP53* (17q) tumour suppressors, which are all commonly observed in various cancers [32].

At the gene level, somatic mutations were most frequent in *CTNNB1* and *TP53*, consistent with the TCGA and almost all other HCC cohorts globally. However, the NQ cohort lacked mutations in other notable frequently mutated genes. For example, we detected no *ARID2* mutations and only a single *ARID1A* mutation, considered some of the most frequently mutated HCC genes [33]. Also, the common *TERT* promoter mutation was unfortunately undetectable in our analysis due to our use of transcriptomic data. Surprisingly, we detected eukaryotic translation initiation factor 3 subunit I (*EIF3I*) variants at the same frequency as *TP53* variants. *EIF3I* variants have not been previously studied in HCC, though it is highly expressed in colorectal cancer and its knockdown impedes growth in vivo [34]. Unfortunately, our cohort lacks the statistical power to determine whether these mutation frequencies are significantly different from the TCGA, and further investigation is recommended to determine whether the drivers of HCC are significantly different in this community, which could impact decision-making regarding new research and therapies to target these genes.

As is known in HCC, we observed highly heterogenous mutations, making the case for using personalized driver prioritization algorithms [35]. We used TARGET-SL [19] to find a consensus ranking of personalized driver genes. Many high-ranking driver gene predictions were well-known cancer drivers that have been previously discussed in much detail including *CTNNB1*, *TP53*, *RB1*, *CREBBP*, and *TCF7L2* [36]. However, several lesser-known drivers were predicted with low frequency but very high ranking and have some previous research context to HCC. For example, rare apolipoprotein B (*APOB*) variants have been previously reported to be associated with

MASLD patients and increased risk of HCC [37]. Additionally, the elevated expression of dopamine receptor D4 (*DRD4*) is linked to liver cancer stem cell phenotype and poorer prognosis in patients [38], though genetic variants have not been investigated. *MED1* is an essential component of mediator complex, which orchestrates *PPARα*, *PPARγ*, *CAR*, and GR-mediated transcriptional regulation, and has been found to influence hepatocyte proliferation in the HL-7702 cell line [39]. Meanwhile, a single SNP in *RBL2* has been found to have a modest association with disease-free survival in HCC patients, though the same SNP was not detected in our cohort [40]. Variants in *SLC11A2*, also known as divalent metal transporter 1 (DMT), have not been investigated in HCC.

Finally, we used TARGET-SL to predict druggable essential genes and small-molecule inhibitors. A druggable essential gene was predicted in 63.2% of patients, and in 52.6% of patients an already licensed drug was predicted that could be utilized in drug repurposing. It should be noted that we only considered the "rare essential genes" output from TARGET-SL, which filters out essential genes that are common across many samples to identify more targeted predictions, and the rate of identifying druggable options increased to 89.5% when considering "all essential genes." While we were unable to verify the validity of these predictions in this study, we have previously shown TARGET-SL to have a modest success rate [19]. Regardless, the ability to predict a druggable target in nearly two-thirds of patients is potentially important for a disease with such poor prognosis and limited treatment options.

Limitations of the study include limited sample material and small sample size and the use of transcriptomic information only. It is likely that our research design has failed to capture many potentially causative mutations, either due to low-depth or them being non-exonic. Genomic sequencing would have greatly improved our ability to detect genetic variants but would have dramatically increased the required cost. Therefore, here we have demonstrated personalized driver gene and drug predictions at a very low cost of under USD 130 per sample for short-read RNA-seq, representing a clinically relevant approach to personalized oncology.

In summary, we sought to determine the applicability of treatment decisions derived from the TCGA HCC cohort to patients in NQ. We observed highly correlated gene expression and copy number patterns, consistent with accepted knowledge about the disease. We also observed heterogeneous mutations, most frequently in *CTNNB1* and *TP53*, though an unexpectedly high frequency of *EIF3I* mutations. This information is vital to NQ-based clinical decision-makers as new targeted therapies become available. Finally, as a secondary aim, we successfully applied TARGET-SL using RNA-seq data only, providing low-cost drug prediction for patients and identifying rare and novel driver genes.

Statement of Ethics

The study was performed in concordance with the NHMRC National Statement on Ethical Conduct in Human Research (NHMRC 2007, updated 2018) and approved by the Townsville Health Service Human Research Ethics Committee (HREC/16/QTHS/245) and the James Cook University Human Ethics Committee (H7985). All participants in the study provided written informed consent for the publication of de-identified sequencing and clinical information.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

This work was supported by Tour De Cure [RSP-379-FY2023 to R.G.]; James Cook University Postgraduate Research Scholarship (to R.G.); Tropical Australian Academic Health Centre Limited – Research Seed Grant [SF000121 to L.H.]; Townsville Hospital Health Service-Study Education Research Trust Account, Project and Capacity Building Grants (09 2020, RPG1 2023 and RCG2 2023, M.W. R.K., M.B.D., P.P., M.A.F., and L.H.); National Health and Medical Research Council Investigator Grants (#5121190, M.A.F.; #1196405, U.S.); and Cancer Council NSW project grant (RG20-12, U.S.). The funders had no role in the design, data collection, data analysis, and reporting of this study.

Author Contributions

R.G. performed data collection and analysis and prepared the manuscript. M.W., E.J.S, M.B.D., R.K., and P.P. assisted in data collection and reviewed the manuscript. M.A.F., U.S., and L.H. assisted in conceptualization of the project and reviewed the manuscript.

Data Availability Statement

The data that support the findings of this study are openly available in the Gene Expression Omnibus (GEO) (GSE304512).

References

1 Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49. https://doi.org/10.3322/caac.21660
2 Australian Institute of Health and Welfare. Cancer in Australia 2019. Canberra: AIHW; 2019.
3 Australian Institute of Health and Welfare. Cancer in Australia 2021. Canberra: AIHW; 2021.
4 Altekruse SF, Devesa SS, Dickie LA, McGlynn KA, Kleiner DE. Histological classification of liver and intrahepatic bile duct cancers in SEER registries. J Registry Manag. 2011;38(4):201–5.
5 Yeoh YKJ, Dore GJ, Lockart I, Danta M, Flynn C, Blackmore C, et al. Temporal change in aetiology and clinical characteristics of hepatocellular carcinoma in a large cohort of patients in New South Wales, Australia. Intern Med J. 2024;54(4):602–12. https://doi.org/10.1111/imj.16252
6 Falette Puisieux M, Pellat A, Assaf A, Ginestet C, Brezault C, Dhooge M, et al. Therapeutic management of advanced hepatocellular carcinoma: an updated review. Cancers. 2022;14(10):2357. https://doi.org/10.3390/cancers14102357
7 Wu Y, Liu Z, Xu X. Molecular subtyping of hepatocellular carcinoma: a step toward precision medicine. Cancer Commun. 2020;40(12):681–93. https://doi.org/10.1002/cac2.12115
8 Ally A, Balasundaram M, Carlsen R, Chuah E, Clarke A, Dhalla N, et al. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. Cell. 2017;169(7):1327–41.e23. https://doi.org/10.1016/j.cell.2017.05.046
9 Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. Nat Rev Clin Oncol. 2018;15(10):599–616. https://doi.org/10.1038/s41571-018-0073-4
10 Lubel JS, Roberts SK, Strasser SI, Thompson AJ, Philip J, Goodwin M, et al. Australian recommendations for the management of hepatocellular carcinoma: a consensus statement. Med J Aust. 2021;214(10):475–83. https://doi.org/10.5694/mja2.50885
11 Singal AG, Llovet JM, Yarchoan M, Mehta N, Heimbach JK, Dawson LA, et al. AASLD Practice Guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. Hepatology. 2023;78(6):1922–65. https://doi.org/10.1097/HEP.0000000000000466
12 European Association for the Study of the Liver. EASL Clinical Practice Guidelines on the management of hepatocellular carcinoma. J Hepatol. 2025;82(2):315–74. https://doi.org/10.1016/j.jhep.2024.08.028
13 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinforma Oxf Engl. 2013;29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635
14 Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics. 2011;12(1):323. https://doi.org/10.1186/1471-2105-12-323
15 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8
16 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–303. https://doi.org/10.1101/gr.107524.110
17 Fantini D, Vidimar V, Yu Y, Condello S, Meeks JJ. MutSignatures: an R package for extraction and analysis of cancer mutational signatures. Sci Rep. 2020;10(1):18217. https://doi.org/10.1038/s41598-020-75062-0
18 Bařinka J, Hu Z, Wang L, Wheeler DA, Rahbarinia D, McLeod C, et al. RNA-seqCNV: analysis of large-scale copy number variations from RNA-seq data. Leukemia. 2022;36(6):1492–8. https://doi.org/10.1038/s41375-022-01547-8
19 Gillman R, Field MA, Schmitz U, Hebbard L. TARGET-SL: precision essential gene prediction using driver prioritisation and synthetic lethality. Brief Bioinform. 2025;26(3):bbaf255. https://doi.org/10.1093/bib/bbaf255
20 Liu YC, Tsai JJ, Weng YS, Hsu FT. Regorafenib suppresses epidermal growth factor receptor signaling-modulated progression of colorectal cancer. Biomed Pharmacother. 2020;128:110319. https://doi.org/10.1016/j.biopha.2020.110319
21 Park S, Hall MN. Metabolic reprogramming in hepatocellular carcinoma: mechanisms and therapeutic implications. Exp Mol Med. 2025;57(3):515–23. https://doi.org/10.1038/s12276-025-01415-2
22 Sun J, Zhang Z, Cai J, Li X, Xu X. Identification of hub genes in liver hepatocellular carcinoma based on weighted gene Co-expression network analysis. Biochem Genet. 2025;63(3):2120–39. https://doi.org/10.1007/s10528-024-10803-8
23 Arnold JN, Mitchell DA. Tinker, tailor, soldier, cell: the role of C-type lectins in the defense and promotion of disease. Protein Cell. 2023;14(1):4–16. https://doi.org/10.1093/procel/pwac012
24 Xia J, Ding H, Liu S, An R, Shi X, Chen M, et al. C-Type lectin receptors-triggered antifungal immunity may synergize with and optimize the effects of immunotherapy in hepatocellular carcinoma. J Inflamm Res. 2023;16:19–33. https://doi.org/10.2147/JIR.S394503
25 Zheng G, Wu L, Bouamar H, Cserhati M, Chiu YC, Hinck CS, et al. Ficolin-3 induces apoptosis and suppresses malignant property of hepatocellular carcinoma cells via the complement pathway. Life Sci. 2024;357:123103. https://doi.org/10.1016/j.lfs.2024.123103
26 Wang Z, Zhu Q, Li X, Ren X, Li J, Zhang Y, et al. TOP2A inhibition reverses drug resistance of hepatocellular carcinoma to regorafenib. Am J Cancer Res. 2022;12(9):4343–60.
27 Chen B, Pan Y, Xu X, Wu F, Zheng X, Chen SY, et al. Inhibition of EPS8L3 suppresses liver cancer progression and enhances efficacy of sorafenib treatment. Biomed Pharmacother. 2020;128:110284. https://doi.org/10.1016/j.biopha.2020.110284
28 Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. Nature. 2020;578(7793):94–101. https://doi.org/10.1038/s41586-020-1943-3
29 Mukai S, Kanzaki H, Ogasawara S, Ishino T, Ogawa K, Nakagawa M, et al. Exploring microsatellite instability in patients with advanced hepatocellular carcinoma and its tumor microenvironment. JGH Open. 2021;5(11):1266–74. https://doi.org/10.1002/jgh3.12660
30 Müller M, Bird TG, Nault JC. The landscape of gene mutations in cirrhosis and hepatocellular carcinoma. J Hepatol. 2020;72(5):990–1002. https://doi.org/10.1016/j.jhep.2020.01.019
31 Damrauer JS, Smith MA, Walter V, Thennavan A, Mose LE, Selitsky SR, et al. Genomic characterization of rare molecular subclasses of hepatocellular carcinoma. Commun Biol. 2021;4(1):1150. https://doi.org/10.1038/s42003-021-02674-1
32 Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010;463(7283):899–905. https://doi.org/10.1038/nature08822
33 Lee JS. The mutational landscape of hepatocellular carcinoma. Clin Mol Hepatol. 2015;21(3):220–9. https://doi.org/10.3350/cmh.2015.21.3.220
34 Zhang Y, Wan X, Yang X, Liu X, Huang Q, Zhou L, et al. eIF3i promotes colorectal cancer cell survival via augmenting PHGDH translation. J Biol Chem. 2023;299(9):105177. https://doi.org/10.1016/j.jbc.2023.105177
35 Gillman R, Field MA, Schmitz U, Karamatic R, Hebbard L. Identifying cancer driver genes in individual tumours. Comput Struct Biotechnol J. 2023;21:5028–38. https://doi.org/10.1016/j.csbj.2023.10.019

36 Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive characterization of cancer driver genes and mutations. Cell. 2018;173(2):371–85.e18. https://doi.org/10.1016/j.cell.2018.02.060
37 Mureddu M, Pelusi S, Jamialahmadi O, Periti G, Ronzoni L, Heigadeh H, et al. Carriage of rare pathogenic apob variants predispose to severe masld and hcc. Dig Liver Dis. 2025;57:S64–5. https://doi.org/10.1016/j.dld.2025.01.119
38 Yang Z, Zhang P, Zhao Y, Guo R, Hu J, Wang Q, et al. DRD4 promotes chemoresistance and cancer stem cell-like phenotypes by mediating the activation of the Akt/β-catenin signaling axis in liver cancer. Br J Cancer. 2024;131(7):1212–23. https://doi.org/10.1038/s41416-024-02811-7
39 Liu J, Zhu J, Zhang X, Jia Y, Lee X, Gao Z. Hsa-miR-637 inhibits human hepatocyte proliferation by targeting Med1-interacting proteins. Liver Res. 2021;5(2):88–96. https://doi.org/10.1016/j.livres.2021.01.002
40 Liu S, Yang TB, Nan YL, Li AH, Pan DX, Xu Y, et al. Genetic variants of cell cycle pathway genes predict disease-free survival of hepatocellular carcinoma. Cancer Med. 2017;6(7):1512–22. https://doi.org/10.1002/cam4.1067