

Article

Deciphering Tumour Microenvironment of Liver Cancer through Deconvolution of Bulk RNA-Seq Data with Single-Cell Atlas

Shaoshi Zhang ^{1,2}, Wendi Bacon ³, Maikel P. Peppelenbosch ², Folkert van Kemenade ³ and Andrew Peter Stubbs ^{1,*}

¹ Department of Pathology and Clinical Bioinformatics, Erasmus University Medical Centre, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands

² Department of Gastroenterology and Hepatology, Erasmus University Medical Centre, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands

³ School of Life, Health & Chemical Sciences, Faculty of Science, Technology, Engineering & Mathematics, The Open University, Milton Keynes MK7 6AA, UK

* Correspondence: a.stubbs@erasmusmc.nl

Simple Summary: ScRNA-seq is a powerful tool for investigating the cancer microenvironment, but the cost of analysing every scientific scenario is prohibitive. Fortunately, deconvolution of bulk RNA-seq data with scRNA-seq cell atlas reference datasets provides a cheaper strategy. In this study, we validated the feasibility of deciphering the microenvironment of liver cancer through the estimation of cell fractions with Cibersortx and scRNA-seq atlas reference datasets. Five cell types are associated with patient outcomes, showing that deconvolution is a useful method for characterising the tumour microenvironment.

Abstract: Liver cancers give rise to a heavy burden on healthcare worldwide. Understanding the tumour microenvironment (TME) underpins the development of precision therapy. Single-cell RNA sequencing (scRNA-seq) technology has generated high-quality cell atlases of the TME, but its wider application faces enormous costs for various clinical circumstances. Fortunately, a variety of deconvolution algorithms can instead repurpose bulk RNA-seq data, alleviating the need for generating scRNA-seq datasets. In this study, we reviewed major public omics databases for relevance in this study and utilised eight RNA-seqs and one microarray dataset from clinical studies. To decipher the TME of liver cancer, we estimated the fractions of liver cell components by deconvoluting the samples with Cibersortx using three reference scRNA-seq atlases. We also confirmed that Cibersortx can accurately deconvolute cell types/subtypes of interest. Compared with non-tumorous liver, liver cancers showed multiple decreased cell types forming normal liver microarchitecture, as well as elevated cell types involved in fibrogenesis, abnormal angiogenesis, and disturbed immune responses. Survival analysis shows that the fractions of five cell types/subtypes significantly correlated with patient outcomes, indicating potential therapeutic targets. Therefore, deconvolution of bulk RNA-seq data with scRNA-seq atlas references can be a useful tool to help understand the TME.

Keywords: liver cancer; tumour microenvironment; deconvolution



Citation: Zhang, S.; Bacon, W.; Peppelenbosch, M.P.; van Kemenade, F.; Stubbs, A.P. Deciphering Tumour Microenvironment of Liver Cancer through Deconvolution of Bulk RNA-Seq Data with Single-Cell Atlas. *Cancers* **2023**, *15*, 153. <https://doi.org/10.3390/cancers15010153>

Academic Editor: Diego F. Calvisi

Received: 11 November 2022

Revised: 21 December 2022

Accepted: 24 December 2022

Published: 27 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Liver cancer is one of the leading causes of cancer-related mortality worldwide, making up 4.7% of newly diagnosed cases but 8.2% of deaths [1]. Hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA), which are frequently tallied together, constitute the major burden of liver cancer [2]. With a 5-year survival of 18%, liver cancer ranks as the second-most lethal cancer [3]. The poor prognosis of liver cancer is partially due to the insufficiency

of effective therapies. Surgical interventions yield the best outcomes but are limited by high recurrence rates or easy loss of operative windows. Liver transplantation achieves better long-term survival but is hampered by an inadequate supply of donor organs [4,5]. Systematic non-specific chemotherapy resulted in disappointing results for both HCC and CCA [5,6]. Innovative agents targeting angiogenesis [7,8], fibrogenesis [9], and regulation of immune responses [10,11] have shown the potential to improve outcomes. Such advances suggest that a shift from a cancer-centric to a tumour microenvironment (TME)-centric perspective is required in the future development of precision therapy [12,13].

Single-cell sequencing (scRNA-seq) technology delivers in-depth interrogation of the TME. The analysis of complex cancer tissues at the single-cell level through scRNA-seq has brought insights into the heterogeneity and progression of cancer, as well as escape from immune surveillance, drug resistance, and intercellular communication [14]. However, scRNA-seq technology is expensive, requires specific tissue collection methods, and can be difficult to implement. The cheaper and more common bulk RNA-seq studies make up the largest body of work in this area, filling public repositories, including flagship projects such as The Cancer Genome Atlas (TCGA) and its resulting resource the Pan-Cancer Atlas [15]. To make the most out of these available datasets, numerous algorithms have been proposed that enhance the informativeness of bulk transcriptome analysis using reference scRNA-seq data. Machine learning is the major approach of such methodologies. For example, stemness indices within the Pan-Cancer Atlas were estimated with the Progenitor Cell Biology Consortium datasets and one-class logistic regression [16]. Support vector regression was also used to estimate the abundance of a specific component in bulk RNA-seq samples [17,18]. Another group of deconvolution algorithms, such as MuSiC [19] and Cibersortx [20], focus on profiling cell fractions of the bulk transcriptomic data. Meanwhile, emerging scRNA-seq atlases (e.g., Human Cell Atlas) provide high-quality references, allowing accurate deconvolution of bulk RNA-seq data in increasingly wider contexts (e.g., profiling of TME for head and neck cancers [21]).

In this study, we reviewed major omics databases and selected ten studies that compared transcriptomes between HCC/CCA and normal liver. The cell fractions of tumour and non-tumour tissues were estimated with Cibersortx and three scRNA-seq reference atlases. The included studies contain two cohort studies of HCC, allowing us to investigate the clinical implications of TME abnormalities through survival analysis and gene set enrichment analysis (GSEA) [22]. We determined that the TMEs of liver cancer lack multiple cell types (e.g., hepatocytes) that form normal liver microarchitectures, and instead are enriched with components involved in fibrogenesis, abnormal angiogenesis, and irregular immune activities. Among all the cell types/subtypes in the HCC TME, hepatocytes and mature B cells are positively correlated with patient outcomes, while cholangiocytes, bi-potent stem cells, plasma B cells, and regulatory T (T_{reg}) cells correlate with negative outcomes.

2. Materials and Methods

2.1. Data Obtainment

We searched three public omics databases—the Gene Expression Omnibus (GEO), The Cancer Genome Atlas (TCGA), and the International Cancer Genome Consortium (ICGC)—for studies engaging liver cancers. Inclusion criteria were: (1) histologically identified sample series from human tissues in clinical studies (including cohort studies and case review series); (2) whole transcriptomes by microarray or RNA-seq; (3) with non-tumorous or normal liver tissues as controls (for comparison) or follow-up information (for interrogation of clinical outcomes); (4) studies covering two major liver cancers (HCC and CCA) were included and all other studies were excluded, i.e., metastatic liver cancers and non-malignant hyperplasia (e.g., hemangioma). Finally, nine curated datasets (eight RNA-seq and one microarray) were selected for this study (Table S1).

2.2. Pre-Processing of Microarray Data

Microarray studies with raw data (CEL files) from the GEO database were obtained via R package SCAN.UPC [23], which provides one-step pre-processing through empirical correction of major bias (GC content-related). This normalisation method for microarray datasets proves reliable for downstream analysis [24]. All gene names were translated into HGNC symbol with R package BioMart [25]. Duplication of features in expression matrices were collapsed by MaxMean strategy [26].

2.3. Pre-Processing of RNA-Seq Data

TCGA-LIHC was obtained via UCSCXenaTools [27] and datasets in ICGC were obtained via the official web portal [28]. Expression matrices of other studies from the GEO database were retrieved according to the authors' instructions. All datasets recruited in this study have been listed in Supporting Information. All gene names were translated to HGNC symbols with R package BioMart [25]. Feature duplications in expression matrices were removed with the summation method [26].

2.4. Deconvolution of Cell Types with Cibersortx and Three Atlases

Three single-cell RNA-seq datasets were used in this study: (1) GSE115469 (Normal Liver), (2) GSE146409 (TME-Stroma), and (3) GSE156337 (TME-Immune). GSE115469 is a liver subset of the Human Cell Atlas [28]. GSE146409 contains the TME of liver cancer (HCC and CCA), including malignant cells [29]. GSE156337 contains the HCC microenvironment, without malignant cells [30]. This dataset identified high-quality immune cells.

All expression matrices were normalised to 10,000 counts/cell and packed into an H5AD file with authors' annotation as metadata for subsequent estimations. Both the signature matrix of scRNA-seq datasets and the expression matrix of bulk RNA-seq datasets were transformed into tab-delimited tables. The signature matrices of reference scRNA-seq were built with the Cibersortx protocol for "scRNA-seq" ("S"). Deconvolution was performed with the "Impute Cell Fractions" module. In validation experiments, batch correction was disabled in within-study validation and activated in "S" mode (with single-cell expression matrix as reference) in cross-study validation (two groups of validation experiments will be described below). In the estimation of real clinical data, batch correction was enabled in "S" mode for RNA-seq datasets and "B" mode (with single-cell expression matrix collapsed into bulk matrix before used as reference) for microarray datasets. All other Cibersortx parameters were the default configurations [20].

Cibersortx fails when the cell type tree of the reference atlas is complicated (e.g., >10 cell types). In this situation, collapsing some branches of the cell type tree would complete the calculation [20]. In our study, for example, when we used the Normal Liver atlas, the cell type tree was divided into two groups (immune and non-immune groups). Similarly, the TME-Stroma atlas was divided into three groups (mesenchymal, vasculature and immune groups), and the TME-Immune atlas was divided into two groups (immune and non-immune groups) (All cell type trees are shown in File S2). To evaluate the influence of adjusting cell type tree, we performed a group of validation experiments in which cell subtypes in the Normal Liver atlas were collapsed (for example, let macrophage = inflammatory macrophage + non-inflammatory macrophage, T cell = alpha-beta T cell + gamma-delta T cell, B cell = mature B cell + plasma B cell, and LSEC = periportal LSEC + central venous LSEC). To increase the matching in cross-study experiments (described below), the cell-type trees of the other two atlases were also adjusted (for TME-Stroma atlas, let macrophage = scar-associated macrophage + Kupffer cell + tissue macrophage, and for TME-Immune atlas, let T cell = CD4⁺ T cell + CD8⁺ T cell + T_{reg} cell).

2.5. Accuracy and Robustness of Cell Fraction Estimation

The accuracy and robustness of deconvolution with Cibersortx were tested in two groups of experiments: intra-study validation and cross-study validation. In intra-study validation experiments, the signature matrix of scRNA-seq and the pseudo-bulk for testing

were generated from the same scRNA-seq atlas. The advantage of this method is full coverage of all cell types/subtypes. In cross-study validation experiments, the signature matrix of scRNA-seq and the pseudo-bulk dataset were generated from different atlases.

Pseudo-bulk datasets were generated using the random module of NumPy in the following procedure: the expression of the scRNA-seq atlas was separated into two groups, cell type of interest and all the remaining cell type/subtypes collapsed into a single group labelled as “others”; 10% of cells in each group were selected using Choice function of NumPy, then two representative expression vectors (V) were generated by calculating the mean value of each gene; a random number f (between 0 and 100) was generated by Uniform function of NumPy; finally, the expression vector of the pseudo-bulk sample was generated by $V_{celltype} \times f + V_{others} \times (100 - f)$. f was used as benchmarking target.

Two parameters were used to evaluate the accuracy of the Cibersortx estimation: (1) Pearson correlation coefficient (PCC) between predefined proportions (f) and estimated fractions (f'); (2) mean absolute error ($MAE = 1/n \sum_{i=1}^n |f'_i - f_i|, i = 1, \dots, n$) and direction (overestimation or underestimation).

2.6. Survival analysis, Statistics, and Data Visualisation

To demonstrate the added value of our deconvolution analysis, we investigated the survival impact of estimated cell fractions on two HCC cohorts (TCGA-LIHC, GSE14520). The patient cohort was first ordered based on descending order of estimated cell fractions and then separated into high- and low-level groups. All separation possibilities (from 1:n-1 to n-1:1) were tested with log-rank tests. The one with the lowest P-value in log-rank tests was selected as the optimised separation. If all the P-values were above 0.05, the cohort was equally separated into two groups (median-point separation).

Survival lengths were transformed into months and observed events (labelled as “1”) were transformed into “True” (Boolean value, according to the requirement of Scikit-Survival [31]). Kaplan–Meier (K-M) curves were then fitted with Scikit-Survival [31] and plotted with the Step function of Matplotlib. The log-rank test was calculated with Lifelines [32]. All boxplots of this study were generated with Matplotlib.

Pathway activities were estimated by PROGENy [33]. GSEA was performed with GSEAp. The input gene list for GSEA was the marker genes selected by Cibersortx for the cell types. Our study shows a demonstration with “WikiPathway 2021—Human” as the reference. GSEAp allows dozens of different libraries, and we provide scripts in our GitHub repository.

For better reproducibility, all the analysis scripts including pre-processing have been shared with GitHub (https://github.com/ErasmusMC-Bioinformatics/OIO_Shaoshi, accessed date 27 December 2022). The supplements and pre-processed datasets were shared with Zenodo (<https://doi.org/10.5281/zenodo.7467268>, accessed date 27 December 2022).

3. Results

3.1. Cibersortx Estimation of Cell Fraction

In this study, we aim to decipher the TME of liver cancer by estimating the cell fractions in transcriptomes. This requires an accurate and robust model with well-annotated single-cell atlases. We adopted a state-of-the-art deconvolution algorithm (Cibersortx) and three scRNA-seq atlases (Normal Liver, TME-Stroma, and TME-Immune). These atlases describe more than 20 cell types or subtypes. Figure 1A outlines the workflow of this study, and the cell-type trees of the three atlases are outlined in File S2.

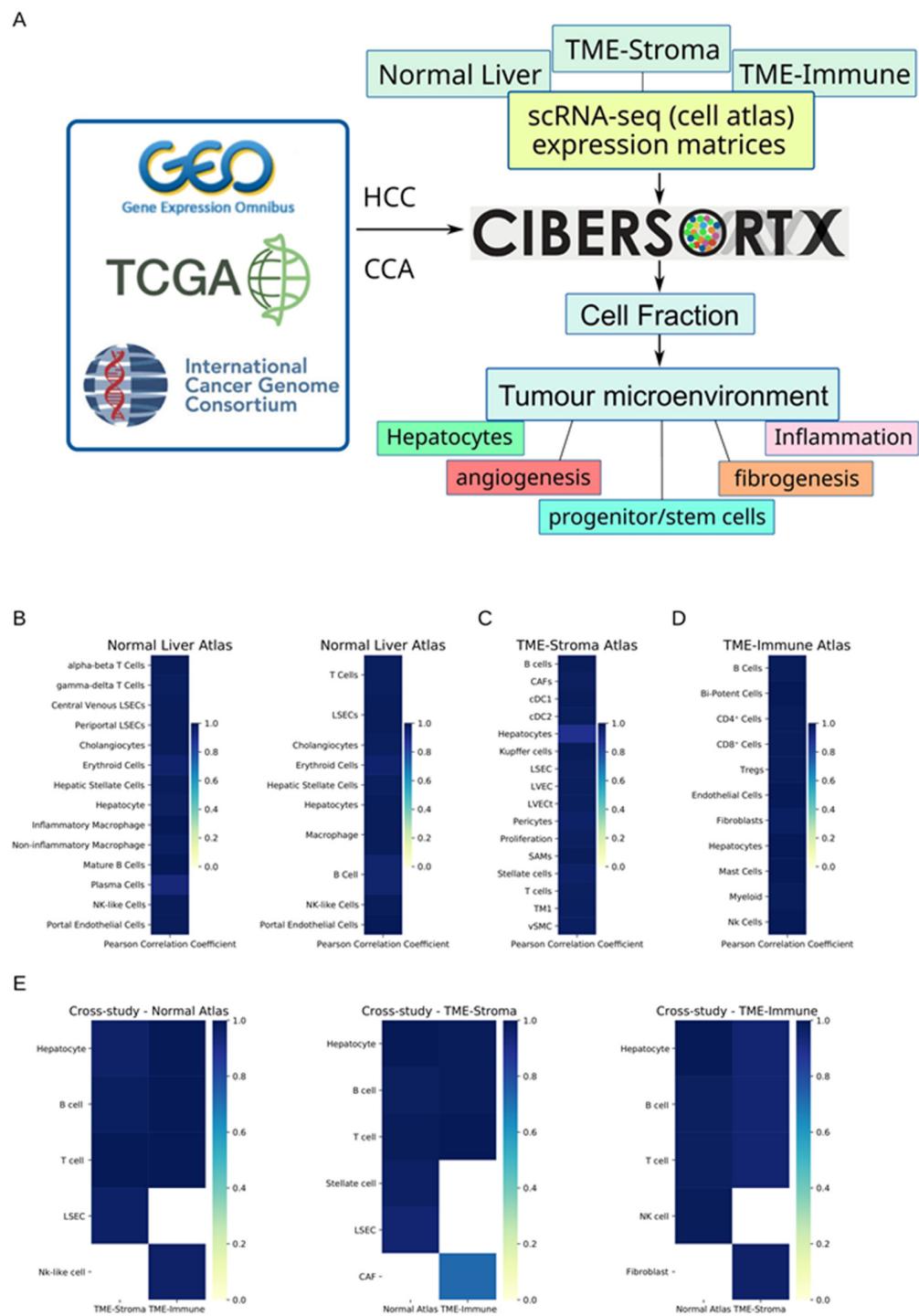


Figure 1. Study workflow and in silico validation of Cibersortx. (A) A general workflow of deciphering tumour microenvironment. Cibersortx estimates the cell fraction with scRNA-seq atlas and bulk RNA-seq/microarray data. Three expression matrices of scRNA-seq study were used as reference atlases. Through estimation of the fractions of more than 20 cell types/subtypes, biological events could be inferred. (B–D) Performances of Cibersortx deconvolution in intra-study validation experiments. (E) Performances of Cibersortx deconvolution in cross-study validation experiments. The title of each panel indicates which reference atlas was used in the deconvolution, two ticks on the bottom indicate which dataset was used to generate the pseudo-bulk samples.

We first performed intra-study validation experiments to test whether all identified cell types from each atlas could be accurately deconvoluted with Cibersortx. In this group of experiments, pseudo-bulk datasets were generated by the same atlas (Figure 1). In this validation mode, most cells achieved ideal PCCs (File S4). The PCCs compare the relationship between predefined cell fractions in generated pseudo-bulk samples and estimated values by Cibersortx. However, MAEs vary between different cell types (File S4). To evaluate the effect of merging some cell types, we did a group of tests by combining the major subtypes in the Normal Liver atlas (File S4). Both panels in Figure 1B show the result of intra-study validation with the Normal Liver atlas. Although both experiments show high levels of PCCs, subtle differences in accuracy exist. For example, after merging eight cell subtypes into four major cell types, the PCC of hepatocyte increases (0.979 vs. 0.992, File S4, the same the following), cholangiocyte decreases (0.9915 vs. 0.9844), and hepatic stellate cell (HSC) decreases (0.9901 vs. 0.986).

Although Cibersortx allows “partial deconvolution”, in which the samples may contain cell type/subtypes not present in the reference atlas, we performed a group of cross-study validation experiments to assess this impact. The results of these experiments demonstrate that the PCC between the presets and prediction remain high (Figure 1E, File S4, Page 19) whilst the MAEs vary significantly between cell types (File S4, Page 19). Our study suggests that Cibersortx normally guarantees high levels of PCCs but MAEs vary when using partial deconvolution method. Thus, we adopted a protocol in all subsequent analyses such that if a cell type (e.g., hepatocytes) was present in multiple atlases, the one with the best performances in both intra-study and cross-study validation experiments was chosen as the reference for clinical data.

3.2. Difference of Cell Fraction between Tumour and Non-Tumour Liver Tissue

To determine the difference in cell fraction between tumour and non-tumour tissues, we compared seven RNA-seq datasets which provide paired tissues collected from HCC/CCA cohort studies or case review studies. LIRI-JP is an RNA-seq study that includes primary liver cancers, and secondary liver cancers from stomach, colon, prostate, etc., with adjacent non-tumorous liver tissues as controls. GSE119336 is an RNA-seq study comparing CCA and non-tumour liver. The other five RNA-seq studies compare HCC and non-tumour liver tissues. Three of these studies included cases with HBV infection as the risk factor.

The three atlases provide more than 20 cell types, allowing a panoramic view of TMEs. All of these cell types can be largely divided into three groups: (1) cell types underpinning the fundamental physiology of livers, e.g., hepatocytes, cholangiocytes; (2) cell types participating in the pathological evolution of cancer formation, e.g., HSCs, LVECs; (3) immune cells. Results of related cell types are described in the following groups.

3.2.1. Hepatocytes and Cholangiocytes

These two cell types constitute the major functional units of livers—liver lobule and bile ducts [28]. As both the Normal Liver and TME-Immune atlases have hepatocytes, we made a comparison of results by deconvolution with two different atlases for the same dataset. Figure 2A,B show the results of deconvolution for the same datasets with different atlases. Compared with the results determined with the TME-immune atlas, deconvolution with the normal atlas returned higher levels of hepatocytes for some datasets (such as GSE119336) and lower levels for others (such as GSE94660). However, both experiments arrived at the same result whereby the decrease of hepatocytes and elevation of cholangiocytes are common in liver cancer. Similarly, File S3, pages 1–3 generated by deconvolution with TME-Stroma resulted in the same conclusion. Compared with HCC, the decrease of hepatocytes in CCA is more significant (Figure 2A,B).

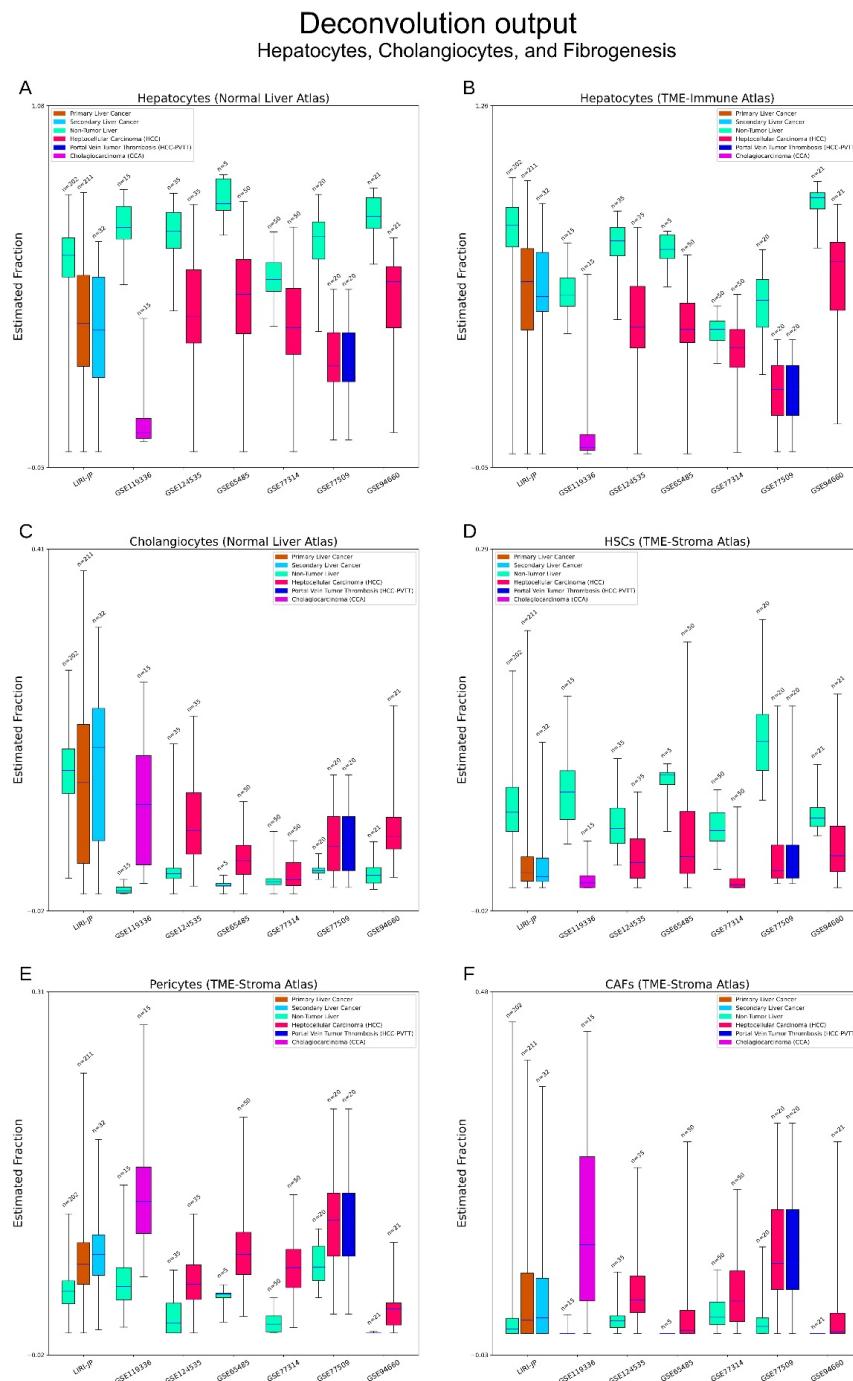


Figure 2. Deconvolution output—hepatocytes, cholangiocytes, and fibrogenesis. **(A,B)** A comparison of hepatocytes estimation by Cibersortx with two different atlases. Subtle differences of cell fractions can be seen between two experiments, but they arrived at the same conclusion, namely that hepatocytes decrease in cancerous tissue. In CCAs, this phenomenon is more prominent. **(C)** Estimation of cholangiocytes. Elevation of cholangiocytes can be widely seen in cancerous tissues. **(D–F)** Estimation of hepatic stellate cells (HSCs), pericytes and cancer-associated fibroblasts (CAFs). These three cell types are key components participating the fibrogenesis of liver cancer. HSCs are liver-specific pericytes, and often activated in liver cancer. Pericytes (broad sense) can be hardly detected in normal liver but are widely present in liver cancer. CAFs are uniquely characterised fibroblasts in cancer. They can hardly be detected in normal liver but are common in cancerous tissues.

3.2.2. Fibrogenesis

HSCs are tissue-specific equivalents of pericytes, and pericytes can be rarely detected in normal livers [29,34]. Fibroblasts proliferate following the activation of HSCs. Compared with non-tumorous tissues, a decrease in HSCs and the elevation of pericytes can be seen in liver cancers. CAFs are rarely detected in normal liver and are more often seen in liver cancers. The opposite alterations between HSCs and pericytes/CAFs suggest active fibrogenesis in tumours (Figure 2D–F).

3.2.3. Vasculature

Liver sinusoidal endothelial cells (LSECs) form the wall of hepatic sinusoids and participate in immune responses. Vascular smooth muscle cells (vSMCs) are also key components of blood vessels. Tumour liver vascular endothelial cells (LVECts) are variants of normal LVECs [29]. Decreased LSECs, vSMCs, and LVECs can be seen in cancerous tissues, while the elevation of LVECt can be observed in liver cancers (Figure 3A–D). These results suggest the emergence of abnormal angiogenesis.

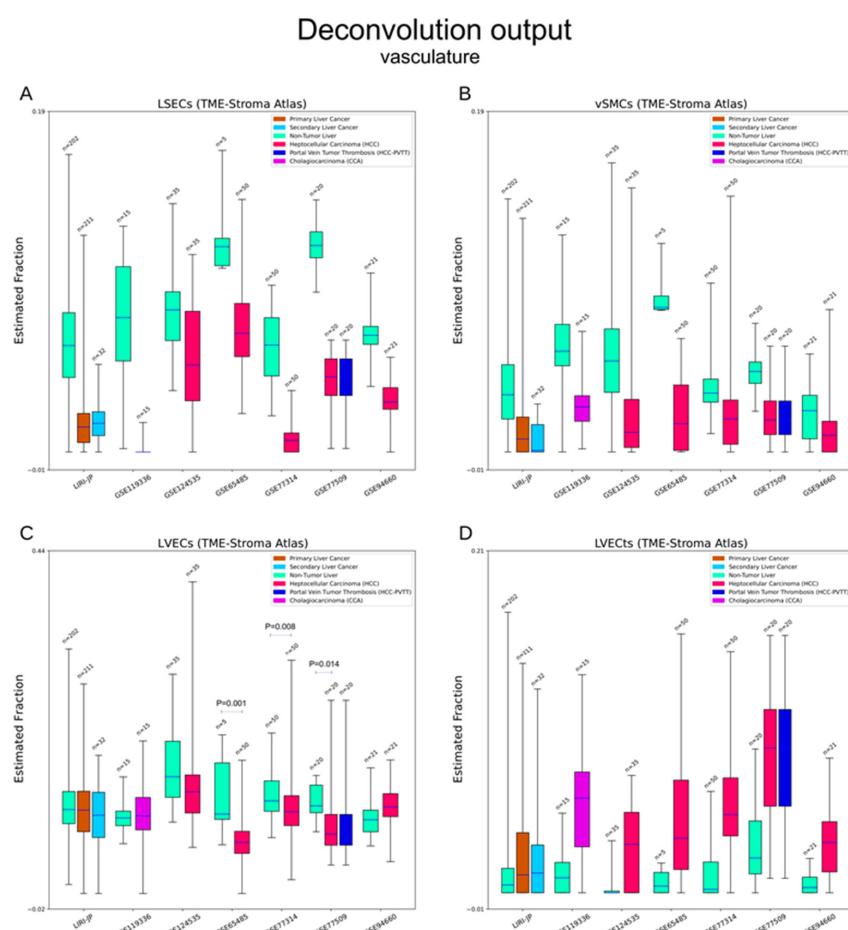


Figure 3. Deconvolution output—vasculature. (A–D) Estimation of liver sinusoidal endothelial cells (LSECs), vascular smooth muscle cells (vSMCs), liver vascular endothelial cells (LVECs), and tumour LVECs (LVECts). These four cell types/subtypes are key components of blood vessels in both normal and cancerous livers. Decreases of normal structural blocks (LSECs, vSMCs) and abnormal angiogenesis (LVECts) can be seen in liver cancers. Not all the differences are statistically significant. Pairs with $p < 0.05$ have been labelled.

3.2.4. Immune Cells

T cells: In the three atlases, five subsets of T cells were identified: T cells with alpha-beta ($\alpha\beta$) TCR chains or gamma-delta ($\gamma\delta$) chains (Normal Liver atlas) [28]; CD4⁺ (helper),

CD8⁺ (cytotoxic) T cells, and T_{reg} cells (TME-Immune atlas) [30]. We observed obvious elevations of $\alpha\beta$ T cells but no clear alterations of $\gamma\delta$ T cells in liver cancers. CD4⁺ T cells rise sharply in CCA while CD8⁺ T cells elevate moderately in HCC. Finally, T_{reg} cells are rarely detected in normal livers whereas elevations are common in liver cancers. T_{reg} cells have been recognised as a suppressor of tumour immune responses. Liver cancers also show high levels of overall T cells (CD4⁺+CD8⁺+T_{reg} cells). (Figure 4A–F).

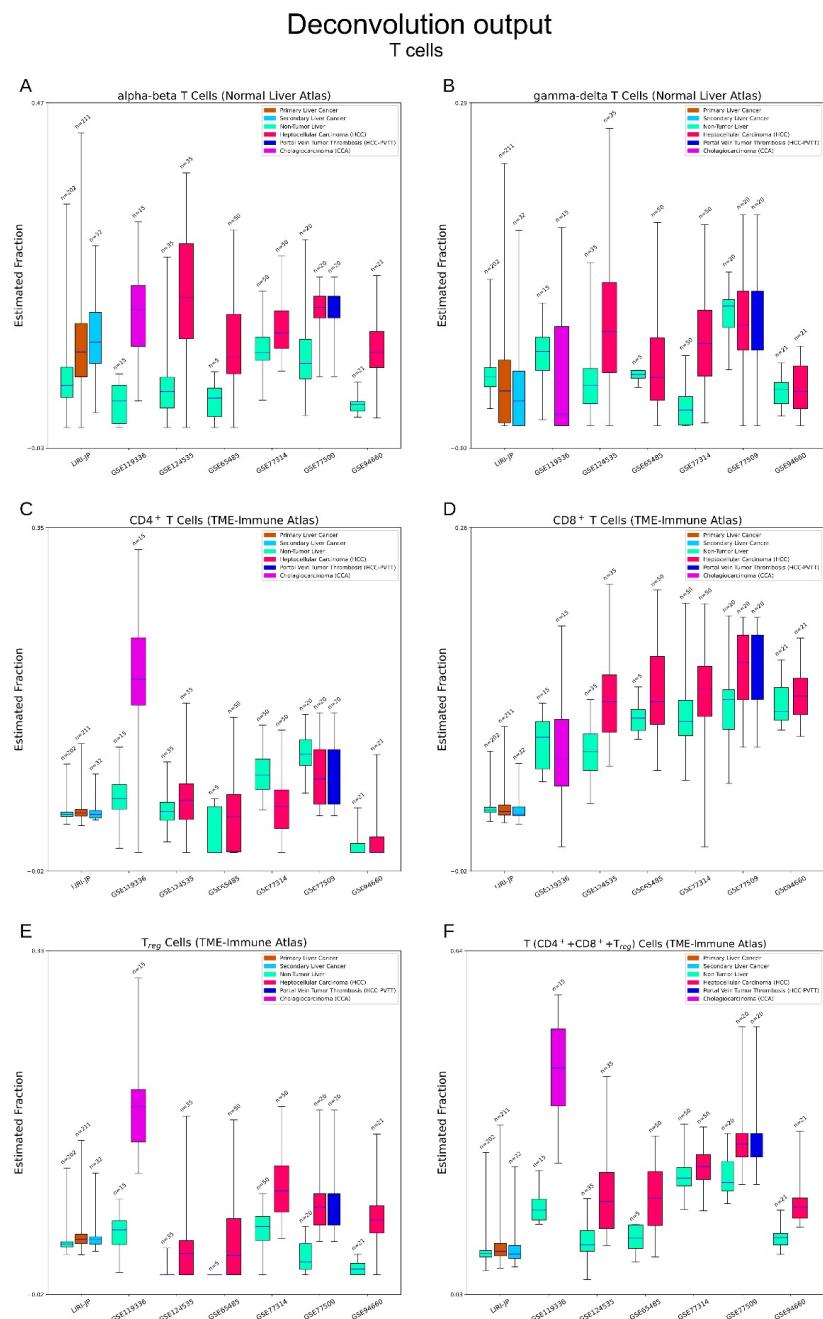


Figure 4. Deconvolution output—T cells. (A–E) Estimation of T cells. Alpha–beta ($\alpha\beta$) T cells and gamma–delta ($\gamma\delta$) T cells are from the Normal Liver atlas. CD4⁺/CD8⁺ T cells and regulatory T (T_{reg}) cells are from the TME-Immune atlas. Overlap of cell type trees may exist between the two atlases. Obvious elevations of $\alpha\beta$ T cells and obscure changes of $\gamma\delta$ T cells can be seen in liver cancers. CD4⁺ T cells rise prominently in CCAs and CD8⁺ T cells moderately in HCCs. (F) Estimation of overall T cells. Higher activities of T cells can be seen in liver cancers.

B cells: The Normal Liver atlas identified two subtypes of B cells, mature B cells (antigen inexperienced) and plasma B cells (antigen secreting) [28]. Variation of mature B cells do not show a direct association with liver cancers. Plasma B cells were rarely detected in normal liver but were widely detected in liver cancers (Figure 5A,B).

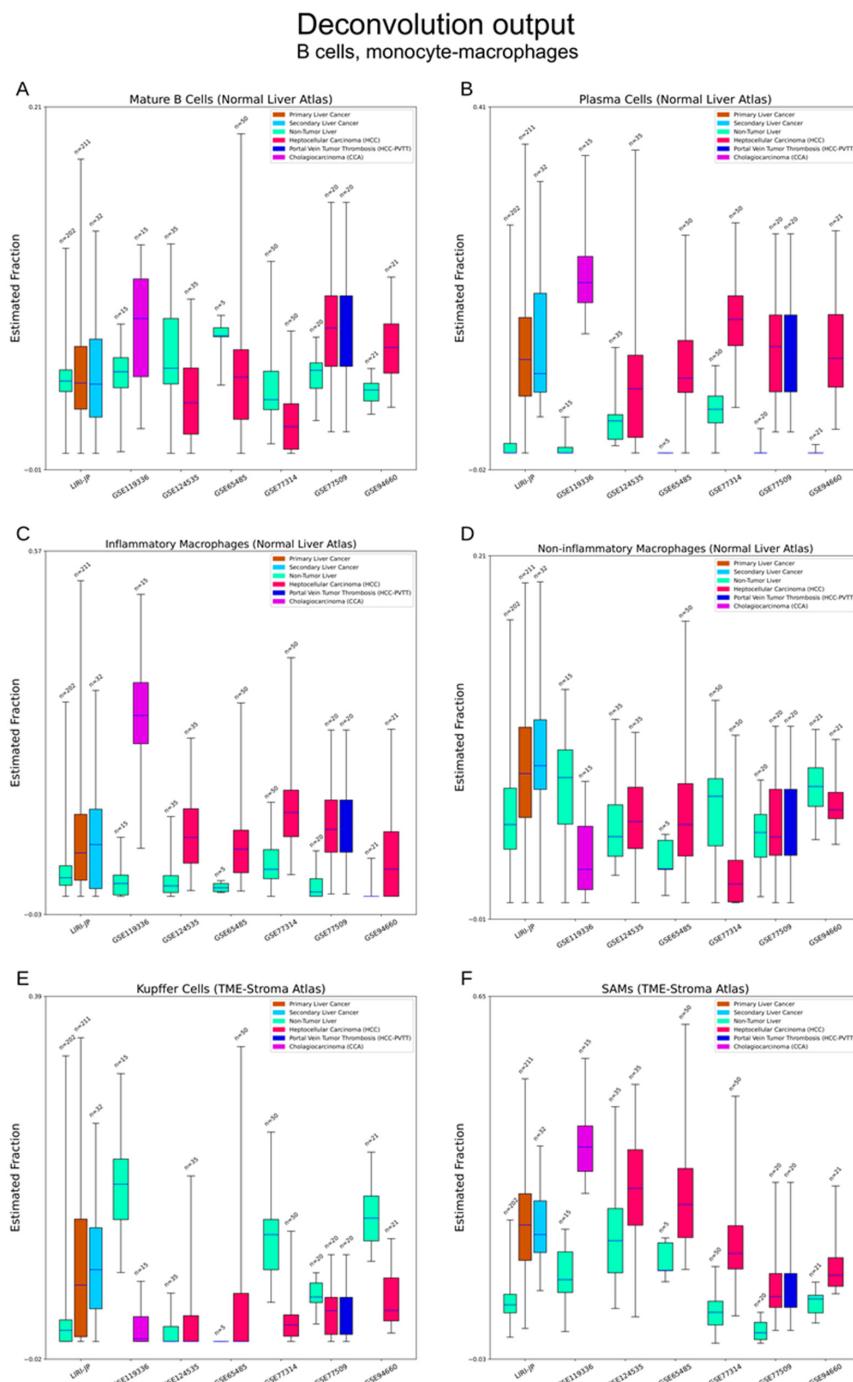


Figure 5. Deconvolution output—B cells, monocyte–macrophages. (A,B) Estimation of B cells. Plasma B cells can rarely be detected in normal livers but common in liver cancers. (C–F) Estimation of macrophages. The Normal Liver atlas separates macrophages into inflammatory and non-inflammatory subtypes. The TME-Stroma atlas identified Kupffer cells, scar-associated macrophages (SAMs) and tissue macrophages (TMs, Figure 6E). Overlap of cell type trees may exist between the two atlases. Increase of inflammatory macrophages can be seen in liver cancers. SAMs are a pathological variant of macrophages and participate in the process of liver fibrosis. Their elevation can be seen in liver cancers.

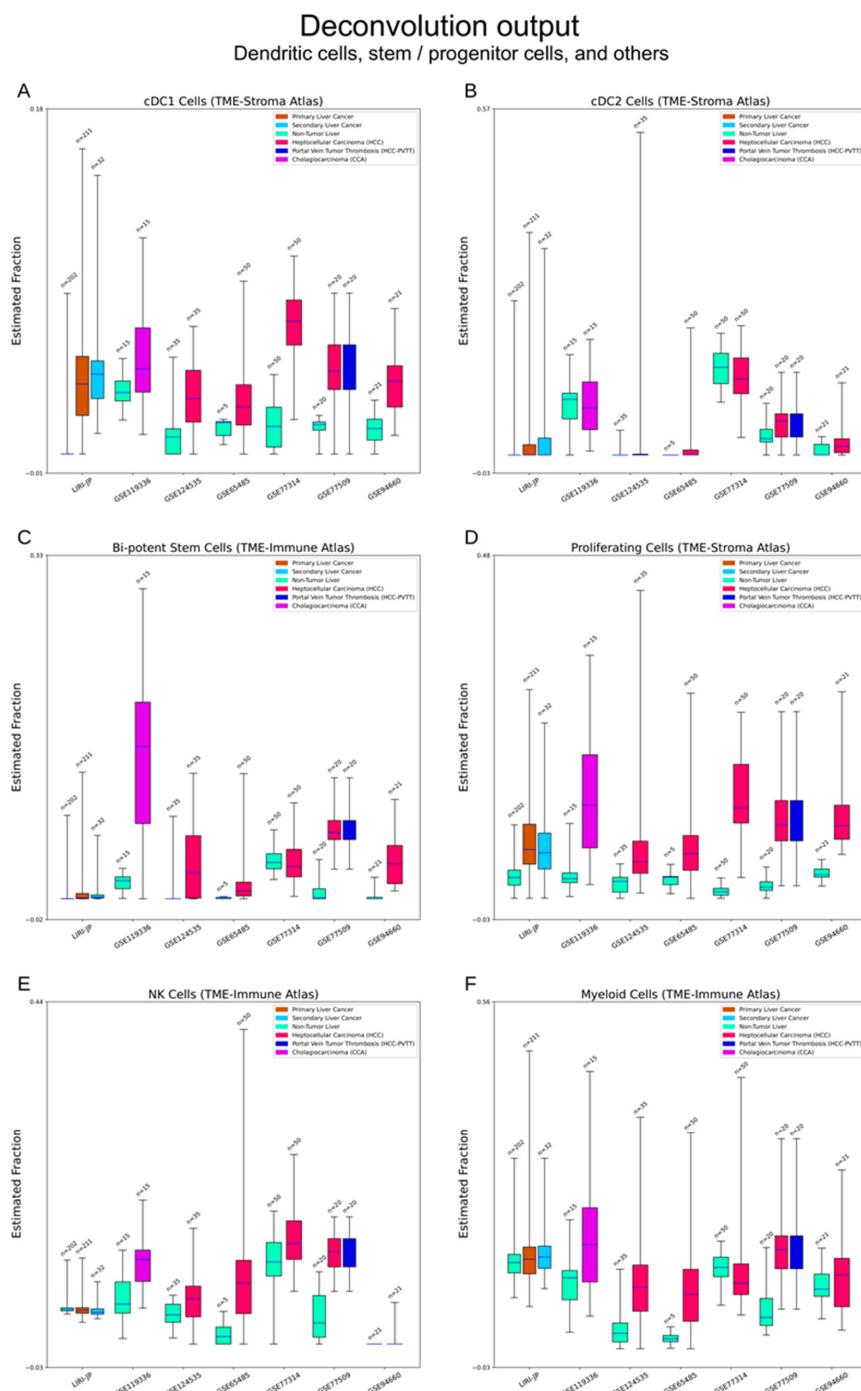


Figure 6. Deconvolution output—dendritic cells, stem/progenitor cells, and others. (A,B) Estimation of conventional dendritic cells (cDCs). Elevations of cDC1s can be seen in liver cancers. (C,D) Estimation of two types of proliferative cells. Bi-potent stem cells are a group of late-stage pluripotent cells with potential to differentiate into hepatocytes and cholangiocytes. The TME-Stroma atlas did not clarify the exact characteristics of proliferating cells. Its signature genes suggest its pluripotent origin. Elevations of these two cell types can be seen in liver cancers. (E,F) Estimation of natural killer (NK) cells and myeloid cells. NK cells belong to innate immune branch and myeloid cells are the hematopoiesis-originated immune branch. Deconvolution of these two cell types shows altered activities in liver cancers, but no direct association can be drawn.

Macrophages: The Normal Liver atlas contained inflammatory and non-inflammatory macrophages [28]. The TME-Stroma atlas separated Kupffer cells, tissue monocytes (TMs), and scar-associated macrophages (SAMs) [29]. SAMs are often recruited in the process of liver fibrosis [29,35]. We observed elevations of inflammatory macrophages and SAMs in liver cancers. No obvious differences in Kupffer cells, TMs and non-inflammatory macrophages were seen (Figure 5C–F).

Dendritic cells: As the quintessential antigen-presenting cell, the dendritic cell (DC) is another hot research interest for the potential of immunotherapy. The TME-Stroma atlas was used to estimate conventional DC1 and DC2 (cDC1 and cDC2) cell types [29,36]. Deconvolution suggests that cDC1 cells are elevated in liver cancers while cDC2 cells are not (Figure 6A,B).

In summary, liver cancers show higher levels of overall immune cells, involving both the innate (monocyte–macrophages) and adaptive branches (T, B cells), as well as auxiliary components (dendritic cells). Meanwhile, suppressive components such as T_{reg} cells can also be observed, suggesting the disordered responses in tumours.

3.2.5. Bi-Potent Stem Cells and Proliferative Cells

This group involves two cell types that can proliferate and differentiate. Bi-potent stem cells (from TME-Immune atlas) were named for their potential to differentiate into both hepatocytes and cholangiocytes [30]. Proliferating cells were identified in the TME-Stroma atlas and elevation of these two cell types (HCC and CCA) was common in tumours [28]. Bi-potent cells were rare in normal livers and their elevation in CCA is prominent (Figure 6C,D).

3.2.6. Other Cell Types

These three atlases also identified some immune cell types which exist in the liver with a relatively low abundance. TME-Immune identified a cluster of natural killer (NK) cells (a key component of the innate immune branch) and a cluster of myeloid cells (the liver-resident precursors of monocytes–macrophages) [30]. The Normal Liver atlas isolated a cluster of NK-like cells, which may be an ambiguous mixture of natural killer T (NKT) cells and NK cells [28]. Different atlases may have some cell types/subtypes with the same labels. However, calculated signature matrices suggest that they have different scopes, e.g., HSCs in the Normal Liver atlas, Figure 2D, vs. HSCs in the TME-Stroma atlas, File S3, Page 1).

3.3. Cell Fraction of HCC TME Correlates with Clinical Outcome

Finally, we investigated whether cellular alteration affects clinical outcome of HCC through survival analysis. In public repositories, TCGA-LIHC is the highest-cited cohort of a liver cancer study, with well-annotated follow-up information and substantial sample size (370 HCC patients). TCGA-LIHC is a pooled study of five cohorts with mixed risk factors. Available survival analyses include overall survival (OS) and disease-free survival (DFS) [37].

The distributions of estimated cell fractions show two typical shapes, “Sigmoid” or “Exponential” (File S5). Using an optimisation strategy (lowest log-rank test p-value), the patient cohort was typically separated at inflection points, although this separation may fail in cases of negative results or meaningless grouping (e.g., separating one case into a group). In these circumstances, we used the median-point strategy to finish complete KM curves (Figure 7, Files S5 and S6).

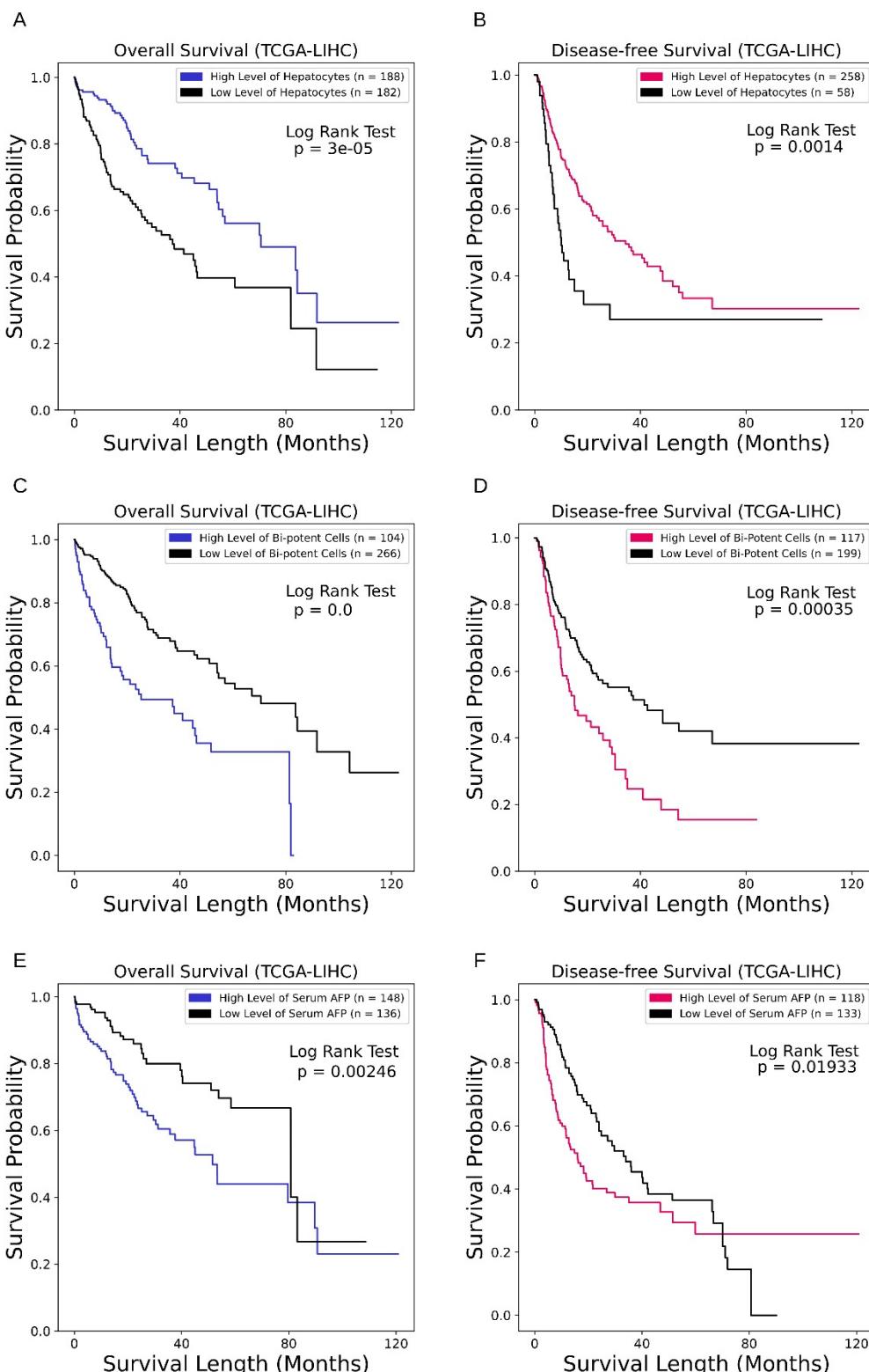


Figure 7. Impact of cell fractions on patients' survivals. (A,B) Impact of hepatocyte fraction. High-level hepatocytes show both longer OS and DFS lengths. (D-E) Impact of bi-potent cell fraction. High-level bi-potent cells show lower OS and DFS lengths. (C,F) Impact of serum AFP at patient admission.

Among all the estimated cell types, hepatocytes and bi-potent cells show substantial impacts on patients' outcomes. The estimated fractions of hepatocytes show a sigmoid-shaped distribution. The optimisation strategy (lowest log-rank test p-value) separates the cohort at a close-to-median point in OS analysis, and at an inflection point for DFS analysis (Figure 7C). High fractions of hepatocytes are associated with longer OS and DFS (Figure 7A,B). Estimated fractions of bi-potent cells show exponential distribution. The optimisation strategy isolated a subset with the cell fractions close to zero (Figure 7F). Those with high fractions of bi-potent cells show lower OS and DFS (Figure 7D,E). TCGA also included serum alpha-fetoprotein (AFP) at patient admission. Although the differences of OS and DFS between patients with high- or low-level AFP reach statistical significance, crossings of KM curves exist. In contrast, deconvoluted hepatocytes and bi-potent cells show better discrimination (Figure 7).

GSE14520 is a study that recruited more than 200 HBV-related HCC cases and provides both OS and DFS information [38,39]. Its transcriptomic tests are based on microarray platforms, which may provide less accuracy than RNA-seq. In our study, rare cell types were not often detected in the deconvolution of microarray data. However, given the subtle difference of study protocol, it still provided alternative evidence about the impact of cell fractions on patients' outcomes. A compilation of the survival analyses for GSE14520 is available in File S6, with a summary in Table S2. Consistent and significant results for both OS and DFS in the two cohorts include: hepatocyte (positive), cholangiocyte (negative), bi-potent stem cell (negative), mature B cell (positive), plasma B cell (negative), and T_{reg} cell (negative).

Pathway analysis provides useful information for the identification of therapeutic targets. Figure 8 shows the estimation of pathway activities by PROGENy [33]. Only a small cell population show high activities of specific pathways. For example, LVECt, proliferating cells, bi-potent cells, and mast cells are EGFR active. GSEA is another useful method. File S7 shows the examples using the signatures from the three atlases generated by Cibersortx with the library "WikiPathway 2021—Human" as the source of pathway definitions.

3.4. Cell Abundance Estimation by Support Vector Regression

We also conducted a parallel series of analyses by support vector regression. Supporting Information describes the detailed methodology. Both methods achieved the same conclusions for most cell types. Of note, a subtle difference exists between the prediction by SVR and Cibersortx deconvolution. SVR estimates abundance between samples while the summation of deconvoluted fractions is equal to one. A substantial alteration of mass cell types (most often hepatocytes in liver) may significantly impact trace components in deconvolution.

All Signature matrices and scripts have been shared online (see GitHub/Zenodo address).

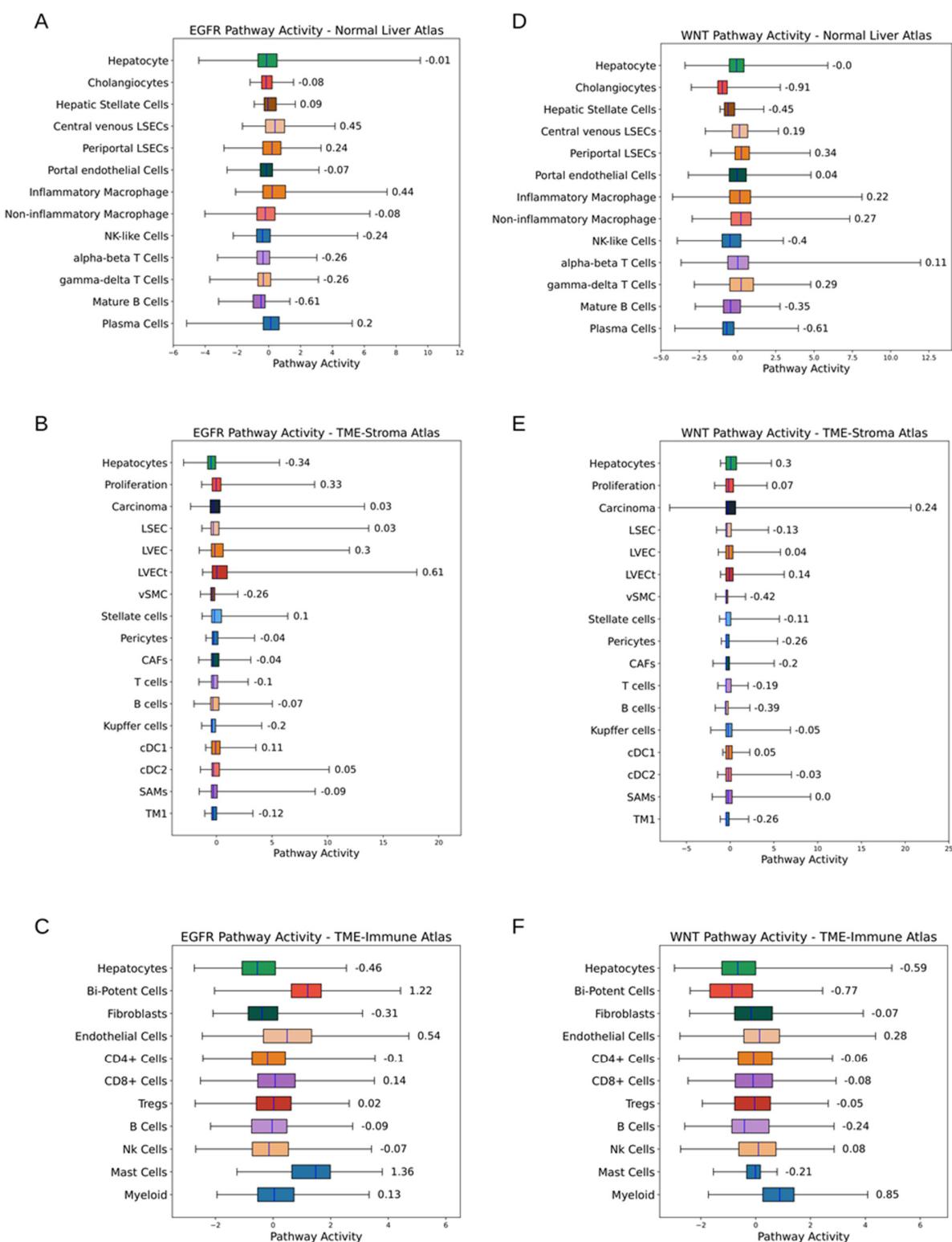


Figure 8. Pathway activities of cell types in the three atlases (estimated by PROGENy). (A–C) EGFR activities. Central venous LSECs, inflammatory macrophages, LVECt, bi-potent cells, mast cells, etc., are EGFR active, suggesting the involvement in angiogenesis, immune responses, and proliferation. (D–F) WNT activities. Periportal LSECs, inflammatory and non-inflammatory macrophages, gamma-delta T cells, hepatocytes (TME-Strom), carcinoma cells (TME-Immune), myeloid cells, etc., are WNT active, suggesting the involvement in immune activities.

4. Discussion

Deconvolution algorithms draw high interest in estimating cell fractions with scRNA-seq atlases. Cibersortx and MuSiC are two state-of-the-art algorithms in benchmarking studies [40,41]. We adopted Cibersortx for better reproducibility while implementation with MuSiC involves R scripting and manual selection of markers. In our study, in silico experiments determined that Cibersortx can accurately predict the fractions of cell types. We also performed a parallel series of analyses with SVR to reassure the conclusions by Cibersortx.

Cibersortx estimates cell fractions by quantifying the abundance of signature genes, which warrants careful consideration when interpreting the results. Cell fractions can be defined in diverse ways. In conventional histological studies, cell proportion calculates by volume, cell number, mass, etc., [42]. The prediction of Cibersortx seems to be close to the definition of “fraction by cell number”. However, the expression of signature genes varies between cells, tissues, individuals, and the different disease states, leading to an ambiguity in the concept of “fraction”. This gap becomes prominent when the biomedical conditions of the reference scRNA-seq atlases and the bulk RNA-seq samples differ (necessitating “partial deconvolution”). In addition, cell clusters between atlases with the same label may not be identical. In our study, we preserved all the signature matrices generated by Cibersortx for better comparison. Therefore, we recommend taking into account all of these factors when interpreting the biomedical implications of the deconvoluted results [20].

Three scRNA-seq atlases help portray biological events in liver cancers, including angiogenesis, fibrogenesis, immunity, and stem cell transformation. Agents targeting angiogenesis made the first breakthrough in liver cancer chemotherapy [43]. Cancer growth signifies abnormal blood supply in the TME, with simultaneous pseudo-hypoxia and neo-vasculature. The TME-Stroma atlas demonstrated the involvement of pericytes and LVECt in tumour angiogenesis [29]. As vessel components and immune barriers, LSECs lose their position in tumour tissues. The retreat of LSECs gives way to metastasis [44]. Our study shows that these alterations are common in liver cancers.

Liver carcinogenesis has close ties to chronic inflammations, either viral hepatitis or steatohepatitis. A normal liver resides liver-specific pericytes (also termed hepatic stellate cells), which can be inflammation-activated. Protracted inflammation leads to stellate cell transformation and fibroblast proliferation [45]. The latter interacts with malignant cells and helps a tumour-favourable microenvironment. Cancer-associated fibroblasts (CAFs) provide an immune-evading and chemotherapy-resistant barrier. They also secret multiple cytokines promoting tumour growth and angiogenesis [46]. Experimental evidence shows that CAF remodelling of extracellular matrix (ECM) helps the metastasis-promoting TME [47,48]. Thus, therapeutic strategies targeting CAFs are in development, such as nanocarriers [9]. Our study shows the universal attendance of CAFs in liver cancers, indicating the potential of CAF-targeting therapies.

Liver cancer evolves with the cross-talk between malignant cells and the immune system. In the early phase of cancer initiation, immune cells actively move towards and fight against transformed cells. Long-term engagement finally resulted in the exhaustion of anti-tumour immunity. Some of the immune cells serve the malignant transformation. Such components include well-established tumour-associated macrophages [49,50], exhaustive and immunosuppressive T cells [51,52], and gradually recognised tumour-infiltrating B cells [53]. Our study provides a helicopter view of the broad attendance of immune cells in tumour tissues and the appearance of unfavourable components such as T_{reg} cells. T_{reg} cells function through the PD-1/PD-L1 pathway, leading to tumour tolerance. A blockade of this communication results in the resurrection of immune responses in a minor group of patients. Further studies show that the therapeutic efficiency of a PD1 inhibitor depends on the interaction between the TME and other immune components (e.g., CD8⁺ T cells) [10,54].

The cancer stem cell (CSC) hypothesis proposed that a small cell population harbouring embryonic characteristics fuel tumour growth. It is difficult to identify CSCs except by tracing their descendants bearing specific features. Reported biomarkers of liver CSCs

include CD133, CD90, epithelial cell adhesion molecule (EpCAM), etc., [55]. EpCAM⁺ cells were proposed as a tumour-initiating component in HCC development [56]. EpCAM expresses in foetal livers, hepatic progenitor cells, carcinoma cells, etc., but not in mature hepatocytes [57]. In our study, EpCAM was selected by Cibersortx as a signature gene for cholangiocytes (Normal Liver atlas), proliferating cells (TME-Stroma), and bi-potent stem cells (TME-Immune), indicating the close relationship between these cell components and liver CSCs. Wnt–beta-catenin signalling activates EpCAM expression, which is associated with AFP elevation and foreshadows negative outcomes [58,59]. Pathway analysis by PR-GOGENy suggests that hepatocytes and carcinoma cells are WNT active in the TME-Stroma atlas. PROGENy suggests high EGFR activities of bi-potent stem cells (TME-Immune) and proliferating cells (TME-Stroma). EGFR is responsible for the maintenance of multiple CSC phenotypes [60]. Survival analysis suggests the negative impacts of bi-potent stem cells on patient outcomes, providing alternative clinical evidence for the tumour-initiating hypothesis of EpCAM⁺ cells. Although the interpretation of these findings warrants careful consideration, our study demonstrates that deconvolution can also help understand the mechanism of cancer formation.

5. Conclusions

In this study, we deciphered the TME of liver cancer by estimating the cell fractions of a sample given a transcriptome. By estimating more than 20 cell types/subtypes within bulk RNA-seq data using three atlases and Cibersortx, we found disruptions of normal liver architecture, abnormal fibrogenesis and angiogenesis, as well as disturbed immune responses in HCC and CCA. Survival analysis demonstrated that five cell types/subtypes highly correlated with patient outcomes.

Deconvolution algorithm and emerging scRNA-seq atlases allow the decomposition of bulk RNA-seq data into cell-type fractions. By linking the cell fractions of samples and clinical follow-up information, we provide an innovative approach for the discovery of potential therapeutic targets. In the future, with the advent of more high-quality scRNA-seq atlases, deconvolution could be a powerful data mining tool for uncovering the intricate nature of the TME of liver cancer and revealing valuable information in the vast amount of available transcriptomic data.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers15010153/s1>, File S1—Estimation by Support Vector Regression. File S2—Cell Type Hierarchy. File S3—Estimation by Cibersortx and Support Vector Regression—Other Cell Types. File S4—In Silico Validation of Cibersortx and Support Vector Regression. File S5—Survival Impacts of Cell Fractions Estimated by Cibersortx—TCGA-LIHC. File S6—Survival Impacts of Cell Fractions Estimated by Cibersortx—GSE14520. File S7—Pathway Analysis. Table S1—Summary of the Datasets Used in This Study. Table S2—Summary of Survival Impact. Refs. [41,61,62] cite in the Supporting Information file.

Author Contributions: Conceptualization, S.Z. and A.P.S.; methodology, S.Z. and A.P.S.; software, S.Z.; validation, S.Z.; formal analysis, S.Z.; investigation, S.Z. and A.P.S.; resources, S.Z. and A.P.S.; data curation, S.Z. and A.P.S.; writing—original draft preparation, S.Z. and A.P.S.; writing—review and editing, S.Z., W.B., F.v.K. and A.P.S.; visualization, S.Z.; supervision, M.P.P., F.v.K. and A.P.S.; project administration, A.P.S.; funding acquisition, M.P.P., F.v.K. and A.P.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was covered by the Human Subjects Protection and Data Access Policies. These are set of policies developed by NCI and NHGRI to protect the privacy of participants donating specimens to TCGA. Included are TCGA's informed consent policy, data access policy and information about HIPAA Privacy Rule compliance. January 2014.

Data Availability Statement: All data is publicly available. The code used to reproduce this work is available from our GitHub (https://github.com/ErasmusMC-Bioinformatics/OIO_Shaoshi, accessed on 27 December 2022). The supplements have been uploaded to Zenodo (<https://doi.org/10.5281/zendodo.7467268>, accessed on 27 December 2022).

Acknowledgments: The authors would like to thank Ju Jie, Willem de Koning, Teodora-Elena Trandafir, as well as all the Bioinformatics team, Erasmus MC.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef] [PubMed]
2. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer Statistics, 2020. *CA Cancer J. Clin.* **2020**, *70*, 7–30. [CrossRef] [PubMed]
3. Jemal, A.; Ward, E.M.; Johnson, C.J.; Cronin, K.A.; Ma, J.; Ryerson, A.B.; Mariotto, A.; Lake, A.J.; Wilson, R.; Sherman, R.L.; et al. Annual Report to the Nation on the Status of Cancer, 1975–2014, Featuring Survival. *J. Natl. Cancer Inst.* **2017**, *109*, djx030. [CrossRef] [PubMed]
4. Llovet, J.M.; Kelley, R.K.; Villanueva, A.; Singal, A.G.; Pikarsky, E.; Roayaie, S.; Lencioni, R.; Koike, K.; Zucman-Rossi, J.; Finn, R.S. Hepatocellular Carcinoma. *Nat. Rev. Dis. Prim.* **2021**, *7*, 6. [CrossRef] [PubMed]
5. Brindley, P.J.; Bachini, M.; Ilyas, S.I.; Khan, S.A.; Loukas, A.; Sirica, A.E.; Teh, B.T.; Wongkham, S.; Gores, G.J. Cholangiocarcinoma. *Nat. Rev. Dis. Prim.* **2021**, *7*, 65. [CrossRef] [PubMed]
6. Deng, G.L.; Zeng, S.; Shen, H. Chemotherapy and Target Therapy for Hepatocellular Carcinoma: New Advances and Challenges. *World J. Hepatol.* **2015**, *7*, 787–798. [CrossRef]
7. Morse, M.A.; Sun, W.; Kim, R.; He, A.R.; Abada, P.B.; Mynderse, M.; Finn, R.S. The Role of Angiogenesis in Hepatocellular Carcinoma. *Clin. Cancer Res.* **2019**, *25*, 912–920. [CrossRef]
8. Simone, V.; Brunetti, O.; Lupo, L.; Testini, M.; Maiorano, E.; Simone, M.; Longo, V.; Rolfo, C.; Peeters, M.; Scarpa, A.; et al. Targeting Angiogenesis in Biliary Tract Cancers: An Open Option. *Int. J. Mol. Sci.* **2017**, *18*, 418. [CrossRef]
9. Kaps, L.; Schuppan, D. Targeting Cancer Associated Fibroblasts in Liver Fibrosis and Liver Cancer Using Nanocarriers. *Cells* **2020**, *9*, 2027. [CrossRef]
10. Sangro, B.; Sarobe, P.; Hervás-Stubbs, S.; Melero, I. Advances in Immunotherapy for Hepatocellular Carcinoma. *Nat. Rev. Gastroenterol. Hepatol.* **2021**, *18*, 525–543. [CrossRef]
11. Fiste, O.; Ntanasis-Stathopoulos, I.; Gavriatopoulou, M.; Lontos, M.; Koutsoukos, K.; Dimopoulos, M.A.; Zagouri, F. The Emerging Role of Immunotherapy in Intrahepatic Cholangiocarcinoma. *Vaccines* **2021**, *9*, 422. [CrossRef] [PubMed]
12. Jin, M.Z.; Jin, W.L. The Updated Landscape of Tumor Microenvironment and Drug Repurposing. *Signal Transduct. Target. Ther.* **2020**, *5*, 166. [CrossRef] [PubMed]
13. Binnewies, M.; Roberts, E.W.; Kersten, K.; Chan, V.; Fearon, D.F.; Merad, M.; Coussens, L.M.; Gabrilovich, D.I.; Ostrand-Rosenberg, S.; Hedrick, C.C.; et al. Understanding the Tumor Immune Microenvironment (TIME) for Effective Therapy. *Nat. Med.* **2018**, *24*, 541–550. [CrossRef] [PubMed]
14. Zhang, Y.; Wang, D.; Peng, M.; Tang, L.; Ouyang, J.; Xiong, F.; Guo, C.; Tang, Y.; Zhou, Y.; Liao, Q.; et al. Single-cell RNA Sequencing in Cancer Research. *J. Exp. Clin. Cancer Res.* **2021**, *40*, 81. [CrossRef] [PubMed]
15. Cooper, L.A.D.; Demicco, E.G.; Saltz, J.H.; Powell, R.T.; Rao, A.; Lazar, A.J. PanCancer Insights from The Cancer Genome Atlas: The Pathologist’s Perspective. *J. Pathol.* **2018**, *244*, 512–524. [CrossRef]
16. Malta, T.M.; Sokolov, A.; Gentles, A.J.; Burzykowski, T.; Poisson, L.; Weinstein, J.N.; Kamińska, B.; Huelsken, J.; Omberg, L.; Gevaert, O.; et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* **2018**, *173*, 338–354.e15. [CrossRef]
17. Chiu, Y.J.; Hsieh, Y.H.; Huang, Y.H. Improved Cell Composition Deconvolution Method of Bulk Gene Expression Profiles to Quantify Subsets of Immune Cells. *BMC Med. Genomics* **2019**, *12*, 169. [CrossRef]
18. Tsoucas, D.; Dong, R.; Chen, H.; Zhu, Q.; Guo, G.; Yuan, G.C. Accurate Estimation of Cell-Type Composition from Gene Expression Data. *Nat. Commun.* **2019**, *10*, 2975. [CrossRef]
19. Wang, X.; Park, J.; Susztak, K.; Zhang, N.R.; Li, M. Bulk Tissue Cell Type Deconvolution with Multi-Subject Single-Cell Expression Reference. *Nat. Commun.* **2019**, *10*, 380. [CrossRef]
20. Newman, A.M.; Steen, C.B.; Liu, C.L.; Gentles, A.J.; Chaudhuri, A.A.; Scherer, F.; Khodadoust, M.S.; Esfahani, M.S.; Luca, B.A.; Steiner, D.; et al. Determining Cell Type Abundance and Expression from Bulk Tissues with Digital Cytometry. *Nat. Biotechnol.* **2019**, *37*, 773–782. [CrossRef]
21. Qi, Z.; Liu, Y.; Mints, M.; Mullins, R.; Sample, R.; Law, T.; Barrett, T.; Mazul, A.L.; Jackson, R.S.; Kang, S.Y.; et al. Single-Cell Deconvolution of Head and Neck Squamous Cell Carcinoma. *Cancers (Basel)* **2021**, *13*, 1230. [CrossRef] [PubMed]
22. 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. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef] [PubMed]

23. Piccolo, S.R.; Sun, Y.; Campbell, J.D.; Lenburg, M.E.; Bild, A.H.; Johnson, W.E. A Single-Sample Microarray Normalization Method to Facilitate Personalized-Medicine Workflows. *Genomics* **2012**, *100*, 337–344. [CrossRef] [PubMed]
24. Piccolo, S.R.; Withers, M.R.; Francis, O.E.; Bild, A.H.; Johnson, W.E. Multiplatform Single-Sample Estimates of Transcriptional Activation. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17778–17783. [CrossRef]
25. 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.* **2009**, *4*, 1184–1191. [CrossRef]
26. Miller, J.A.; Cai, C.; Langfelder, P.; Geschwind, D.H.; Kurian, S.M.; Salomon, D.R.; Horvath, S. Strategies for Aggregating Gene Expression Data: The CollapseRows R Function. *BMC Bioinform.* **2011**, *12*, 322. [CrossRef]
27. Wang, S.; Liu, X. The UCSCXenaTools R Package: A Toolkit for Accessing Genomics Data from UCSC Xena Platform, from Cancer Multi-Omics to Single-Cell RNA-Seq. *J. Open Source Softw.* **2019**, *4*, 1627. [CrossRef]
28. MacParland, S.A.; Liu, J.C.; Ma, X.Z.; Innes, B.T.; Bartczak, A.M.; Gage, B.K.; Manuel, J.; Khuu, N.; Echeverri, J.; Linares, I.; et al. Single Cell RNA Sequencing of Human Liver Reveals Distinct Intrahepatic Macrophage Populations. *Nat. Commun.* **2018**, *9*, 4383. [CrossRef]
29. Massalha, H.; Bahar Halpern, K.; Abu-Gazala, S.; Jana, T.; Massasa, E.E.; Moor, A.E.; Buchauer, L.; Rozenberg, M.; Pikarsky, E.; Amit, I.; et al. A Single Cell Atlas of the Human Liver Tumor Microenvironment. *Mol. Syst. Biol.* **2020**, *16*, e9682. [CrossRef]
30. Sharma, A.; Seow, J.J.W.; Dutertre, C.A.; Pai, R.; Blériot, C.; Mishra, A.; Wong, R.M.M.; Singh, G.S.N.; Sudhagar, S.; Khalilnezhad, S.; et al. Onco-Fetal Reprogramming of Endothelial Cells Drives Immunosuppressive Macrophages in Hepatocellular Carcinoma. *Cell* **2020**, *183*, 377–394.e21. [CrossRef]
31. Pölsterl, S. Scikit-Survival: A Library for Time-to-Event Analysis Built on Top of Scikit-Learn. *J. Mach. Learn. Res.* **2020**, *21*, 1–6.
32. Davidson-Pilon, C. Lifelines: Survival Analysis in Python. *J. Open Source Softw.* **2019**, *4*, 1317. [CrossRef]
33. Schubert, M.; Klänger, B.; Klünemann, M.; Sieber, A.; Uhlitz, F.; Sauer, S.; Garnett, M.J.; Blüthgen, N.; Saez-Rodriguez, J. Perturbation-Response Genes Reveal Signaling Footprints in Cancer Gene Expression. *Nat. Commun.* **2018**, *9*, 20. [CrossRef] [PubMed]
34. June, S.L.; Semela, D.; Iredale, J.; Shah, V.H. Sinusoidal Remodeling and Angiogenesis: A New Function for the Liver-Specific Pericyte? *Hepatology* **2007**, *45*, 817–825. [CrossRef]
35. Tacke, F.; Zimmermann, H.W. Macrophage Heterogeneity in Liver Injury and Fibrosis. *J. Hepatol.* **2014**, *60*, 1090–1096. [CrossRef]
36. Lurje, I.; Hammerich, L.; Tacke, F. Dendritic Cell and T Cell Crosstalk in Liver Fibrogenesis and Hepatocarcinogenesis: Implications for Prevention and Therapy of Liver Cancer. *Int. J. Mol. Sci.* **2020**, *21*, 7378. [CrossRef]
37. Ally, A.; Balasundaram, M.; Carlsen, R.; Chuah, E.; Clarke, A.; Dhalla, N.; Holt, R.A.; Jones, S.J.M.; Lee, D.; Ma, Y.; et al. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* **2017**, *169*, 1327–1341.e23. [CrossRef]
38. Roessler, S.; Jia, H.L.; Budhu, A.; Forgues, M.; Ye, Q.H.; Lee, J.S.; Thorgeirsson, S.S.; Sun, Z.; Tang, Z.Y.; Qin, L.X.; et al. A Unique Metastasis Gene Signature Enables Prediction of Tumor Relapse in Early-Stage Hepatocellular Carcinoma Patients. *Cancer Res.* **2010**, *70*, 10202–10212. [CrossRef]
39. Roessler, S.; Long, E.L.; Budhu, A.; Chen, Y.; Zhao, X.; Ji, J.; Walker, R.; Jia, H.L.; Ye, Q.H.; Qin, L.X.; et al. Integrative Genomic Identification of Genes on 8p Associated with Hepatocellular Carcinoma Progression and Patient Survival. *Gastroenterology* **2012**, *142*, 957–966.e12. [CrossRef]
40. Jin, H.; Liu, Z. A Benchmark for RNA-Seq Deconvolution Analysis under Dynamic Testing Environments. *Genome Biol.* **2021**, *22*, 102. [CrossRef]
41. Avila Cobos, F.; Alquicira-Hernandez, J.; Powell, J.E.; Mestdagh, P.; De Preter, K. Benchmarking of Cell Type Deconvolution Pipelines for Transcriptomics Data. *Nat. Commun.* **2020**, *11*, 5650. [CrossRef] [PubMed]
42. Blouin, A.; Bolender, R.P.; Weibel, E.R. Distribution of Organelles and Membranes between Hepatocytes and Nonhepatocytes in the Rat Liver Parenchyma. A Stereological Study. *J. Cell Biol.* **1977**, *72*, 441–455. [CrossRef] [PubMed]
43. Sampat, K.R.; O’Neil, B. Antiangiogenic Therapies for Advanced Hepatocellular Carcinoma. *Oncologist* **2013**, *18*, 430–438. [CrossRef]
44. Yang, M.; Zhang, C. The Role of Liver Sinusoidal Endothelial Cells in Cancer Liver Metastasis. *Am. J. Cancer Res.* **2021**, *11*, 1845–1860. [PubMed]
45. Tsuchida, T.; Friedman, S.L. Mechanisms of Hepatic Stellate Cell Activation. *Nat. Rev. Gastroenterol. Hepatol.* **2017**, *14*, 397–411. [CrossRef]
46. Affo, S.; Yu, L.X.; Schwabe, R.F. The Role of Cancer-Associated Fibroblasts and Fibrosis in Liver Cancer. *Annu. Rev. Pathol. Mech. Dis.* **2017**, *12*, 153–186. [CrossRef]
47. Sahai, E.; Astsaturov, I.; Cukierman, E.; DeNardo, D.G.; Egeblad, M.; Evans, R.M.; Fearon, D.; Greten, F.R.; Hingorani, S.R.; Hunter, T.; et al. A Framework for Advancing Our Understanding of Cancer-Associated Fibroblasts. *Nat. Rev. Cancer* **2020**, *20*, 174–186. [CrossRef]
48. Kubo, N.; Araki, K.; Kuwano, H.; Shirabe, K. Cancer-Associated Fibroblasts in Hepatocellular Carcinoma. *World J. Gastroenterol.* **2016**, *22*, 6841–6850. [CrossRef]
49. Pittet, M.J.; Michielin, O.; Migliorini, D. Clinical Relevance of Tumour-Associated Macrophages. *Nat. Rev. Clin. Oncol.* **2022**, *19*, 402–421. [CrossRef]
50. Noy, R.; Pollard, J.W. Tumor-Associated Macrophages: From Mechanisms to Therapy. *Immunity* **2014**, *41*, 49–61. [CrossRef]
51. Tanaka, A.; Sakaguchi, S. Regulatory T Cells in Cancer Immunotherapy. *Cell Res.* **2017**, *27*, 109–118. [CrossRef] [PubMed]

52. Takeuchi, Y.; Nishikawa, H. *International Immunology*; Oxford University Press: Oxford, UK, 2016; pp. 401–409.
53. Laumont, C.M.; Banville, A.C.; Gilardi, M.; Hollern, D.P.; Nelson, B.H. Tumour-Infiltrating B Cells: Immunological Mechanisms, Clinical Impact and Therapeutic Opportunities. *Nat. Rev. Cancer* **2022**, *22*, 414–430. [[CrossRef](#)] [[PubMed](#)]
54. Plitas, G.; Rudensky, A.Y. Regulatory T Cells in Cancer. *Annu. Rev. Cancer Biol.* **2020**, *4*, 459–477. [[CrossRef](#)]
55. Lee, T.K.W.; Guan, X.Y.; Ma, S. Cancer Stem Cells in Hepatocellular Carcinoma—from Origin to Clinical Implications. *Nat. Rev. Gastroenterol. Hepatol.* **2022**, *19*, 26–44. [[CrossRef](#)] [[PubMed](#)]
56. Yamashita, T.; Ji, J.; Budhu, A.; Forgues, M.; Yang, W.; Wang, H.Y.; Jia, H.; Ye, Q.; Qin, L.X.; Wauthier, E.; et al. EpCAM-Positive Hepatocellular Carcinoma Cells Are Tumor-Initiating Cells With Stem/Progenitor Cell Features. *Gastroenterology* **2009**, *136*, 1012–1024.e4. [[CrossRef](#)]
57. Yoon, S.M.; Gerasimidou, D.; Kuwahara, R.; Hytioglou, P.; Yoo, J.E.; Park, Y.N.; Theise, N.D. Epithelial Cell Adhesion Molecule (EpCAM) Marks Hepatocytes Newly Derived from Stem/Progenitor Cells in Humans. *Hepatology* **2011**, *53*, 964–973. [[CrossRef](#)]
58. Yamashita, T.; Budhu, A.; Forgues, M.; Xin, W.W. Activation of Hepatic Stem Cell Marker EpCAM by Wnt- β -Catenin Signaling in Hepatocellular Carcinoma. *Cancer Res.* **2007**, *67*, 10831–10839. [[CrossRef](#)] [[PubMed](#)]
59. Yamashita, T.; Forgues, M.; Wang, W.; Jin, W.K.; Ye, Q.; Jia, H.; Budhu, A.; Zanetti, K.A.; Chen, Y.; Qin, L.X.; et al. EpCAM and α -Fetoprotein Expression Defines Novel Prognostic Subtypes of Hepatocellular Carcinoma. *Cancer Res.* **2008**, *68*, 1451–1461. [[CrossRef](#)]
60. Talukdar, S.; Emdad, L.; Das, S.K.; Fisher, P.B. EGFR: An Essential Receptor Tyrosine Kinase-Regulator of Cancer Stem Cells. *Adv. Cancer Res.* **2020**, *147*, 161–188. [[CrossRef](#)]
61. Robinson, M.D.; Oshlack, A. A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data. *Genome Biol.* **2010**, *11*, R25. [[CrossRef](#)]
62. Stuart, T.; Butler, A.; Hoffman, P.; Hafemeister, C.; Papalexi, E.; Mauck, W.M.; Hao, Y.; Stoeckius, M.; Smibert, P.; Satija, R. Comprehensive Integration of Single-Cell Data. *Cell* **2019**, *177*, 1888–1902.e21. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.