
# Single-cell genomics and spatial transcriptomics: discovery of novel cell states and cellular interactions in liver physiology and disease biology

Antonio Saviano, Neil C. Henderson, Thomas F. Baumert

PII: S0168-8278(20)30372-X

DOI: https://doi.org/10.1016/j.jhep.2020.06.004

Reference: JHEPAT 7794

To appear in: *Journal of Hepatology*

Received Date: 31 January 2020

Revised Date: 28 May 2020

Accepted Date: 2 June 2020

Please cite this article as: Saviano A, Henderson NC, Baumert TF, Single-cell genomics and spatial transcriptomics: discovery of novel cell states and cellular interactions in liver physiology and disease biology, *Journal of Hepatology* (2020), doi: https://doi.org/10.1016/j.jhep.2020.06.004.

# Single-cell genomics and spatial transcriptomics: discovery of novel cell states and cellular interactions in liver physiology and disease biology

Antonio Saviano[1,2,3], Neil C. Henderson[4,5§], Thomas F. Baumert[1,2,3,6§]

[1]Inserm, U1110, Institut de Recherche sur les Maladies Virales et Hépatiques, Strasbourg, France

[2]Université de Strasbourg, Strasbourg, France

[3]Pôle Hépato-digestif, Institut-Hospitalo-Universitaire, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

[4]Centre for Inflammation Research, University of Edinburgh, Edinburgh EH16 4TJ, UK

[5]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK

[6]Institut Universitaire de France, Paris, France.

[§]Correspondence should be addressed to:

Prof. Thomas F. Baumert, MD, Inserm U1110, Institut de Recherche sur les Maladies Virales et Hépatiques, Université de Strasbourg, 3 rue Koeberlé, 67000 Strasbourg, France. Email: thomas.baumert@unistra.fr

Prof. Neil C. Henderson, MD, PhD, Centre for Inflammation Research, The Queen's Medical Research Institute, University of Edinburgh, Edinburgh EH16 4TJ, UK. Email: Neil.Henderson@ed.ac.uk

**Keywords:** single-cell, single-cell RNA sequencing, spatial transcriptomics, zonation, liver diseases, hepatocellular carcinoma, microenvironment, cirrhosis, fibrosis, non-parenchymal cells.

**Word count:** summary 198 words, main text only 4998 words.

**Figures:** 7

**Tables:** 1

**Authors' contribution:** AS performed literature research and data collection and analysis, designed the figures and wrote the original draft. AS, NCH, TFB reviewed and wrote the manuscript. NCH and TFB supervised the research.

**Financial support:** A.S. is the recipient of a fellowship of Region-Alsace, IHU and LabEx HEPSYS. N.C.H. acknowledges the support of the Wellcome Trust (Senior Research Fellowship in Clinical Science, ref. 103749), Medical Research Council and the Chan Zuckerberg Initiative. T.F.B. acknowledges support by the European Union (ERC-AdG2014 HEPCIR #671231, H2020 HEPCAR #667273, ERC PoC-HEPCAN #862551), Fondation ARC Paris and IHU Strasbourg (TheraHCC2.0, IHU201901299), the Foundation of the University of Strasbourg and Roche Institute (HEPKIN), the Agence Nationale de Recherches sur le Sida (ANRS) and the US National Institutes of Health (R21CA209940, R03AI131066, R01CA233794, U19AI12386). This work has been published under the framework of the LABEX ANR-10-LABX-0028_HEPSYS and PLAN CANCER 2014-2019 HCCMICTAR and benefits from a funding from the state managed by the French National Research Agency as part of the Investments for the Future Program, INCa (National Institute for Cancer) and Inserm.

The authors do not declare any conflict of interest.

## KEY POINTS

- Single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics are revolutionary techniques which allow the study of liver cell composition, physiology and disease development in unprecedented detail.
- ScRNA-seq comprises multiple technologies and the choice of platform used should be guided by the biological question, the study design and endpoints required.
- Gathering spatial information from single-cell data is challenging and several sequencing strategies and computational frameworks have been developed to overcome this issue.
- ScRNA-seq has uncovered substantial functional heterogeneity within the main liver cell lineages in health and disease, identifying zonation of multiple lineages across the liver lobule, and identification of novel progenitor populations.
- Liver zonation is not restricted to hepatocytes, but it is extended to non-parenchymal cells such as liver sinusoidal endothelial cells and stellate cells.
- ScRNA-seq of cirrhotic liver samples has allowed investigation of the cellular interactome regulating the human liver fibrotic niche.
- Notch signaling is a central pathway involved in cell interactions in the human liver fibrotic niche.
- ScRNA-seq has uncovered cellular heterogeneity within the tumor microenvironment of primary liver cancers.

## SUMMARY

Transcriptome analysis allows the study of gene expression of human tissues and it is a valuable tool to characterize liver function, gene expression changes during liver disease, identify prognostic markers or signatures, and to facilitate discovery of new therapeutic targets. In contrast to whole tissue RNA sequencing analysis, single-cell RNA-sequencing (scRNA-seq) and spatial transcriptomics enables the study of transcriptional activity at the single cell or spatial level. ScRNA-seq has paved the way to the discovery of previously unknown cell types and subtypes in normal and diseased liver, the study of rare cells such as liver progenitor cells as well as the functional role of non-parenchymal cells in chronic liver

disease and cancer. By adding spatial information to scRNA-seq data, spatial transcriptomics transforms understanding of tissue functional organization and cell-to-cell interactions in their native environment. These approaches have recently been applied to investigate liver regeneration, organization and division of labor of hepatocytes and non-parenchymal cells, and to profile the single cell landscape of chronic liver diseases and cancer. Here we review the principles and technologies behind scRNA-seq and spatial transcriptomics approaches, highlighting the recent discoveries and novel insights these methodologies have yielded in both liver physiology and disease biology.

## INTRODUCTION

Sequencing technologies are increasingly used to study phenotypes and drivers of liver disease. Whole tissue RNA sequencing has been primarily used to identify major differences in gene expression between normal and diseased conditions. Advanced computational analyses have established gene signatures to predict patients' prognosis and classify primary liver cancers[1, 2] but these tools have yet to be fully integrated into clinical practice. Whole tissue RNA sequencing provides an average readout of the RNA content of a sample, which represents mixed RNA signals from the different cells present within the tissue and, therefore, it is significantly influenced by cell type prevalence. This approach is unable to study rare cell populations, cellular heterogeneity (i.e. cell subsets among major cell types), specific pathogenic cell subpopulations, or to dissect cancer clonal evolution and microenvironment. In the era of immunotherapy and precision medicine, higher resolution sequencing data are required to characterize heterogeneous tissues and complex diseases such as chronic liver disease and cancer.

Recent technological advances enabled genome-wide RNA profiling in individual cells, a technique termed single-cell RNA sequencing (scRNA-seq)[3-6]. In scRNA-seq, liver tissue is dissociated, single cells captured, and RNA sequencing is performed using several workflows (**Figure 1, 2**). ScRNA-seq generates very large datasets of thousands of gene transcripts per cell. These datasets are usually represented in a compressed 2D space, e.g. t-distributed stochastic neighbor embedding (*t*-SNE) map[7], where each cell is a dot and the distance between cells is a function of their similarity (**Figure 1A**). In this 2D space, cells can be clustered according to their similarity and single or multiple genes can be plotted on separate *t*-SNE maps. ScRNA-seq allows discovery, identification and/or study of rare cell types, cell subtypes, disease-specific cell-types and cell-to-cell interactions via ligand-

receptor expression analysis (**Figure 1A**). Furthermore, computational analyses, such as pseudo-time diffusion mapping[8] or RNA velocity[9], allow *in silico* lineage tracing and analysis of development trajectories between cell types (e.g. from progenitor cells to differentiated hepatocytes) or among cell subtypes (e.g. from cytotoxic to exhausted T cells) (**Figure 1B**).

A major challenge of scRNA-seq data is to match the cell RNA profile with cell position within the tissue (i.e. spatial information). This is particularly important in liver biology because the liver is spatially organized in functional lobules and acini[10]. To address this need, spatially-resolved RNA sequencing, paired-cell sequencing, complex computational algorithms and direct spatial transcriptomic techniques – in which scRNA-seq is performed on tissues sections using spatially organized RNA capture probes – have recently been developed.

Here, we summarize and discuss the technical principles of scRNA-seq and spatial transcriptomic approaches and present their application and discoveries regarding liver organization, regeneration, and cell-cell interactions in chronic liver disease and cancer.

## FROM LIVER TISSUE TO SINGLE-CELL RNA SEQUENCING

The initial steps in a scRNA-seq experiment involve tissue dissociation and isolation of single cells which can be obtained by a variety of methods such as fluorescence-activated cell sorting (FACS), magnetic separation using specific antibodies, chip-based or microdroplet-based microfluidic technologies, micromanipulation using an inverted microscope and a motorized micromanipulation platform or laser microdissection[11]. FACS is one of the most widely used techniques and allows the selection of specific cell populations from heterogeneous tissues. High-throughput microdroplet-based microfluidic technologies

(e.g. 10X Chromium) are increasingly used because of high capture efficiency and low costs. Microfluidic technologies are based on the dispersion of single cells into water-in-oil droplets, containing uniquely barcoded beads and primers, using a continuous oil flow as depicted in **Figure 2**. The choice of single-cell capture method greatly depends on the cell types of interests, their prevalence in the tissue, and costs.

After cell isolation, scRNA-seq libraries are generated by cell lysis, reverse transcription into complementary DNA (cDNA), second-strand synthesis and cDNA amplification by polymerase chain reaction (PCR) or *in vitro* transcription (IVT) followed by deep sequencing. These steps vary across the different scRNA-seq protocols (**Figure 2**). Smart-Seq2 is a protocol which uses template-switching technologies for the reverse transcription and PCR technologies for the amplification allowing the sequencing of full-length transcripts and the study of splicing events and allele-specific expression[6, 12, 13]. Smart-Seq2 is limited by high costs and, therefore, different protocols have evolved to allow adequate RNA coverage and reduced costs. These protocols involve the capture of the RNA poly(A) tail with the insertion into the cDNA of random unique molecular identifiers (UMIs) and pre-specified cellular barcodes (**Figure 2**). The presence of both cellular barcodes and UMIs in each single cDNA allow pooling of cDNAs from different cells for the amplification and sequencing steps which reduces significantly the costs per run. The cell of origin is inferred using the cellular barcodes and gene expression is quantified by counting and normalizing UMIs per single cells. In terms of performance, Smart-seq2 and CEL-seq2 showed the highest sensitivity, while Drop-seq has reduced costs but capture efficiency and resolution are lower[3]. Among the different microdroplet-based microfluidic technologies, 10X Chromium results in higher sensitivity and less technical noise[14]. Finally, the combination of two or more scRNA-seq techniques, e.g. a microdroplet-based system and Smart-Seq2, can

be synergistic, increasing the probability of capturing both rare cell types and low abundance transcripts[15].

ScRNA-seq comprises multiple technologies and the choice of platform used should be guided by the biological question. The appropriate technique or combination of techniques should be chosen in the context of the study design and endpoints required (e.g. study of rare cell types or lowly expressed genes or splicing variant analysis). Smart-seq2 is preferred when analyzing splicing, transcriptome annotations or genome integrations while high-throughput microdroplet-based microfluidic technologies are preferred for broader cell coverage at shallower sequencing read depths.

## LIVER PHYSIOLOGY AT SINGLE-CELL LEVEL

### REWIND THE TAPE: GATHER SPATIAL INFORMATION FROM SINGLE-CELL DATA TO STUDY LIVER ZONATION

One of the first applications of scRNA-seq has been the study of liver zonation in mice and humans. The liver is a highly organized tissue, and the porto-central axis of the acinus is a fundamental functional unit during homeostasis and disease development. Hepatocyte function varies along this axis, with hepatocytes classically divided into three zones. A major challenge in the use of scRNA-seq for the study of liver physiology is the integration of individual cell RNA data with spatial information. To overcome this hurdle, specific sequencing strategies and bioinformatic analyses have been developed (**Table 1**), allowing new insights into liver zonation (**Figure 3**). Halpern *et al.* studied liver zonation in mice combining scRNA-seq with single-molecule RNA fluorescence in situ hybridization (smRNA-FISH) to perform spatially-resolved RNA-sequencing[16]. At first, they used smRNA-FISH to

assess at high-resolution the spatial distribution of known zonated landmark genes allowing their fine porto-central profiling. Secondly, scRNA-seq of mouse hepatocytes was performed and the porto-central profile of landmark genes was used to assign a porto-central position to each single cell (for review see[17]). Spatially-resolved scRNA-seq data of the mouse liver discovered that (1) major determinants of liver zonation were not only oxygen gradient and WNT signaling[18], but also RAS signaling, which activates periportal genes, and pituitary signals which inhibits periportal genes (**Figure 3B**); (2) zonation is not always monotonic and some genes, e.g. *Hamp* encoding for hepcidin, have the highest expression in the mid-layers of the lobule (**Figure 3A**); (3) genes encoding for biliary acid metabolism enzymes are differently expressed along the porto-central axis suggesting the spatial zonation of entire metabolic processes; (4) metabolites produced in periportal areas are uptaken by pericentral hepatocytes in a process called spatial recycling.

Once the spatial transcript data of a certain cell type is known, paired-cell sequencing is an elegant technique to infer the zonation of other cell types forming with it strong cell-to-cell interactions (for review see[17]). Halpern and colleagues sequenced doublets of hepatocytes and liver sinusoidal endothelial cells (LSECs) and used hepatocyte single-cell zonation data[16] to infer the zonation of LSECs[19]. This analysis showed that LSECs genes are significantly zonated, and pericentral LSECs are enriched with WNT signaling genes and modulators - major determinants of hepatocyte zonation - suggesting that LSECs might shape hepatocyte zonation.

When surface proteins are available as spatial markers, spatial sorting is a strategy using FACS to sort cells from a specific area. Combining two or more inversely zonated markers allows sorting of cells from specific liver lobule areas facilitating not only scRNA-seq but also multi-omics analyses[20]. Mass spectrometry proteomics and RNA-seq on spatially sorted hepatocytes allowed the mapping of protein zonation and the correlation of gene

expression with protein expression in specific liver zones. Bulk microRNA (miRNA) microarray measurement after spatial sorting on mouse liver revealed that miRNAs are zonated along the porto-central axis[20]. MiRNA are short non-coding RNA oligonucleotides which target specific messenger RNAs to increase their degradation or to decrease their translation[21]. Forty-five percent of known and validated hepatocyte miRNA were found to be mildly pericentral zonated (79%) or strongly periportal zonated (11%)[20] with their targets inversely zonated. Study of mouse miRNA zonation via spatial sorting revealed their inverse correlation with WNT-related genes suggesting a potential role of miRNAs in determining hepatocyte zonation.

Computational analysis can also help with inference of spatial information from scRNA-seq data when spatial organization is the main source of heterogeneity in a tissue (**Table 1**). Aizarani *et al.* applied diffusion pseudotime analysis to model zonation of hepatocytes and LSECs in human healthy liver[22]. This computational analysis was able to (1) profile for the first time, at single-gene level, the porto-central zonation of human hepatocytes and LSECs, (2) discover that LSECs' genes are highly zonated and (3) demonstrate that both hepatocyte and LSECs have genes with non-monotonic zonation patterns. More than 60% of LSECs genes were found to be zonated: periportal LSECs were enriched in genes involved in hormone signaling and metabolism (e.g. incretin and angiotensinogen metabolism) while central/mid LSECs were enriched in genes involved in platelet activation, immunity regulation and scavenger functions. Interestingly, scavenger and platelet activation genes were also enriched in central/mid-zone hepatocytes suggesting a functional co-zonation of hepatocytes and LSECs (**Figure 3C and D**).

More complex computational algorithms, enabling spatial information from scRNA-seq data, have recently been developed. NovoSpaRc is a computational framework allowing *de novo* spatial reconstruction of single-cell gene expression cartographies with or without the

use of known spatial information and marker genes[23]. NovoSpaRc assumes that physically apposed cells probably share similar transcriptomic profiles and that physical distance can be a function of transcriptomic difference. The algorithm can reconstruct, in a virtual space, the organization of symmetric tissues, e.g. normal liver and intestine, but also early embryos and charts of complex tissues such as cerebellum and kidney. However, novoSpaRc has not yet been used to investigate liver spatial organization and function.

One of the challenges of inferring spatial information from standard scRNA-seq data is the requirement for careful, follow-on validation by direct spatial techniques such as immunohistochemistry, immunofluorescence or FISH. To overcome this issue, systems which allow *in situ* spatial transcriptomics have recently been developed[24-26]. These systems generally consist of a special slide covered by beads carrying oligos composed by a polyd(T) tail for RNAs capture, a spatial barcode defining bead position, an UMI for transcript count, promoters and adaptors for cDNA synthesis, amplification and sequencing, and a cleavage site to detach the oligos from the slide (**Figure 4A**). Frozen liver tissue is cut, placed on the spatial transcriptomic slide, stained by hematoxylin and eosin (H&E) and scanned by a conventional microscopy slide scanner. The tissue is lysed - releasing RNA which is captured by the oligos - the capture oligos are cleaved, and library prepared as for scRNA-seq. Once the sequencing is performed, the H&E image is combined with the coordinates of the spatial barcode beads to produce single-cell spatial transcriptomic data. Indeed, H&E staining provides data on cell position and size and allows for the definition of cell boundaries and the assignment of, in certain protocols, spatial barcodes to a single cell. Single-cell transcriptomic data can then be visualized in the tissue 2D space (**Figure 4B**). These techniques have been successfully used in investigating complex tissues such as the brain[24-26] or breast cancer[26], and hold exceptional promise for the detailed study of liver disease.

## USING SCRNA-SEQ TO IDENTIFY PROGENITOR CELLS IN THE CONTEXT OF LIVER DEVELOPMENT AND REGENERATION

Regeneration is one of the key features of liver physiology, but the precise identity and degree of heterogeneity of hepatobiliary precursor cells has still to be fully clarified. Data has mainly been generated from mouse models, proposing differing progenitor populations ranging from biliary-like progenitor cells to differentiated hepatocytes as the major sources of the hepatic epithelial regenerative response, depending on injury model and experimental context[27-31]. ScRNA-seq, with its ability to help study rare cell types, has recently been used in this area, investigating heterogeneity and signaling pathways within hepatobiliary precursors in both fetal and adult livers.

Single-cell analysis of the human fetal liver has identified a distinct hepatobiliary hybrid progenitor (HHyP) cell capable of lineage commitment towards hepatocytes or biliary epithelial cells[32]. The fetal HHyP belongs to the $EPCAM^{+}/NCAM^{+}/TROP2^{-}$ compartment and showed both cholangiocyte, hepatocyte and classical progenitor markers. This cell can be found in the liver ductal plate which is a single or double layered structure of small cuboidal cells at the interface between hepatoblasts and portal mesenchyme.

Aizarani and colleagues used scRNA-seq to analyze the heterogeneity across $EPCAM^{+}$ cells in healthy human livers to understand whether the adult liver has a cell type analogous to the HHyP. They observed considerable heterogeneity within the $EPCAM^{+}$ compartment, which comprises an $EPCAM^{+}TROP2^{int}CK19^{low}$ progenitor cell with high potential for forming bipotent organoids and committing to either hepatocyte or cholangiocyte fate[22, 33] (**Figure 5**). This previously unknown adult liver progenitor cell is located in the canals of Hering and represents the equivalent of the fetal HHyP and the oval cell described in mice[31].

In the normal liver, progenitor cells are usually quiescent and the mechanisms underlying their commitment and activation following liver injury are still unclear. In animal models, liver injury can be induced in a reproducible fashion with several strategies mimicking different liver pathologies. ScRNA-seq has been successfully applied to mouse models of liver injury to study drivers of liver regeneration and has revealed YAP target genes as a major source of the heterogeneity in the $EPCAM^+$ compartment[34]. This YAP target gene signature represents a dynamic inducible state that is upregulated during liver injury, promoting and sustaining progenitor proliferation and liver regeneration (**Figure 5**)[34].

In summary, scRNA-seq has facilitated the discovery of a bipotent progenitor cell in the $EPCAM^+$ compartment in both fetal and normal adult liver, and whose activation is associated with an upregulation of YAP target genes. To accurately characterize liver regeneration drivers in human disease, further analyses focused on progenitor cell populations in human liver after chronic and acute injury are required.

## NOVEL INSIGHTS INTO CHRONIC LIVER DISEASE AND CANCER MICROENVIRONMENT

### THE PHENOTYPE OF NON-PARENCHYMAL CELLS IN CHRONIC LIVER DISEASE AND CIRRHOSIS

Two human liver single-cell atlases provide a detailed insight into the composition of the normal liver using two complementary sequencing techniques, mCEL-Seq2 (miniaturized CEL-Seq2)[35] and 10X Chromium[36] and constitute a reference point for single-cell based research in liver disease[22, 37].

The liver microenvironment, comprising hepatocytes and non-parenchymal cells (NPCs), plays a key role in the pathogenesis of all chronic liver diseases. In response to chronic hepatocyte damage, immune cells produce pro-inflammatory cytokines and chemokines and activate quiescent hepatic stellate cells (HSCs) into myofibroblasts that are responsible for collagen and extra-cellular matrix accumulation[38, 39] - a hallmark of liver fibrosis[40]. This dysregulation of liver immunity is common across different forms of chronic liver diseases and triggers cellular stress and death, apoptosis, liver fibrosis, and hepatocyte proliferation and liver regeneration[38]. Single-cell studies have been carried out to resolve the heterogeneity and complex cell-to-cell interactions of NPCs in chronic liver diseases and cirrhosis.

**Liver endothelial cells are involved in multiple cell-to-cell interactions and prime differentiation of circulating monocytes into liver macrophages.**

A single-cell study of NPCs in healthy and diet-induced NASH amylin (AMLN) mice[41] focused on the characterization of the NPCs secretome and cell-to-cell interactions[42]. LSECs were found to secrete angiocrine factors and express several genes involved in cell-to-cell interactions. Ligands were expressed by cholangiocytes, HSCs but also LSECs, suggesting extensive interactions with other NPCs as well as autocrine signaling. In NASH liver, LSECs upregulated the expression of genes implicated in lipid metabolism, chemokine release and antigen presentation, whilst genes involved in vascular homeostasis and vascular development were downregulated, inducing a significant disruption of sinusoid capillaries[42].

LSECs are the port of entry of monocyte and other bone-marrow derived cells in the liver lobule. LSEC-to-monocyte interactions are crucial in determining the fate of circulating monocytes and their differentiation into liver macrophages[43]. Livers of Kupffer cell (KC)-

depleted mice are rapidly repopulated by circulating monocytes which acquire a KC-like phenotype. LSECs express DLL4 and TGFβ1 that interacts, respectively, with NOTCH and TGFβ/BMP receptors on monocytes, downregulating monocyte-specific genes. Single-cell analysis of mouse models demonstrated that monocyte-derived macrophages are largely expanded in NASH livers with a unique inflammatory phenotype[44]. Whether dysregulated NASH LSECs determine the phenotype of the NASH monocyte-derived macrophages is still unknown and further studies are needed to elucidate LSEC-to-monocyte interactions in the context of NASH pathogenesis.

**Hepatic stellate cells are spatially and functionally zonated, and are hubs of autocrine and paracrine signaling.**

HSCs had previously been thought to represent a functionally homogeneous population. Dobie *et al.* used scRNA-seq to deconvolve the hepatic mesenchyme in both healthy and fibrotic mouse liver, uncovering spatial zonation of HSC across the hepatic lobule[45]. HSCs partition into topographically diametric lobule regions, designated portal vein-associated HSCs (PaHSCs) and central vein-associated HSCs (CaHSCs). HSCs display functional zonation, with CaHSCs representing the dominant pathogenic collagen-producing cells in a mouse model of CCl4-induced centrilobular fibrosis. Furthermore, LPAR1 (lysophosphatidic acid receptor 1) was identified as a therapeutic target on collagen-producing HSCs, and inhibition of LPAR1 resulted in decreased contractility in human HSCs *in vitro,* and a reduction of liver fibrosis in a choline-deficient high-fat diet rodent model of NASH[45].

ScRNA-seq from mouse liver has also shown that HSCs specifically secrete cytokines acting on LSECs, macrophages and cholangiocytes regulating fibrosis pathways, cytokine

expression, vasoactive hormone signaling and HSC apoptosis via secretion of nerve growth factor[42]. HSCs express both *Il11ra1*, a receptor belonging to the IL-6 family, and its ligand *Il11,* constituting a previously unknown autocrine signal which stimulates the activation of STAT3 and ERK and cytokine secretion[42]. Analysis of HSC gene expression also revealed potential extrahepatic modulation of this cell type. HSCs express vasoactive hormone-responsive receptors mediating both contraction and relaxation which specifically mediate the effect of calcitonin gene-related peptide, PTH and VIP which are not expressed by any liver cell type. In the classical view of liver fibrosis pathogenesis, HSCs are the final effector and the last step of the NPCs activation cascade. Single-cell analysis of HSCs in NASH showed that they upregulate both *Il11* and cytokines and modulate function of LSECs and macrophages, suggesting a more complex bidirectional interaction. Altogether, single-cell profiling of HSC has further extended the concept that this cell type acts as a central hub in the paracrine/autocrine network of liver NPCs in both normal and diseased liver.

**Macrophage phenotype and non-parenchymal cell interactions in the fibrotic niche.**

In homeostasis, the liver is continuously exposed to pathogens and toxins derived from the gut and removes large numbers of microbes and microbe-associated molecules to maintain a tolerant and immunosuppressive environment[46]. Data from the human liver single-cell atlases have shown that the normal liver contains not only immunomodulating macrophages with metabolic and scavenger functions, but also proinflammatory macrophages[22, 37].

In NASH mouse models, scRNA-seq has demonstrated an expansion of macrophages with a proinflammatory phenotype. Macrophages in NASH express high levels

of *Trem2* encoding for an innate immunity scavenger receptor implicated in phagocytosis and clearance of apoptotic cells. This receptor has been described in the pathogenesis of Alzheimer's disease as a microglia metabolism modifier[47], and in human and mouse adipose tissue macrophages in response to pathogenic lipid accumulation[48]. Liver *Trem2*$^{high}$ macrophages were enriched in genes involved in antigen presentation, ECM remodeling, endocytosis and lysosomal degradation suggesting an important role in NASH pathogenesis[42]**.** *Trem2*$^{high}$ macrophages also overexpress *Cd9* which encodes for a tetraspanin protein involved in many cellular processes including cell differentiation, adhesion, and signal transduction[49] and prevents macrophage fusion into multinucleated giant cells[50]. Furthermore, this NASH-associated macrophage expresses *Gpnmb* which is a transmembrane glycoprotein negatively regulating inflammation and previously described in macrophages infiltrating the liver during the recovery phase of CCl4-induced acute liver injury[51, 52]. *Trem2*$^{high}$ macrophages represent over 60% of KCs in NASH livers whilst they were almost undetectable in control mice, and their prevalence is reduced upon treatment with elafibranor or switch from AMLN diet to chow with subsequent improvement of liver inflammation. Ramachandran and colleagues performed scRNA-seq of healthy and cirrhotic human livers and investigated heterogeneity in fibrosis-associated NPCs[53]. Specific macrophage subpopulations were more prevalent in cirrhotic tissue and were annotated as scar-associated macrophages (SAMΦ). SAMΦ were marked by the expression of *TREM2* and *CD9*, and were able to activate HSCs. Self-organizing maps and pseudotime analysis at single-cell level revealed that SAMΦ are derived from blood monocytes. The differentiation process towards SAMΦ fate involved the expression of genes related to antigen processing and presentation, phagocytosis, chemokines, angiogenesis, production of extracellular matrix and wound healing. SAMΦ were also found in the early stages of NAFLD and in a CCl4 mouse model of liver fibrosis. Overall, these data suggest that TREM2$^{+}$ SAMΦ are

monocyte-derived macrophages that represent a conserved innate response to chronic liver damage, promoting mesenchymal cell activation and fibrogenesis. Ongoing studies are investigating ways to manipulate this macrophage subpopulation for therapeutic gain, and more functional data are required to fully understand the contribution of this novel macrophage subtype across different aetiologies of chronic liver disease.

ScRNA-seq analysis also unveiled the complexity of the cellular interactome of the human liver fibrotic niche, identifying not only SAMΦ but PDGFRα$^{+}$ mesenchymal cells (SAMes, scar-associated mesenchymal cells) and two, previously unknown, scar-associated endothelial cell subpopulations (CD34$^{+}$PLVAP$^{+}$VWA1$^{+}$ and CD34$^{+}$PLVAP$^{+}$ACKR1$^{+}$, SAEndo)[53]. Using single-cell data, multi-lineage ligand-receptor interaction analysis and multiplex immunofluorescence, the multi-directional interactions between SAMΦ, SAEndo and SAMes were characterized (**Figure 6**). Multi-lineage modelling of ligand-receptor interactions between these cells revealed intra-scar activity of several pro-fibrogenic pathways including TNFRSF12A, PDGFR and NOTCH signaling. As an example, SAMes and SAEndo interact via non-canonical Notch ligands *DLL4, JAG1, JAG2* with the receptor *NOTCH3* expressed on SAMes and primary SAEndo from cirrhotic liver cultured together with HSCs promoted collagen production which decreased upon treatment with the Notch-signaling inhibitor Dibenzazepine.

In summary, scRNA-seq revealed novel scar-associated subpopulations of macrophages, endothelial cells and mesenchymal cells inhabiting the fibrotic niche of human liver cirrhosis, and has shed light on how these different cell types interact to promote fibrosis. Fibrogenesis in cirrhosis is a highly complex process characterized by the interaction of multiple different cell lineages which are in various states of differentiation and activation. Development of novel anti-fibrotic therapies will require consideration of the complexity of the

human liver fibrotic niche, and will likely need to modulate multiple therapeutic targets simultaneously to achieve anti-fibrotic efficacy.

## UNRAVELING TUMOR MICROENVIRONMENT AND HETEROGENEITY WITHIN PRIMARY LIVER CANCER

Aizarani *et al.* demonstrated the potential of hepatocellular carcinoma (HCC) single-cell analysis using their normal human cell atlas as a reference to characterize perturbed cell states in HCC[22]. They showed that (1) cancer epithelial cells upregulate pro-inflammatory, WNT and Hedgehog genes, (2) endothelial cells in HCC lose classical sinusoidal markers and display typical macrovascular endothelial cell markers in line with the arterialization process characterizing HCC development and (3) both HCC endothelial cells and macrophages downregulate innate immunity pathways and upregulate receptor tyrosine kinase signaling pathways, targets of the currently approved systemic treatments for HCC such as sorafenib and regorafenib. Interestingly, HCC endothelial cells expressed *CD34* and *PLVAP* at high levels, as also observed in SAEndo[53], suggesting potentially common changes in fibrosis-associated and cancer-associated endothelial cells.

ScRNA-seq analysis has also allowed new insights into the complexity of the immune cell microenvironment in HCC (**Figure 7**). Zheng *et al.* investigated, at single cell level, T cell composition in blood, non-tumor liver and tumor tissues from HCC patients. T regulatory cells (Tregs, $CD4^+CTLA4^+$) with immunosuppressive functions and exhausted $CD8^+$ T cells (Tex) were clonally enriched, with the latter predicted to originate from cytotoxic $CD8^+$ T cells via an intermediate $CD8^+GZMK^+$ T cell subtype. *LAYN,* a transmembrane protein with homology to c-type lectin, was identified as a novel marker of T cell exhaustion and its expression in HCC was found to be associated with higher rates of tumor recurrence[54]. Combining Smart-Seq2

and 10X Chromium approaches, Zhang and colleagues performed scRNA-seq of $CD45^+$ immune cells from tumor, lymph nodes (LN) and ascites to characterize macrophages and dendritic cells (DCs) in HCC[15]. LAMP3 is a DC-specific glycoprotein induced upon DC maturation after inflammatory stimulation[55]. Mature *LAMP3*$^+$ DCs were observed in both HCC and LNs, and these cells were predicted to interact with T and NK cells via IL-15 and PD-1/PD-L1 and, importantly, they were strongly associated with T cell dysfunction[15]. Macrophages in HCC were found to have two main distinct states. Some macrophages resembled myeloid-derived suppressor cells, which have a strong immunosuppressive phenotype and can regulate the function of other immune cell types including T cells and DCs[56, 57]. A second macrophage group were similar to the tumor-associated macrophages (TAM) described in lung cancer[58] with a mixed proinflammatory-immunosuppressive phenotype. TAM-like macrophages express *TREM2* and *GPNMB* similarly to the SAMΦ described in fibrotic livers.

The factors shaping tumor microenvironment (TME) in HCC are still not known. In a recent study, single-cell analysis was used to explore the interconnection between intratumor heterogeneity (ITH) and TME. Data from both HCC and intrahepatic cholangiocarcinoma showed that tumors with higher ITH have a more immunosuppressive TME, are associated with more hypoxia-related genes, higher VEGF expression and lower long-term patient survival. Hypoxia and VEGF secretion from cancer epithelial cells seems to be the main mechanism driving ITH and TME changes in heterogenous cancers giving a supplementary rationale to the anti-VEGF and anti-angiogenic drugs in the treatment of primary liver cancer[59].

Considering the urgent need for new treatment strategies for liver cancer, scRNA-seq could help not only in the identification of new therapeutic targets but also in the development of more refined tumor classification, allowing more accurate tailoring of a patient's treatment.

Preliminary classification using scRNA-seq has been already developed[60], but will need prospective validation before being incorporated into routine use.

Collectively scRNA-seq has helped characterize the cellular phenotypes of various cell types within the HCC microenvironment, and has shed light on the interplay between cancer epithelial cells and TME. HCC is a complex cellular ecosystem, including clonal Tregs, clonal *CD8*$^{+}$*LAYN*$^{+}$ Tex, pre-exhausted *CD8*$^{+}$*GZMK*$^{+}$ *cells, LAMP3*$^{+}$ DC, myeloid-derived suppressor cells, TAM-like macrophages and *PLVAP*$^{+}$ endothelial cells resembling endothelial cells inhabiting the liver fibrotic niche. ScRNA-seq and spatial transcriptomic approaches will be valuable tools to help increase our understanding of the cellular and molecular mechanisms regulating the TME, which should in turn aid in the identification of novel treatment targets for hepatobiliary cancers. Furthermore, these approaches should also be informative with regard to development of more precise tumor classification and patient stratification, thereby refining clinical trial design in this area.

## CHALLENGES AND PERSPECTIVES

While scRNA-seq and the associated cutting-edge computational analyses have revolutionized investigation of complex organs and tissues and hold great promise for enabling future discoveries in the liver field, several challenges still need to be addressed. Dissociation is a critical step that can induce transcriptomic changes[61] and should be carefully optimized to obtain the maximum dissociation yield without inducing biases. Furthermore, scRNA-seq is expensive and the analysis of single-cell data is time consuming and requires skilled bioinformatics support. Direct spatial transcriptomic techniques can potentially overcome some of these issues but their sensitivity and validity in liver-related studies is still to be determined. Finally, technologies are rapidly moving towards the

development of multi-omics single cell approaches that will allow the characterization of proteomic, gene expression and DNA mutations in the same cell[62]. Single-cell multi-omics will allow an even more comprehensive understanding of liver biology and disease at single-cell resolution. Efforts are needed to reduce the costs of single-cell genomics technologies, and then identify histological or radiological surrogate markers to help characterize and stratify liver disease, which in turn will help predict drug response or patients' prognosis without having to recourse to full single-cell analysis of samples from patients.

## CONCLUSIONS

ScRNA-seq is a revolutionary technique which has already been successfully applied to study the biology of healthy and diseased liver at unprecedented resolution, capturing the heterogeneity of cell types and states and characterizing cell-to-cell interactions. The choice of scRNA-seq approach relies on study design, endpoints and costs, and often entails a compromise between costs and gene coverage. Computational algorithms, direct spatial transcriptomics and combination of scRNA-seq and spatial techniques enable the study of single-cell gene expression in complex, highly spatially organized tissues.

ScRNA-seq has already delivered transformative new discoveries in the understanding of liver zonation, regeneration, and the biology of chronic liver disease and cancer. Liver disease biology involves multiple cell types and complex cell-to-cell interactions, and scRNA-seq allows detailed investigation of these multi-cellular microenvironments. The challenge now is to fully harness and translate this new knowledge into effective novel therapeutic approaches to address the major clinical challenges in hepatology.

## ACKNOWLEDGMENTS

A.S. is the recipient of a fellowship of Region-Alsace, IHU and LabEx HEPSYS. N.C.H. acknowledges the support of the Wellcome Trust (Senior Research Fellowship in Clinical Science, ref. 103749), Medical Research Council and the Chan Zuckerberg Initiative. T.F.B. acknowledges support by the European Union (ERC-AdG2014 HEPCIR #671231, H2020 HEPCAR #667273, ERC PoC-HEPCAN #862551), Fondation ARC Paris and IHU Strasbourg (TheraHCC2.0, IHU201901299), the Foundation of the University of Strasbourg and Roche Institute (HEPKIN), the Agence Nationale de Recherches sur le Sida (ANRS) and the US National Institutes of Health (R21CA209940, R03AI131066, R01CA233794, U19AI12386). This work has been published under the framework of the LABEX ANR-10-LABX-0028_HEPSYS and PLAN CANCER 2014-2019 HCCMICTAR and benefits from a funding from the state managed by the French National Research Agency as part of the Investments for the Future Program, INCa (National Institute for Cancer) and Inserm.

The authors do not declare any conflict of interest.

## REFERENCES


1. Boyault S, Rickman DS, de Reynies A, Balabaud C, Rebouissou S, Jeannot E, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology 2007;45:42-52.

2. Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Research 2009;69:7385-7392.

3. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative Analysis of Single-Cell RNA Sequencing Methods. Mol Cell 2017;65:631-643 e634.

4. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 2015;161:1202-1214.

5. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol 2016;17:77.

6. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 2013;10:1096-1098.

7. Li W, Cerise JE, Yang Y, Han H. Application of t-SNE to human genetic data. J Bioinform Comput Biol 2017;15:1750017.

8. Haghverdi L, Buttner M, Wolf FA, Buettner F, Theis FJ. Diffusion pseudotime robustly reconstructs lineage branching. Nat Methods 2016;13:845-848.

9. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. Nature 2018;560:494-498.

10. Gebhardt R. Metabolic zonation of the liver: regulation and implications for liver function. Pharmacol Ther 1992;53:275-354.

11. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 2018;50:96.

12. Ramskold D, Luo S, Wang YC, Li R, Deng Q, Faridani OR, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 2012;30:777-782.

13. Deng Q, Ramskold D, Reinius B, Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science 2014;343:193-196.

14. Zhang X, Li T, Liu F, Chen Y, Yao J, Li Z, et al. Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems. Mol Cell 2019;73:130-142 e135.

15. Zhang Q, He Y, Luo N, Patel SJ, Han Y, Gao R, et al. Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma. Cell 2019;179:829-845 e820.

16. Halpern KB, Shenhav R, Matcovitch-Natan O, Toth B, Lemze D, Golan M, et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature 2017;542:352-356.

17. Ben-Moshe S, Itzkovitz S. Spatial heterogeneity in the mammalian liver. Nat Rev Gastroenterol Hepatol 2019;16:395-410.

18. Planas-Paz L, Orsini V, Boulter L, Calabrese D, Pikiolek M, Nigsch F, et al. The RSPO-LGR4/5-ZNRF3/RNF43 module controls liver zonation and size. Nat Cell Biol 2016;18:467-479.

19. Halpern KB, Shenhav R, Massalha H, Toth B, Egozi A, Massasa EE, et al. Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. Nat Biotechnol 2018;36:962-970.

20. Ben-Moshe S, Shapira Y, Moor AE, Manco R, Veg T, Bahar Halpern K, et al. Spatial sorting enables comprehensive characterization of liver zonation. Nat Metab 2019;1:899-911.

21. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 2010;466:835-840.

22. Aizarani N, Saviano A, Sagar, Mailly L, Durand S, Herman JS, et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature 2019;572:199-204.

23. Nitzan M, Karaiskos N, Friedman N, Rajewsky N. Gene expression cartography. Nature 2019;576:132-137.

24. Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, et al. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. Science 2019;363:1463-1467.

25. Stahl PL, Salmen F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 2016;353:78-82.

26. Vickovic S, Eraslan G, Salmen F, Klughammer J, Stenbeck L, Schapiro D, et al. High-definition spatial transcriptomics for in situ tissue profiling. Nat Methods 2019;16:987-990.

27. Furuyama K, Kawaguchi Y, Akiyama H, Horiguchi M, Kodama S, Kuhara T, et al. Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. Nature Genetics 2011;43:34-41.

28. Font-Burgada J, Shalapour S, Ramaswamy S, Hsueh B, Rossell D, Umemura A, et al. Hybrid Periportal Hepatocytes Regenerate the Injured Liver without Giving Rise to Cancer. Cell 2015;162:766-779.

29. Wang B, Zhao L, Fish M, Logan CY, Nusse R. Self-renewing diploid Axin2(+) cells fuel homeostatic renewal of the liver. Nature 2015;524:180-185.

30. Tarlow BD, Pelz C, Naugler WE, Wakefield L, Wilson EM, Finegold MJ, et al. Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes. Cell Stem Cell 2014;15:605-618.

31. Deng X, Zhang X, Li W, Feng RX, Li L, Yi GR, et al. Chronic Liver Injury Induces Conversion of Biliary Epithelial Cells into Hepatocytes. Cell Stem Cell 2018;23:114-122 e113.

32. Segal JM, Kent D, Wesche DJ, Ng SS, Serra M, Oules B, et al. Single cell analysis of human foetal liver captures the transcriptional profile of hepatobiliary hybrid progenitors. Nat Commun 2019;10:3350.

33. Huch M, Gehart H, van Boxtel R, Hamer K, Blokzijl F, Verstegen MM, et al. Long-term culture of genome-stable bipotent stem cells from adult human liver. Cell 2015;160:299-312.

34. Pepe-Mooney BJ, Dill MT, Alemany A, Ordovas-Montanes J, Matsushita Y, Rao A, et al. Single-Cell Analysis of the Liver Epithelium Reveals Dynamic Heterogeneity and an Essential Role for YAP in Homeostasis and Regeneration. Cell Stem Cell 2019;25:23-38 e28.

35. Sagar, Herman JS, Pospisilik JA, Grun D. High-Throughput Single-Cell RNA Sequencing and Data Analysis. Methods Mol Biol 2018;1766:257-283.

36. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun 2017;8:14049.

37. MacParland SA, Liu JC, Ma XZ, Innes BT, Bartczak AM, Gage BK, et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat Commun 2018;9:4383.

38. Marrone G, Shah VH, Gracia-Sancho J. Sinusoidal communication in liver fibrosis and regeneration. J Hepatol 2016;65:608-617.

39. Saviano A, Roehlen N, Virzi A, Suarez AAR, Hoshida Y, Lupberger J, et al.: Stromal and Immune Drivers of Hepatocarcinogenesis. In: Hoshida Y, ed. Hepatocellular Carcinoma: Translational Precision Medicine Approaches. Cham (CH), 2019; 317-331.

40. Pellicoro A, Ramachandran P, Iredale JP, Fallowfield JA. Liver fibrosis and repair: immune regulation of wound healing in a solid organ. Nat Rev Immunol 2014;14:181-194.

41. Clapper JR, Hendricks MD, Gu G, Wittmer C, Dolman CS, Herich J, et al. Diet-induced mouse model of fatty liver disease and nonalcoholic steatohepatitis reflecting clinical disease progression and methods of assessment. Am J Physiol Gastrointest Liver Physiol 2013;305:G483-495.

42. Xiong X, Kuang H, Ansari S, Liu T, Gong J, Wang S, et al. Landscape of Intercellular Crosstalk in Healthy and NASH Liver Revealed by Single-Cell Secretome Gene Analysis. Mol Cell 2019;75:644-660 e645.

43. Sakai M, Troutman TD, Seidman JS, Ouyang Z, Spann NJ, Abe Y, et al. Liver-Derived Signals Sequentially Reprogram Myeloid Enhancers to Initiate and Maintain Kupffer Cell Identity. Immunity 2019;51:655-670 e658.

44. Krenkel O, Hundertmark J, Abdallah AT, Kohlhepp M, Puengel T, Roth T, et al. Myeloid cells in liver and bone marrow acquire a functionally distinct inflammatory phenotype during obesity-related steatohepatitis. Gut 2019.

45. Dobie R, Wilson-Kanamori JR, Henderson BEP, Smith JR, Matchett KP, Portman JR, et al. Single-Cell Transcriptomics Uncovers Zonation of Function in the Mesenchyme during Liver Fibrosis. Cell Rep 2019;29:1832-1847 e1838.

46. Jenne CN, Kubes P. Immune surveillance by the liver. Nature Immunology 2013;14:996.

47. Ulland TK, Song WM, Huang SC, Ulrich JD, Sergushichev A, Beatty WL, et al. TREM2 Maintains Microglial Metabolic Fitness in Alzheimer's Disease. Cell 2017;170:649-663 e613.

48. Jaitin DA, Adlung L, Thaiss CA, Weiner A, Li B, Descamps H, et al. Lipid-Associated Macrophages Control Metabolic Homeostasis in a Trem2-Dependent Manner. Cell 2019;178:686-698 e614.

49. Nakazawa Y, Sato S, Naito M, Kato Y, Mishima K, Arai H, et al. Tetraspanin family member CD9 inhibits Aggrus/podoplanin-induced platelet aggregation and suppresses pulmonary metastasis. Blood 2008;112:1730-1739.

50. Takeda Y, Tachibana I, Miyado K, Kobayashi M, Miyazaki T, Funakoshi T, et al. Tetraspanins CD9 and CD81 function to prevent the fusion of mononuclear phagocytes. J Cell Biol 2003;161:945-956.

51. Katayama A, Nakatsuka A, Eguchi J, Murakami K, Teshigawara S, Kanzaki M, et al. Beneficial impact of Gpnmb and its significance as a biomarker in nonalcoholic steatohepatitis. Sci Rep 2015;5:16920.

52. Kumagai K, Tabu K, Sasaki F, Takami Y, Morinaga Y, Mawatari S, et al. Glycoprotein Nonmetastatic Melanoma B (Gpnmb)-Positive Macrophages Contribute to the Balance between Fibrosis and Fibrolysis during the Repair of Acute Liver Injury in Mice. PLoS One 2015;10:e0143413.

53. Ramachandran P, Dobie R, Wilson-Kanamori JR, Dora EF, Henderson BEP, Luu NT, et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. Nature 2019.

54. Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. Cell 2017;169:1342-1356 e1316.

55. Anthony PP, Ishak KG, Nayak NC, Poulsen HE, Scheuer PJ, Sobin LH, et al. The morphology of cirrhosis. Recommendations on definition, nomenclature, and classification by a working group sponsored by the World Health Organization. Journal of Clinical Pathology 1978;31:395-414.

56. Condamine T, Dominguez GA, Youn JI, Kossenkov AV, Mony S, Alicea-Torres K, et al. Lectin-type oxidized LDL receptor-1 distinguishes population of human polymorphonuclear myeloid-derived suppressor cells in cancer patients. Sci Immunol 2016;1.

57. Zhao F, Hoechst B, Duffy A, Gamrekelashvili J, Fioravanti S, Manns MP, et al. S100A9 a new marker for monocytic human myeloid-derived suppressor cells. Immunology 2012;136:176-183.

58. Lavin Y, Kobayashi S, Leader A, Amir ED, Elefant N, Bigenwald C, et al. Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. Cell 2017;169:750-765 e717.

59. Ma L, Hernandez MO, Zhao Y, Mehta M, Tran B, Kelly M, et al. Tumor Cell Biodiversity Drives Microenvironmental Reprogramming in Liver Cancer. Cancer Cell 2019;36:418-430 e416.

60. Zhang Q, Lou Y, Yang J, Wang J, Feng J, Zhao Y, et al. Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas. Gut 2019;68:2019-2031.

61. van den Brink SC, Sage F, Vertesy A, Spanjaard B, Peterson-Maduro J, Baron CS, et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nat Methods 2017;14:935-936.

62. Hou Y, Guo H, Cao C, Li X, Hu B, Zhu P, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. Cell Res 2016;26:304-319.

## TABLES

**Table 1. Spatial transcriptomics and strategies to match scRNA-seq data with spatial information**

| Methods | Required input data other than scRNA-seq | Pros/Cons |
|---|---|---|
| Spatially-resolved RNA-seq[16] | Accurate spatial pattern of two or more marker genes | High resolution and accurate. |
| Paired-scRNA-seq[19] | Spatial pattern of one cell forming strong cell-to-cell interactions with the cell of interest | High resolution and accurate. |
| Spatial sorting analysis[20] | Known extracellular marker proteins to be used for FACS | Known extracellular marker proteins are not always available. Can be used for multi-omics analysis. |
| DPT analysis[22] | None | Cell diversity needs to be correlated with cell position in the tissue. Validation by histology, smRNA-FISH or other imaging techniques is needed. |
| Gene cartography (novoSpaRc)[23] | Optional Marker genes and general tissue organization | Cell diversity needs to be correlated with cell position in the tissue. Marker genes are optional inputs to refine the analysis. Validation by histology, smRNA-FISH or other imaging techniques is needed. |
| In situ spatial transcriptomics[24-26] | Slide-based system | Lower sequencing depth than classical scRNA-seq but higher spatial resolution. High costs. Not data available yet on human liver tissue. |

DPT= diffusion pseudo-time, scRNA-seq= single-cell RNA-sequencing, smRNA-FISH= single-molecule RNA fluorescent *in situ* hybridization.

## FIGURE LEGENDS

**Figure 1. Single-cell RNA-sequencing analyses to study liver pathophysiology.** A) Normal and/or diseased liver tissue are dissociated into a single cell suspension and scRNA-seq is performed. Thousands of transcripts per cell are compressed in a 2D space where each cell is a dot and the distance between cells is a function of their similarity. Cells are can be aggregated in clusters or groups of clusters plotted as different colors and potentially representing cell types or subtypes. ScRNA-seq allows the study of rare cell types, cell state and subtype heterogeneity, disease-specific cell type and cell-to-cell interactions via ligand-receptor analysis. B) Computational analyses such as pseudo-time diffusion mapping or RNA velocity, which analyze cell similarity and diversity, consent to trace differentiation processes, clonal evolution and cell state transitions of a specific cell type or between different cell types (from cell of origin to development A or B).

**Figure 2. Main steps in scRNA-seq workflows and comparison of the most widely used protocols.** Smart-seq2 and CEL-Seq2 are performed in 96 or 384-well plates after FACS sorting, while droplet systems (e.g. 10X Chromium and Drop-Seq) couple cells with barcoded beads containing unique molecule identifier (UMI) and primers forming water-in-oil droplets via a continuous oil flow. Reverse transcription and cDNA amplification are performed by polymerase chain reaction (PCR) in Smart-Seq and 10X Chromium and by *in vitro* transcription (IVT) in CEL-Seq2. In CEL-Seq2 and 10X Chromium protocols, UMI and cell-specific barcodes are added during reverse transcription to allow the pooling of the subsequent steps. Libraries are prepared by fragmentation in CEL-Seq2 and by tagmentation with or without 3'enrichment in Smart-Seq2 and 10X Chromium. Gene coverage is full-length

in Smart-seq2 whereas in CEL-Seq2 and 10X Chromium only the 3' part of the gene is sequenced.

**Figure 3. New concepts in liver zonation derived from scRNA-seq studies.** A) Zonated genes can have a non-monotonic pattern with genes peaking in the mid-layers of liver lobule. B) Determinants of mouse hepatocyte zonation on periportal (left) and pericentral genes (right). C-D) Liver sinusoidal endothelial cell (LSEC) specific zonated pathways. Periportal LSECs are enriched in pathways related to hormone signaling and metabolism while pericentral LSECs are enriched in immune regulatory genes, WNT-related genes, platelet activation and scavenger function pathways. LSECs and hepatocytes show co-zonation in pericentral areas of WNT-related genes (mouse data) and platelet activation and scavenger function pathways (human data). MiRNA = microRNA.

**Figure 4. Model of in situ spatial transcriptomics.** A) Frozen liver tissue is cut and placed on a special slide special slide covered by beads carrying capture oligos composed by a polyd(T) tail for RNAs capture, a spatial barcode defining bead position, an UMI for transcript count, promoters and adaptors for cDNA synthesis, amplification and sequencing and a cleavage site to detach the oligos from the slide. B) The liver tissue on the spatial transcriptomics slide is fixed, stained by hematoxylin and eosin (H&E) and scanned by a conventional microscopy slide scanner. The tissue is lysed to release RNA, the capture oligos are cleaved and the libraries prepared as for scRNA-seq. The H&E image combined with data and coordinates of the spatial barcodes produce high-resolution single-cell gene expression data. GLUL encoding for glutamate-ammonia ligase is a known pericentral

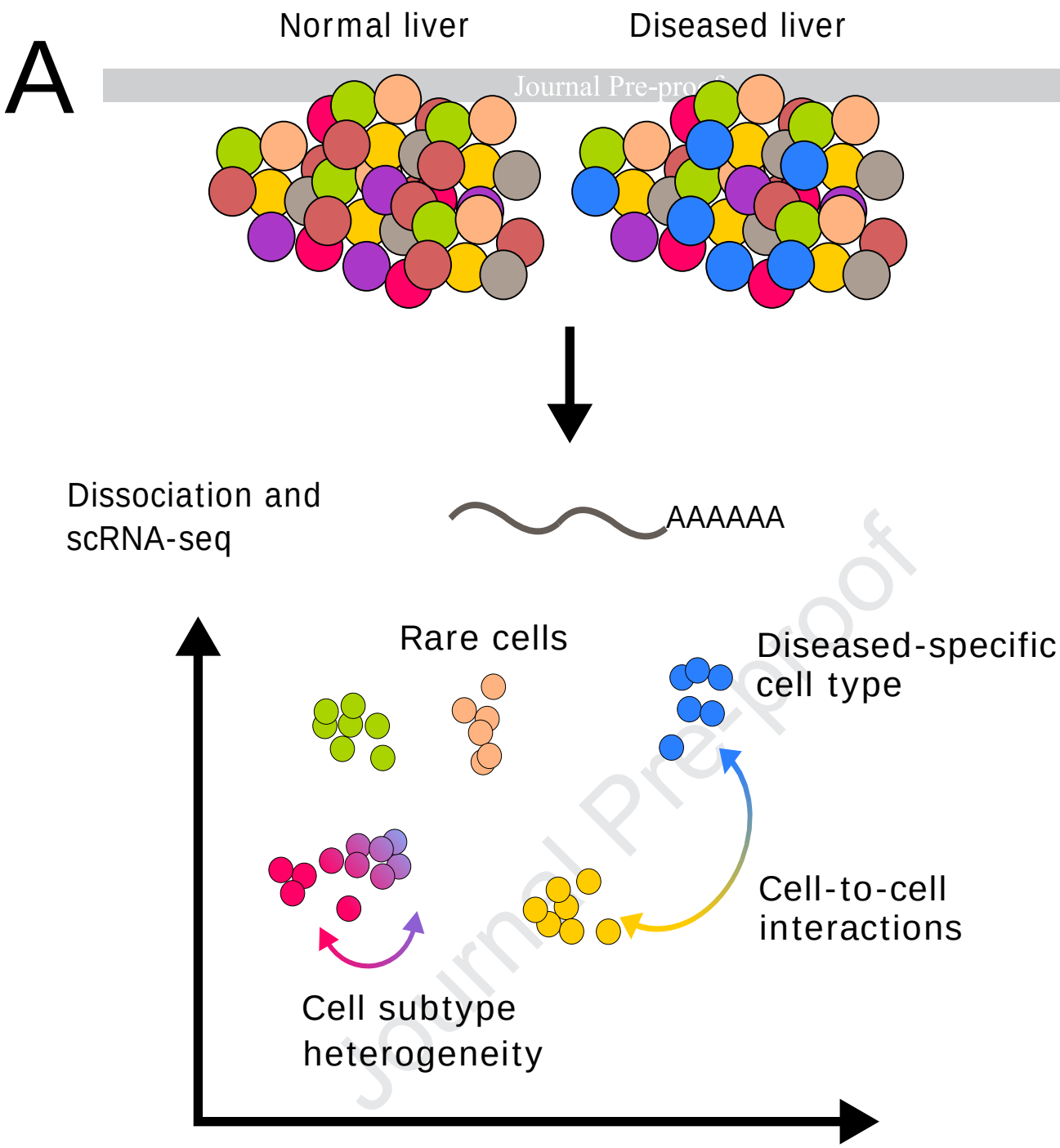
zonated gene and PCK1 encoding for phosphoenolpyruvate carboxykinase 1 is a periportal zonated gene.

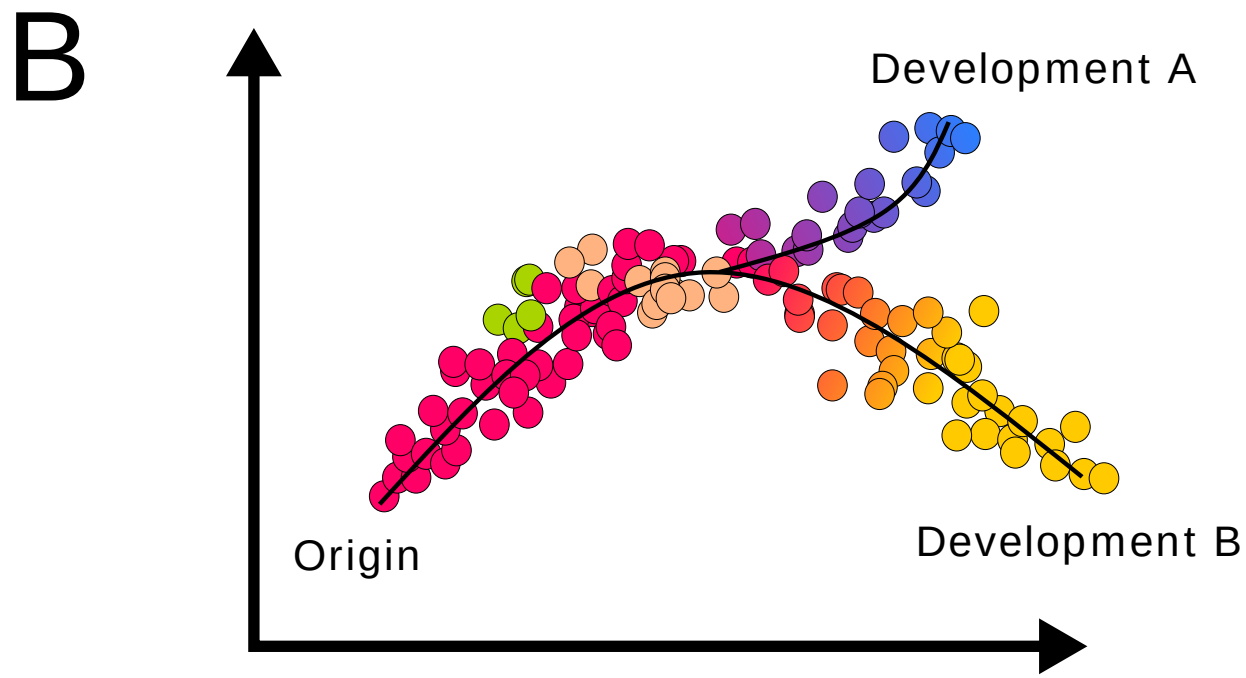
**Figure 5. Bipotent progenitor cell in the normal human liver revealed by scRNA-seq.** ScRNA-seq of human liver identified an $EPCAM^{+}TROP2^{inter}CK19^{low}$ progenitor cell which has the potential to differentiate into cholangiocytes or hepatocytes (human data). Upon liver damage, progenitor cells upregulate WNT- and YAP-target genes promoting liver regeneration (mouse data).

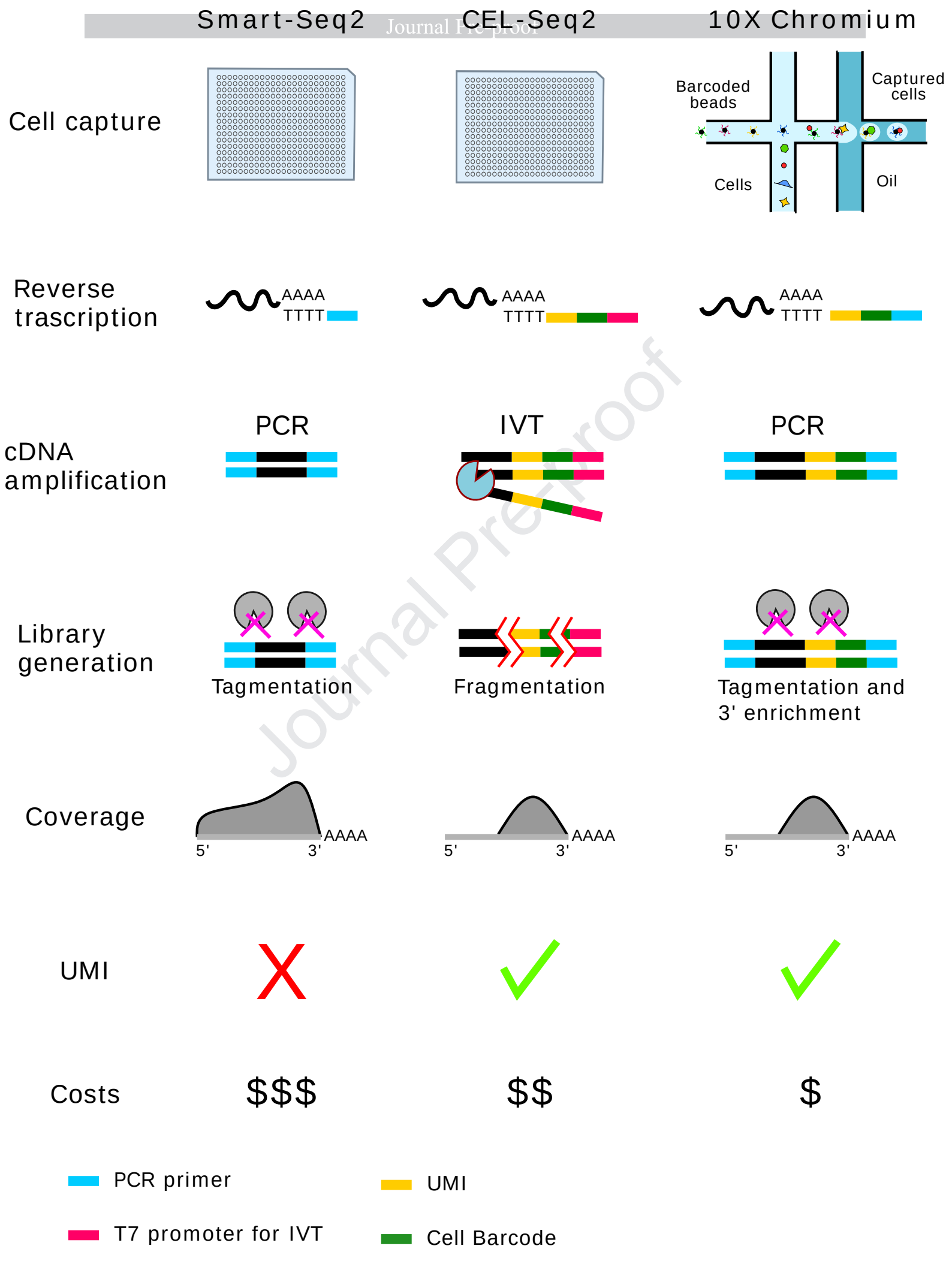
**Figure 6. Intercellular ligand-receptor interactions in the human liver fibrotic niche.** Main receptors and ligands involved in Interactions between scar-associated macrophages, scar-associated mesenchymal cells and scar-associated liver endothelial cells are presented. The most relevant molecules belong to Notch signaling, PDGF and VEGF, TGFβ and TNF family.

**Figure 7. Insights of the tumor microenvironment in hepatocellular carcinoma using scRNASeq.** Hepatocellular carcinoma (HCC) is enriched in clonal $CD4^{+}CTLA4^{+}$ Treg and exhausted $CD8^{+}LAYN^{+}$ lymphocytes expressing the same TCR (T-cell receptor). $CD8^{+}LAYN^{+}$ lymphocytes derived from $CD8^{+}GMZK^{+}$ lymphocytes. $LAMP3^{+}$ dendritic cells (DC) are mature DCs enriched in HCC, interacting with exhausted T cells and Tregs via IL-15 and PD1/PD-L1 axis and capable of migrating into lymph nodes. HCC microenvironment includes also myeloid-derived suppressor cells with strong immunosuppressive functions and

tumor-associated macrophages -like cells which have intermediate proinflammatory-immunosuppressive phenotype and express TREM2.

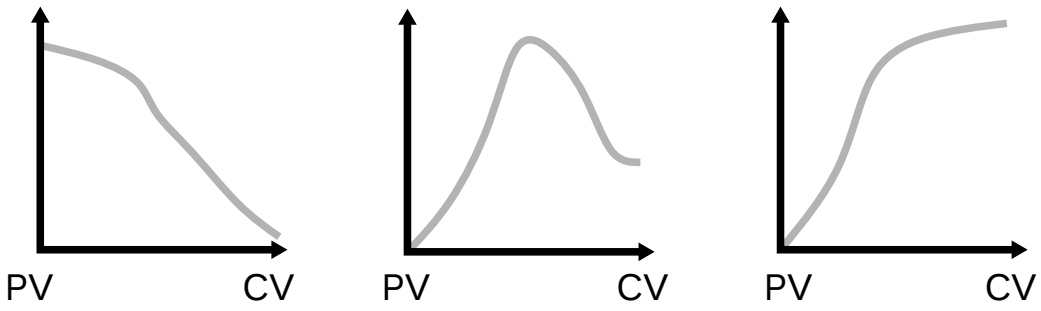




A

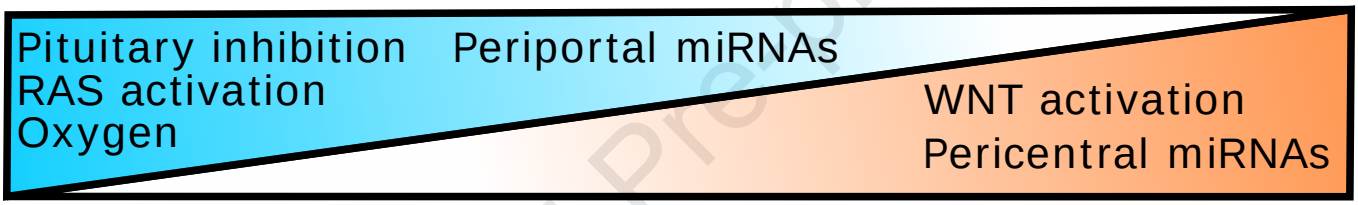
Zonated gene expression patterns




Periportal Mid-layer Pericentral

B

Determinants of mouse hepatocyte zonation

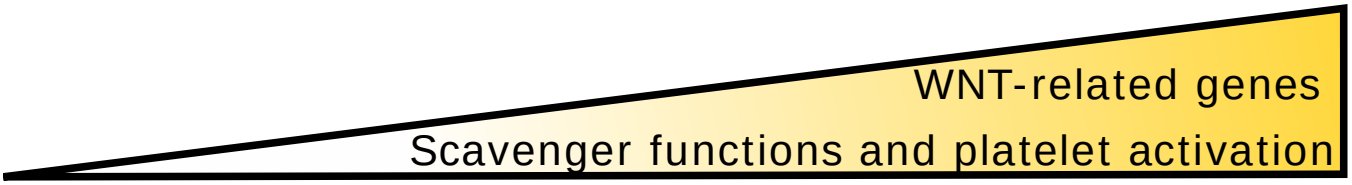



C

Pathways only zonated in endothelial cells

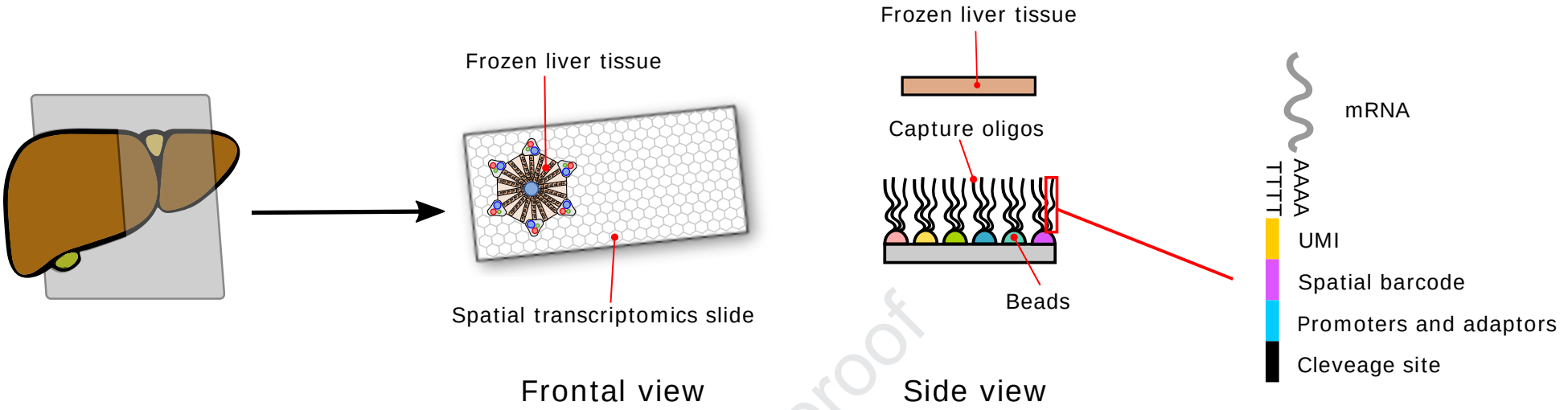

D

Hepatocyte and endothelial cells co-zonated pathways




Portal Vein Central Vein

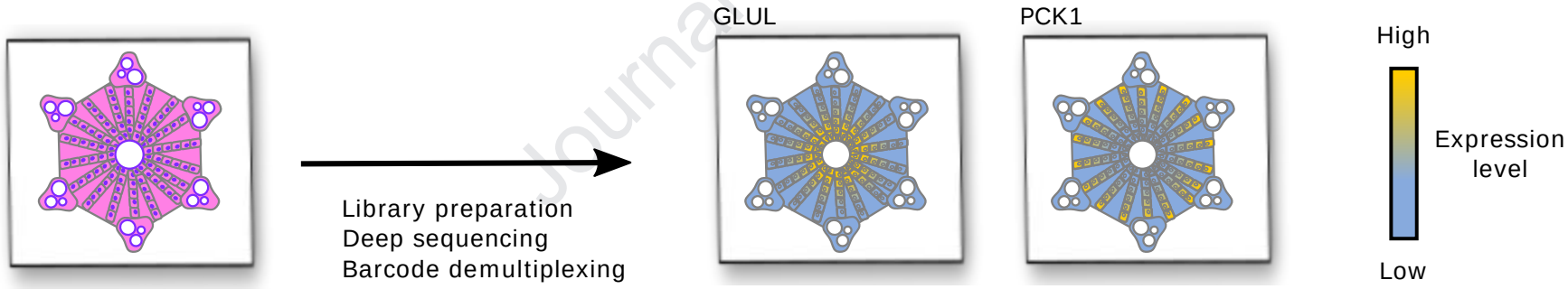
A




Frontal view

Side view

B




Library preparation
Deep sequencing
Barcode demultiplexing

H&E staining for tissue/cell annotation

Spatial transcriptomics data