nature genetics

Article

https://doi.org/10.1038/s41588-026-02600-3

Integrated multi-omics identifies distinct macrophage alterations during progression of metabolic dysfunction-associated steatohepatitis

Received: 11 July 2024

Accepted: 14 April 2026

Published online: 18 May 2026

Check for updates

A list of authors and their affiliations appears at the end of the paper

Metabolic dysfunction-associated steatotic liver disease (MASLD) is a chronic condition impacting over 30% population, yet the dynamic changes in macrophage composition from steatosis to steatohepatitis (metabolic dysfunction-associated steatohepatitis, MASH) remain unclear. Here, by integrating single-nucleus transcriptomics, spatial multi-omics and proteomics on human samples, we delineate the evolving landscape of hepatic macrophages across the MASLD spectrum. Our analysis reveals a progressive depletion of Kupffer cells accompanied by the emergence of diverse, phenotypically distinct macrophage subsets. Spatial multi-omics further demonstrates that disease progression toward MASH is marked by an accumulation of antigen-presenting, phagocytic GPNMB$^+$ macrophages, supported by IL32-producing hepatocytes. These macrophages showed an adaptive metabolic and pro-inflammatory phenotype that is tightly regulated by both spatial context and disease stage. Identified macrophage markers enable patient stratification by disease activity and its stage across independent clinical cohorts. Our study sheds light on the diversity of macrophage identities and metabolic-adaptive phenotypes during the progression of MASLD.

Obesity has a pivotal role in triggering low-grade systemic chronic inflammation, including the activation of macrophages, and serves as a major driver for the development of dysfunction-associated steatotic liver disease (MASLD)[1–3]. MASLD, formerly called nonalcoholic fatty liver disease (NAFLD), is a chronic and progressive condition that is strongly associated with features of the metabolic syndrome and is characterized by excessive accumulation of triglycerides in hepatocytes[4]. It encompasses a spectrum of disease states, ranging from steatosis (MASL or 'fatty liver') to the active form of steatohepatitis (metabolic dysfunction-associated steatohepatitis, MASH). The latter is defined by the presence of lobular inflammation and hepatocyte ballooning, along with varying degrees of fibrosis. Consequently, due to its progression to cirrhosis and hepatocellular carcinoma, MASH has become a prominent indication for liver transplantation[5].

MASLD is an inflammatory disease wherein the crosstalk between macrophages and metabolic stimuli has a pivotal role[6]. In particular, MASH has been implemented in disrupting the self-renewal capacity of the Kupffer cells (KCs), the native resident macrophages, leading to a progressive increase of monocyte-derived macrophages[7]. In mouse MASLD models, single-cell transcriptomic analysis has uncovered the emergence of lipid-associated macrophages (LAMs), a macrophage group broadly characterized by the expression of *Trem2* (ref. 8). LAMs were first identified in obese adipose tissue, where they were shown to promote metabolic homeostasis in a Trem2-dependent manner[9]. A similar transcriptional signature was subsequently observed in steatotic mouse liver, where a distinct subset of LAMs expressing osteopontin (*Spp1*) was identified[8]. Notably, *Spp1*$^+$ LAMs have been implicated in attenuating the progression to MASH, suggesting a

e-mail: quentin.anstee@newcastle.ac.uk; schalk.vandermerwe@uzleuven.be; olivier.govaere@kuleuven.be

context-dependent protective role[10]. Consistent with this protective role, *Trem2*$^+$ monocyte-derived macrophages, originated from the bone marrow, have been reported to localize to fibrotic areas to resolve steatohepatitis and fibrosis[11,12]. Furthermore, in end-stage cirrhotic livers, *TREM2*$^+$ *CD9*$^+$ cells have been defined as scar-associated macrophages (SAMs), of which a subset can express *SPP1*, *GPNMB* or *FABP5*, indicating a functional overlap between LAMs and SAMs[13,14]. Although some of these findings have been spatially confirmed in human samples with early steatosis and mixed alcohol intake[15], translating these findings to human MASLD remains incomplete, especially regarding the transition from MASL to MASH. Current models often oversimplify macrophage heterogeneity in progressive MASLD by categorizing them into broad categories—KCs, monocytes or LAMs/SAMs—while overlooking the dynamic, spatially regulated continuum of macrophages[16]. This gap in translation arises, in part, because animal models do not fully capture key features observed in patients with advanced or 'at-risk' MASH, such as advanced fibrosis, lipogranuloma formation, distinctive lobular inflammation and the presence of infiltrating immune cell aggregates in the portal tract[17–19]. As these patients are more likely to progress to cirrhosis and are a key recruitment group for drug trials, it is crucial to resolve the heterogeneity of the inflammatory component. Notably, a higher degree of portal inflammation in MASLD patients has been linked not only with more severe histological features but also with a poor overall survival[20,21]. To address this critical question, we combined single-nuclei RNA sequencing (snRNA-seq), spatial transcriptomics, proteome multiplexing and functional assays to study the changes in macrophage populations and explored the clinical impact in large and well-characterized patient cohorts (Fig. 1a). This study revealed marked heterogeneity within phagocytic GPNMB$^+$ macrophages, with particular enrichment of a metabolically active subset (MetMacs) within the GPNMB$^+$ cells, that transitions dynamically from MASL to MASH. Within this population, two distinct transcriptional clusters were distinguished by high expression of *HS3ST2* and *LPL*, with the latter encompassing a subset exhibiting LAM features. Functionally, GPNMB$^+$ macrophages mediated clearance of cell debris and lipid droplets, leading to lipogranuloma and ceroid macrophage formation. Their phagocytic capacity and inflammatory activation were further modulated by paracrine signals from IL32-producing stressed hepatocytes. Notably, some of the macrophage markers identified in this study stratified patients by disease activity and its stage across different independent clinical cohorts, underscoring their translational relevance and potential utility as biomarkers in patient risk assessment.

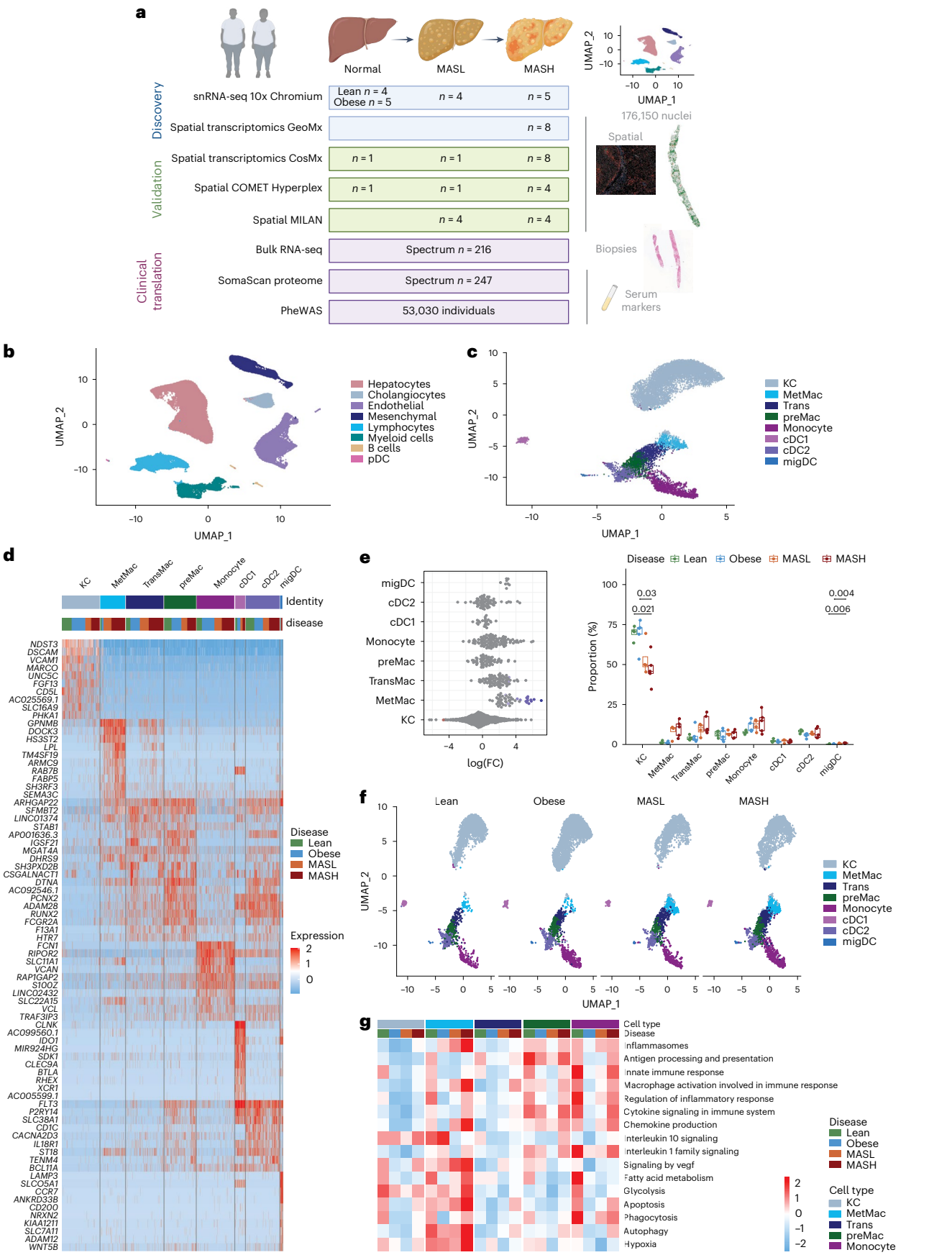
Results

snRNA-seq profiles distinct macrophage subgroups in MASLD

To delineate myeloid cell heterogeneity in the human liver during MASLD progression, we performed snRNA-seq on liver samples from healthy lean (body mass index (BMI) ≤ 25, $n = 4$), healthy obese (BMI ≥ 35, $n = 5$), MASL ($n = 4$) and MASH ($n = 5$; fibrosis stages 1/1/1/2/2) patients (Fig. 1a and Supplementary Table 1). Biopsies underwent histological examination and were categorized into lean, obese with normal liver histology, MASL and MASH. Extraction of nuclei from cryo-preserved samples was followed by the droplet-based 10x Genomics Chromium single-cell gene expression workflow, yielding a dataset comprising 176,150 high-quality nuclei. From this dataset, eight primary hepatic cell types could be annotated, namely, hepatocytes, cholangiocytes, endothelial cells, mesenchymal cells, lymphocytes, myeloid cells, B cells and plasmacytoid dendritic cells (pDCs; Fig. 1b). Of note, the data integrated well across patients and groups, displaying no apparent batch or quality effects (Extended Data Fig. 1a–d and Supplementary Table 2). Further subclustering of hepatocytes revealed a division linked to their zonation, including centrilobular, interzonal and periportal (Extended Data Fig. 1e–g). Interestingly, *IL32*, a marker of ER stress in progressive MASLD[22], was increased in all MASH-hepatocytes (Extended Data Fig. 1h and Supplementary Table 3). In addition, lymphocytes, mesenchymal and endothelial cells were further subclustered based on their gene expression (Extended Data Fig. 2a–c). In-depth analysis of the myeloid populations uncovered eight transcriptionally distinct subpopulations. The most prominent and prevalent among them were KCs, identified by the markers *MARCO* and *CD5L*. The second macrophage population was enriched with various phagocytic, lipid metabolism and cholesterol metabolism markers (*GPNMB*, *HS3ST2*, *LPL*, *FABP5*), and was therefore designated a metabolically active macrophage (MetMac). In addition, transitional macrophages (TransMac; *STAB1*, *MGAT4A*) and premacrophages (*PCNX2*, *ADAM28*, *RUNX2*) exhibited transcriptional similarities with monocytes (*VCAN*, *FCN1*) and mature macrophages. Various conventional dendritic cell (cDC) subtypes were also identified, including cDC1 (*CLNK*, *CLEC9A*), cDC2 (*CD1C*) and migratory DCs (migDC; *LAMP3*, *CCR7*; Fig. 1c,d)[15]. When examining the spectrum of myeloid cell populations from lean to MASH, a gradual shift in cell composition became evident (Fig. 1e). Herein, we observed an MASLD-related increase in migDCs and TransMacs, while in obesity, an elevation of monocytes is suggested. Furthermore, differential abundance testing on neighborhoods of cells using the miloR package revealed a significant decrease in KCs, along with elevated frequencies of MetMacs during MASLD (Fig. 1e,f and Extended Data Fig. 2d). In contrast to KCs, MetMacs exhibited a distinctive response during MASH development, characterized by upregulation of genes associated with fatty acid metabolism, particularly lipoprotein lipase (*LPL*; Fig. 1d,g). In addition, the transcriptional profile of MASH-MetMacs and monocytes shifted toward an inflammatory response, marked by increased activities in chemokine production, cytokine signaling, antigen presentation and macrophage activation (Fig. 1g and Supplementary Table 4). Furthermore, the pro-inflammatory shift in MetMacs during MASH correlates with extended cell–cell interactions with hepatocytes, endothelial cells, mesenchymal cells, *CD4*$^+$ and *CD8*$^+$ T cells, indicating their prominent role in cellular crosstalk during MASH development (Extended Data Fig. 2e and Supplementary Tables 5–8). These snRNA-seq findings indicate changes in cell populations and phenotypic profiles with MASH progression.

Fig. 1 | Single-nucleus map of hepatic macrophages. a, Schematic overview of the discovery, validation and clinical translation cohort and technique. Cohorts were stratified based on liver histology by expert pathologist into normal, MASL and MASH. For individual techniques, the cohorts were further subdivided based on disease stage. Schematic in **a** created in BioRender; Govaere, O. https://BioRender.com/sla50hv (2026). **b**, Annotated UMAP map of 176,150 hepatic cells from lean patients ($n = 4$), obese patients with normal liver histology ($n = 5$), MASL ($n = 4$) and MASH ($n = 5$), showing hepatocytes, cholangiocytes, endothelial cells, mesenchymal cells, lymphocytes, myeloid cells, B cells and pDCs. **c**, UMAP plot of 12,834 myeloid cells (KC, MetMac, TransMac, preMac, cDC, migDC). **d**, Heatmap showing the top ten marker genes for the myeloid populations. Expression is indicated by color saturation, and cluster and disease identity are marked by individual coloring. **e**, Left: MiloR calculated beeswarm plot of differential neighborhood abundance log FC among the lean ($n = 4$), obese ($n = 5$), MASL ($n = 4$) and MASH ($n = 5$) conditions of myeloid populations. Differentially abundant neighborhoods after multiple hypothesis testing at 5% FDR <0.05 are colored. Blue coloration indicates significantly increased and red decreased abundance in the MASLD condition, based on the FDR value. Right: box plot showing the proportion (%) of the myeloid subtypes derived from each condition. Dots indicate individual patients as biological replicates, Holm–Bonferroni adjusted *P* values of two-sided *t* test and the median line indicates the boxes extend from the 25th to 75th percentiles, and the whiskers extend to the minimum and maximum values (right). **f**, Individual UMAP plots of myeloid cells per disease stage. **g**, Heatmap of the average GSVA score of hallmark, KEGG, Reactome and GO BP pathways in the myeloid subtypes per disease, scaled across the subsets and disease. preMac, premacrophage; FC, fold change; GSVA, gene set variation analysis; GO, Gene Ontology; BP, biological process.

https://doi.org/10.1038/s41588-026-02600-3



https://doi.org/10.1038/s41588-026-02600-3

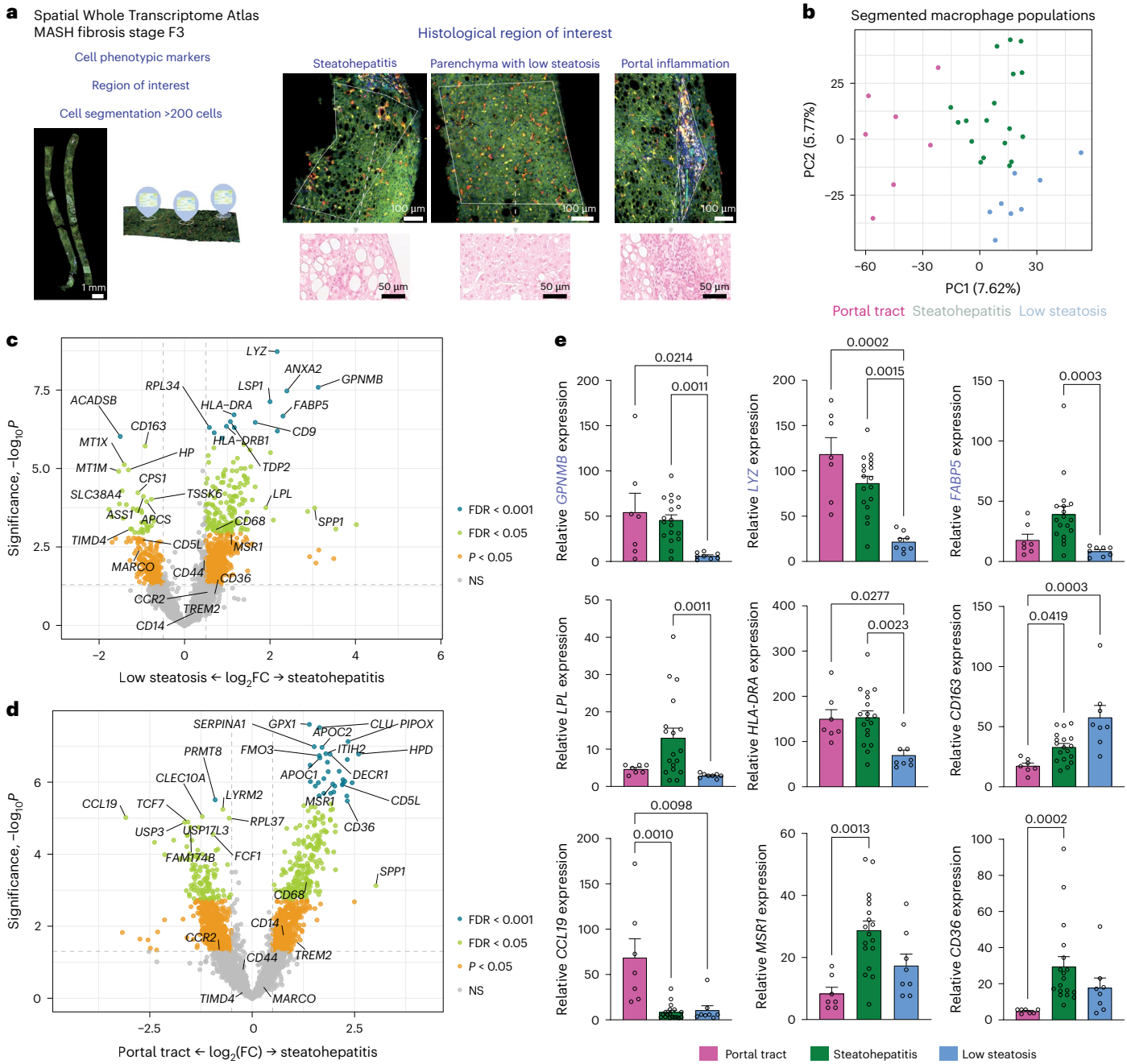


Fig. 2 | Spatial transcriptomics in at-risk MASH patients. a, Workflow of the GeoMx Human Whole Transcriptome Atlas experiment. Human at-risk MASH liver biopsies ($n = 8$) were divided into regions of interest, including steatohepatitis, parenchyma with low steatosis and the portal tract with inflammation. Cells of interest were fluorescently labeled, including macrophage (CD68), immune cell (CD45) and epithelial (panCK) cell populations, followed by sequential segmentation and high-throughput RNA-seq. **b**, PCA plot of the CD68 macrophage population across various regions of interest in all the biopsies. **c**, Volcano plot showing the differentially expressed genes in CD68 segments between the regions of low steatosis ($n = 8$) and steatohepatitis ($n = 18$). **d**, Volcano plot showing the differentially expressed genes in CD68 segments between the regions of the portal tract ($n = 7$) and steatohepatitis ($n = 18$). **e**, Expression levels of selected genes of interest comparing segments from histological regions (portal tract, $n = 7$; steatohepatitis, $n = 18$; and low steatosis, $n = 8$). Data are presented as mean ± s.e.m. (Kruskal–Walis test with Dunn's multiple comparisons test). PCA, principal component analysis. NS, not significant.

Spatial transcriptomics resolves macrophage heterogeneity in at-risk MASH

To elucidate the spatial heterogeneity of macrophage populations in patients with progressive MASH and their association with histopathological features, we used GeoMx Human Whole Transcriptome Atlas profiling on an independent cohort of eight biopsies from 'at-risk' MASH fibrosis stage F3 patients. Patients are classified as 'at-risk' MASH if they have a MASH combined with a fibrosis stage of at least 2 and an NAFLD activity score (NAS) of ≥4 (this is the sum of the histological scores for steatosis, hepatocyte ballooning and lobular inflammation)[23]. Notably, these patients are considered to be at increased risk of further disease progression to cirrhosis. We selected regions of interest based on specific histopathological criteria and used fluorescence markers to segment different macrophage (CD68), immune cell (CD45) and epithelial (pan-keratin (panCK)) cell populations in a sequential manner (Fig. 2a). Whole Transcriptome

https://doi.org/10.1038/s41588-026-02600-3

Atlas RNA profiling was performed on a total of 80 segments, with 77 passing quality controls (Extended Data Fig. 3a). Principal component analysis and hierarchical clustering indicated that the segmented cell populations grouped according to the histopathological region of interest (Fig. 2b and Extended Data Fig. 3b,c). When comparing the parenchymal steatohepatitis-associated (SH-) macrophages with KCs from parenchyma with low steatosis and portal tract (PT-) macrophages, we identified 207 and 348 differentially expressed genes, respectively (false discovery rate (FDR) < 0.05; Fig. 2c,d and Supplementary Tables 9 and 10). PT-macrophages displayed a typical immature/monocyte phenotype, lacking the expression of scavenger receptors (*CD163*, *CD36*, *MSR1*), and expressing lymphocyte-regulatory genes (*CCL19*, *CD48*, *IL7R*, *CCL5*). Compared with KCs in areas of low steatosis, SH-macrophages exhibited features of both monocytes (high expression of *LSP1* and *LYZ*) and mature macrophages (high expression of *MSR1*). Notably, SH-macrophages displayed elevated expression of *GPNMB* alongside an increase in metabolic-related genes, such as *LPL*, *FABP5*, *SPP1* and *ACP5* compared to KCs (Fig. 2c). While genes like *GPNMB*, *LYZ* and *HLA-DRA* were highly expressed in both PT and SH regions, *FABP5*, *LPL*, *MSR1* and CD36 expression was more restricted to SH (Fig. 2e). Pathway analysis revealed an enrichment in phagosome-presenting and antigen-presenting genes in SH-macrophages compared to KCs (Extended Data Fig. 3d). Examining the parenchymal phenotype wherein these SH-macrophages reside, hepatocytes showed an increased expression of ballooned hepatocyte markers *Ubiquitin D* and *SQSTM1/p62*, consistently with the histology, as well as genes linked to cell remodeling (for example, *ANXA2*, *MSN*, *KRT8*; Extended Data Fig. 3e). In addition, CD45$^+$ CD68$^-$ cells from the PT displayed a high expression of immunomodulatory genes as *CCL21* or *TRAF1*, and lower expression of metabolic-associated genes as *SCD* or neutrophil-associated marker *BRI3*, when compared to cells located in the SH region (Extended Data Fig. 3f). To estimate the relative abundance of cell types within each region of interest, SpatialDecon algorithm was applied using a cell profile matrix derived from our snRNA-seq dataset. Validating the accuracy of our segmentation approach, we observed epithelial cell cluster enrichment in panCK regions, myeloid cells in CD68 regions and lymphocyte subtypes in CD45 regions (Extended Data Fig. 3g,h). In addition, we found that the MetMac signature was enriched in areas of steatohepatitis, while the KC cluster signature predominantly mapped to CD68 segments in areas with low steatosis (Extended Data Fig. 3g). Monocyte cluster signatures were enriched in both portal tracts and steatohepatitis regions. Taken together, these findings demonstrate that macrophage gene signatures vary across histopathological locations in patients with at-risk MASH.

Transcriptional heterogeneity within GPNMB$^+$ macrophages

Because *GPNMB* was identified as a top marker of the MetMac cluster in snRNA-seq and showed strong differential expression in GeoMx profiling, we further characterized *GPNMB*$^+$ myeloid cells in detail. Quantification of *GPNMB*$^+$ cells (Unique Molecular Identifier, UMI ≥ 2) across all clusters showed the highest fraction in the MetMac cluster, with additional representation in KC, TransMac, premacrophage, monocyte and cDC2 clusters (Fig. 3a,b). Unsupervised clustering of *GPNMB*$^+$ cells delineated five myeloid subclusters corresponding to distinct cell states (Fig. 3c).

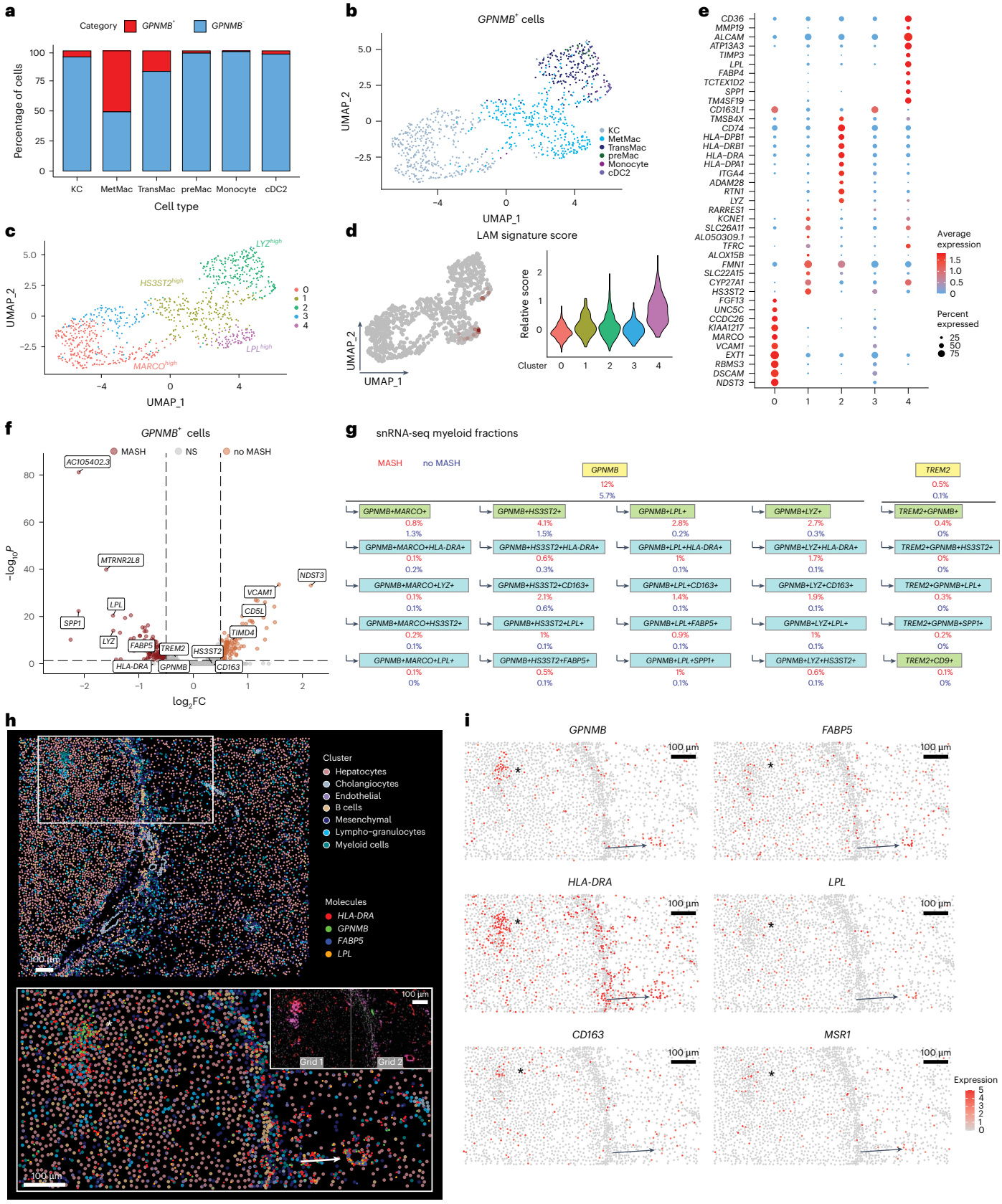
Within the MetMac population, two clusters were defined by elevated expression of *LPL* and *HS3ST2*, indicative of triglyceride hydrolysis and heparan sulfate metabolism, respectively. Additional clusters were enriched for *MARCO* and *LYZ*, reflecting KC-like and monocyte-like identities (Fig. 3c,e). Cluster identifiers, including *LPL* and *FABP5*, along with other spatially resolved transcripts, were uniquely confined to the myeloid compartment, whereas *SPP1* was enriched in nonmyeloid cells such as cholangiocytes (Extended Data Fig. 4a). Feature maps revealed distinct marker dynamics across myeloid subclusters and disease stages (Extended Data Fig. 4b–d), indicating heterogeneity within GPNMB$^+$ cells. To relate *GPNMB*$^+$ populations to LAMs, we applied the originally described Trem2-associated LAM gene set to compute LAM signature scores[9]. A transcriptionally defined subset within the MetMac cluster, and to a lesser extent the TransMac cluster, exhibited elevated LAM signatures (Fig. 3d and Extended Data Fig. 4e,f).

Differential expression analysis of *GPNMB*$^+$ myeloid cells between MASH and non-MASH samples showed increased expression of *LPL*, *FABP5* and *SPP1*, with concomitant downregulation of *CD5L* and *TIMD4* (Fig. 3f and Supplementary Table 11), suggesting changes within the GPNMB$^+$ population composition as disease progresses. To quantify these changes, we assessed the relative abundance of *GPNMB*$^+$ subsets. The proportion of *GPNMB*$^+$ cells within the total myeloid population increased from 5.7% in non-MASH (MASL and normal) to 12% in MASH (Fig. 3g). *GPNMB*$^+$ *LPL*$^+$ and *GPNMB*$^+$ *HS3ST2*$^+$ subsets represented 2.8% and 4.1% of myeloid cells, respectively, with a fraction coexpressing *HLA-DRA* and *CD163* (Fig. 3g). *TREM2*$^+$ *GPNMB*$^+$ populations comprised only 0.4% in MASH, while *GPNMB*$^+$ *LPL*$^+$ *SPP1*$^+$ comprised 1%. Interestingly, *GPNMB*$^+$ *MARCO*$^+$ myeloid cells decreased from 1.3% to 0.8%, indicating a reduction in KC-like cells when progressing toward MASH (Fig. 3g). Collectively, these findings indicate a high heterogeneity in the *GPNMB*$^+$ population. Similar trends were observed in other publicly available data (Extended Data Fig. 4g)[24]. To visualize RNA targets at cellular resolution, we used the CosMx high-plex spatial molecular imaging approach combining a 1000-RNA plex panel with protein segmentation markers (CD68, CD45, panCK and a membranous marker) on nine MASLD samples and one normal liver sample[25]. Cell identification based on Louvain clustering and canonical markers allowed annotation of major cell clusters and digital reconstruction of the histology (Extended Data Fig. 5a,b). This highlighted the change in composition complexity going from early stage MASLD toward end stage (Extended Data Fig. 5c,d), characterized by parenchymal cell loss, structural remodeling and immune cell infiltration (Fig. 3f).

Fig. 3 | Spatial deconvolution and profiling at the cell level identify a metabolic-adaptive macrophage phenotype in *GPNMB*+ macrophages. a, Bar plot visualizing the fraction of *GPNMB*$^+$ (*GPNMB* UMI, count ≥ 2, red) and *GPNMB*$^-$ (*GPNMB* UMI count < 2, blue) cells in the myeloid subclusters of the snRNA-seq data. **b**, UMAP plot of 944 *GPNMB*$^+$ myeloid cells (*GPNMB* UMI count ≥ 2), in the snRNA-seq data, colored based on previously identified myeloid subclusters. **c**, UMAP plot of *GPNMB*$^+$ myeloid cells, colored based on Louvain clustering. **d**, Left: UMAP showing the LAM signature score (*TREM2*, *LIPA*, *LPL*, *CTSB*, *FABP4*, *FABP5*, *LGALS1*, *LGALS3*, *CD9*, *CD36*)[9] in *GPNMB*$^+$ myeloid cells of the snRNA-seq data. Expression is indicated by color saturation (low (gray) to red (high)). Right: violin plots of the LAM signature score in *GPNMB*$^+$ myeloid cells in the snRNA-seq data, the Louvain cluster is marked by individual coloring. **e**, Dot plot of the top ten differentially expressed genes. Expression is indicated by color saturation and the size indicates the percentage of expressing cells in the respective cluster. **f**, Volcano plot showing the differentially expressed genes in *GPNMB*$^+$ macrophages in the snRNA-seq data between MASH and no-MASH patients. **g**, Fractions of myeloid cells from the snRNA-seq data positive (yellow), double-positive (green) or triple-positive (blue) for certain markers based on their gene expression (UMI count ≥ 2), divided into MASH (red) and no-MASH conditions (dark blue). **h**, Representative tissue section of CosMx molecular profiling in end-stage liver disease (total CosMx $n = 10$ of which end-stage $n = 2$), highlighting different hepatic cell populations and markers. Cropped area visualizes *HLA-DRA*, *GPNMB*, *FABP5* and *LPL* probes at a distinct lipogranuloma (arrow) and area of lobular inflammation (asterisk). Insets: fluorescence scans of CD68 (red), CD45 (pink) and panCK (green) protein expression of the lobular inflammation (left) and lipogranuloma area (right). **i**, Spatial plots visualizing the expression level of certain genes in cells at a distinct lipogranuloma (arrow) and area of parenchymal inflammation (asterisk). Expression is indicated by color saturation (from gray to red, representing low to high).

https://doi.org/10.1038/s41588-026-02600-3

Molecular imaging showed that *GPNMB*$^+$ cells were located both in the parenchyma and in the portal tract (Fig. 3f and Extended Data Fig. 5c,d). Interestingly, *GPNMB*$^+$ cells displayed a high expression of *HLA-DRA* and showed heterogeneity depending on the spatial histological location, even within the parenchyma. In lipogranulomas and areas of cell debris, *GPNMB*$^+$ cells displayed high expression of metabolic markers *LPL* and *FABP5* (Fig. 3h,i and Extended Data Fig. 5e). Taken together, these results highlight macrophage heterogeneity in a spatially dependent manner.



https://doi.org/10.1038/s41588-026-02600-3

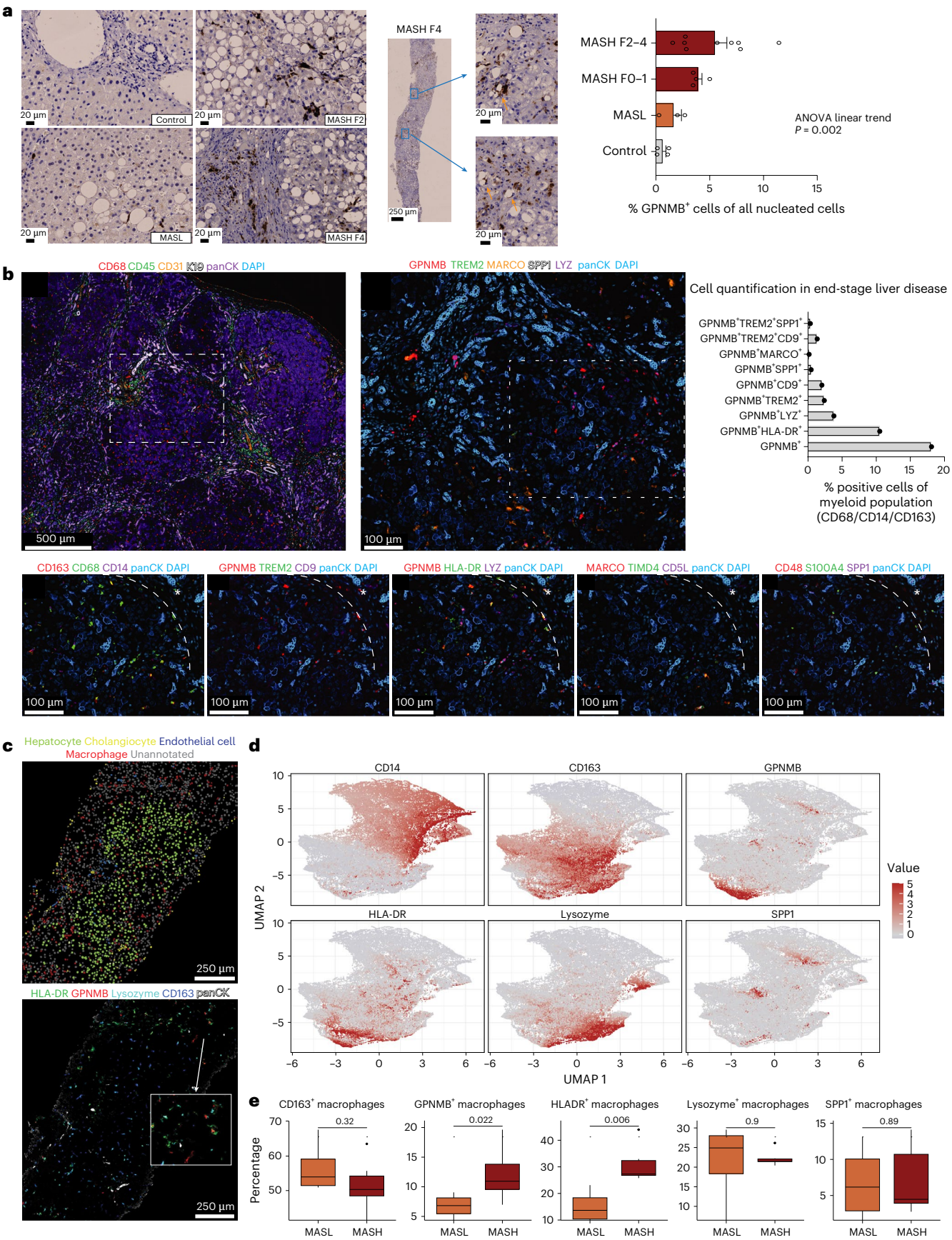


Fig. 4 | Spatial proteomics reveals macrophage heterogeneity during MASLD progression. a, Representative immunohistochemistry images of GPNMB in human liver biopsies of control ($n = 4$), MASL, F2 MASH and F4 MASH patients ($n = 16$). Overview GPNMB staining on MASH F4 biopsy with magnifications indicating granular immunopositivity (arrows) in the portal tract (top) and parenchymal area (bottom). Bar plots indicate the abundance of immunohistochemistry-stained GPNMB$^+$ cells compared with ANOVA trend analysis. Data are presented as mean ± s.e.m. **b**, Representative spatial COMET Hyperplex ($n = 6$) overlayed multicolor fluorescence images of selected markers (left, end-stage liver). Zoom-in of macrophage markers in areas of parenchyma and stromal tissue from the portal tract (asterisk indicates portal tract, dotted line indicates the periportal region). Bar plots show fractions calculated for selected markers in one annotated area of the total myeloid cell population (defined CD68$^+$, CD163$^+$ or CD14$^+$). **c**, Digital reconstruction of a representative human liver sample ($n = 8$) using MILAN, illustrating the major annotated hepatic cell types (top) and a composite fluorescence image of macrophage markers (bottom). Zoom-in depicts a lipogranuloma in the parenchyma (box with arrow indicated). **d**, UMAP showing the expression of various proteins in myeloid cells (CD68$^+$, CD163$^+$ or CD14$^+$). Expression is indicated *z* scores (gray (low) to red (high)). **e**, Box plots comparing the abundance of CD163$^+$, GPNMB$^+$, HLA-DR$^+$, lysozyme$^+$ or SPP1$^+$ macrophages between MASL ($n = 4$) and MASH ($n = 4$). Holm–Bonferroni adjusted *P* values of two-sided *t* test indicated, the boxes extend from the 25th to 75th percentiles and the whiskers extend to the minimum and maximum values. ANOVA, analysis of variance.

Spatial proteomics maps GPNMB$^+$ macrophages in the progression of MASLD

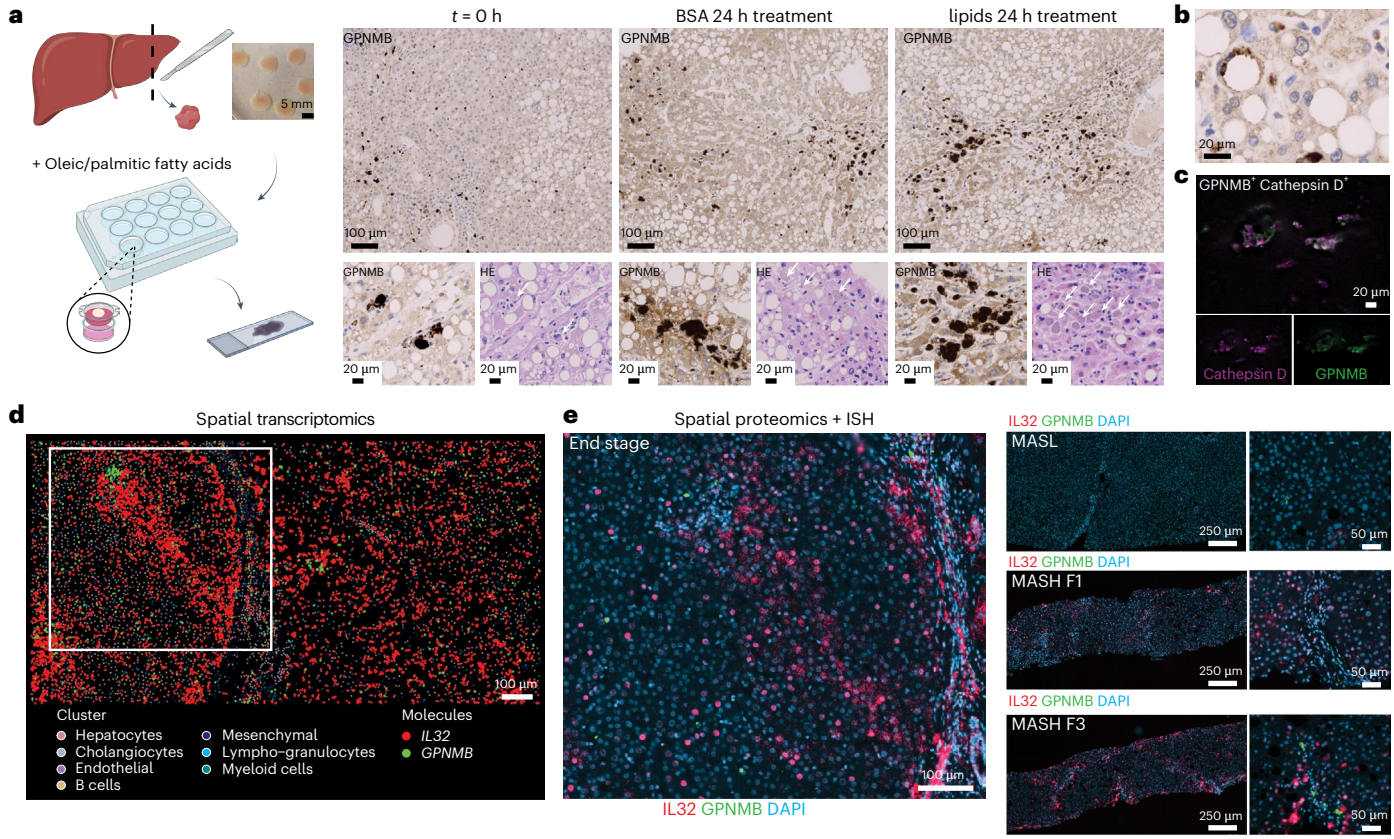
Imperative to our observation on mRNA level is a thorough validation at the protein level. To achieve this, we used standard immunohistochemistry and two approaches of multiplex proteomics—the COMET platform and the multiple iterative labeling by antibody neodeposition (MILAN). The COMET system provides fully automated sequential multiplexing, enabling the visualization of both protein and mRNA molecules within specific regions of interest, whereas the MILAN system is an open platform designed for protein multiplexing across entire biopsy samples[26,27]. Our histopathological analysis involved examining 20 biopsies, including 16 cases of MASLD and 4 controls (Fig. 4a). Notably, in MASL, we observed few GPNMB immunopositive cells in regions with hepatocytic steatosis and within the portal tract. In addition, quantification revealed a progressive rise in the number of GPNMB$^+$ cells toward end-stage MASLD. Some of these cells exhibited large granular immunopositivity in the cytoplasm, particularly in areas with parenchymal cell loss and within the portal tract (Fig. 4a). Furthermore, we applied a spatially resolved single-cell proteomic approach, COMET, to a subset of sequential patient samples ($n = 6$) previously analyzed by CosMx spatial transcriptomics. Antibodies used in the proteome multiplexing were first validated using standard immunohistochemistry and assessed by an expert liver pathologist (Extended Data Fig. 6a). Aligned with our transcriptomics results, the proteome multiplexing revealed substantial heterogeneity among myeloid cells, showing variability for CD163, CD48, CD9, GPNMB, HLA-DR, LYZ, S100A4 and TREM2, as well as the KC markers CD5L, MARCO and TIMD4 (Fig. 4b and Extended Data Fig. 6b). Fractional analysis of myeloid cells (annotated as CD68$^+$, CD163$^+$ or CD14$^+$) within a selected region of end-stage MASLD revealed that GPNMB$^+$ cells comprised 18.1% population, whereas GPNMB$^+$ MARCO$^+$ and GPNMB$^+$ LYZ$^+$ cells, reflecting KC and monocyte features, represented only 0.2% and 3.8%, respectively (Fig. 4b). GPNMB$^+$ TREM2$^+$ cells accounted 2.4% and GPNMB$^+$ SPP1$^+$ cells for 0.5%. Moreover, the number of CD163$^+$ CD5L$^+$ MARCO$^+$ TIMD4$^+$ KCs gradually declined from biopsies with normal histology to MASL and MASH (Extended Data Fig. 6c). Interestingly, in MASH samples, an increase in CD163$^+$ CD5L$^+$ MARCO$^-$ TIMD4$^-$ cells was observed compared to earlier disease stages, reflecting phenotypic shifts within the KC population (Extended Data Fig. 6c). To further investigate the relevance of GPNMB$^+$ cells in the progression from MASL to MASH, we implemented the multiplex staining technique MILAN in an independent cohort of eight MASLD samples. High-dimensional segmentation of the entire biopsies identified a total of 30,000 myeloid cells. These monocytes/macrophages were annotated based on the expression of CD68, CD163 or CD14. Digital reconstruction showed a diverse spatial distribution pattern depending on the cell type and/or marker (Fig. 4c). Uniform manifold approximation and projection (UMAP) analysis revealed that GPNMB, LYZ and HLA-DR expression was predominantly enriched within the CD163$^+$ cells (Fig. 4d). Comparing MASL with MASH, we observed a significant increase in GPNMB$^+$ and HLA-DR$^+$ cells in the latter, accompanied by a trend toward reduced CD163$^+$ cells, while the fraction of LYZ$^+$ and SPP1$^+$ macrophages remained unchanged (Fig. 4e). In addition, the GPNMB$^+$ macrophage cluster exhibited an elevated HLA-DR$^+$ frequency in MASH compared to MASL, along with an increase in CD163$^+$ HLA-DR$^+$ double-positive cells (Extended Data Fig. 6d,e). The distinct GPNMB$^+$ macrophages were detected in both the portal tract and hepatic parenchyma (Supplementary Video). Spatial proximity analysis further revealed that 9% of GPNMB$^+$ macrophages were located within 10 µm of cholangiocytes (Extended Data Fig. 6f), indicating portal infiltration, consistent with our broader findings. Collectively, these results highlight protein-level alterations within the GPNMB$^+$ macrophage population during MASH progression, confirming our transcriptomics results.

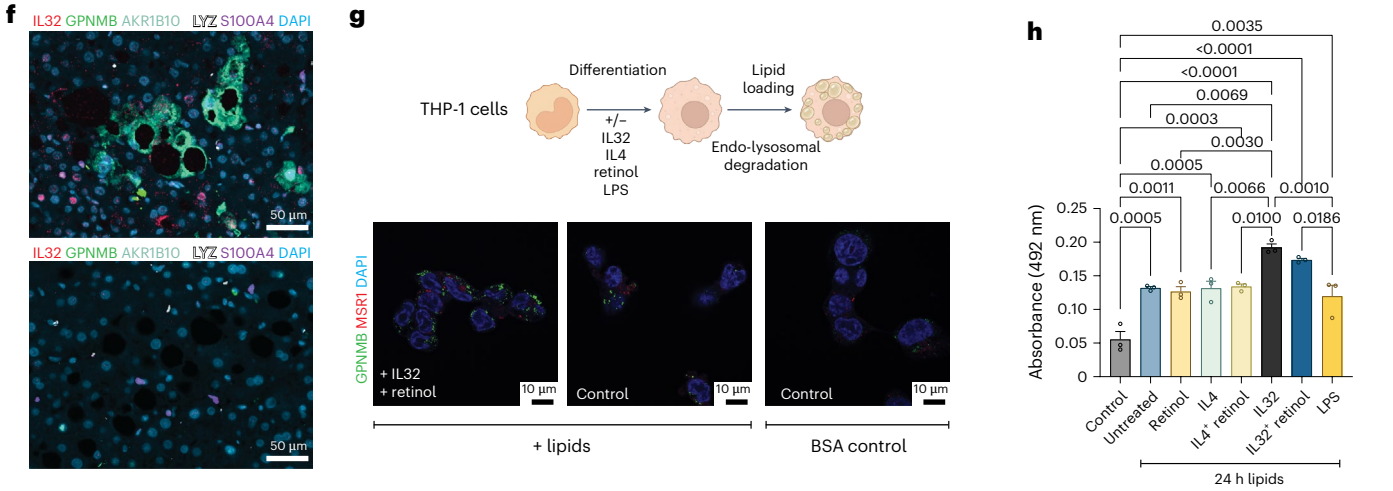
Fig. 5 | IL32-producing hepatocytes promote phagocytic and pro-inflammatory properties of GPNMB$^+$ macrophages. a, Schematic representation of PCLS setup. PCLS were cultured overnight and treated with BSA or lipids (1 mM OA and 500 µM PA) for 24 h ($n = 3$ technical repeats per condition). Representative immunohistochemistry images of GPNMB in PCLS after baseline (0 h) and 24 h after BSA or lipid treatment. Regions of interest with GPNMB aggregates are zoomed in (bottom left), and ceroid macrophage formation is shown in sequential H&E images (bottom right). **b**, Representative GPNMB immunohistochemistry image of early lipogranuloma formation after lipid treatment in PCLS. **c**, High-resolution confocal imaging of GPNMB (green), cathepsin D (magenta) in end-stage human liver (MASH F4, $n = 3$). GPNMB and cathepsin D overlap is highlighted in white. **d**, Representative tissue section of CosMx molecular profiling ($n = 10$) in end-stage liver disease ($n = 2$), highlighting different hepatic cell populations, and *IL32* and *GPNMB* RNA probes. **e**, Consecutive tissue section of COMET Hyperplex ($n = 6$) visualizing GPNMB protein (green) and *IL32* in situ hybridization RNA probe (red) in end-stage MASH ($n = 2$; left), indicating the zonation of *IL32*. Right: representative images and zoom-ins of MASLD biopsies with disease stage MASL ($n = 1$), MASH F0/F1 ($n = 1$) and MASH F3 ($n = 1$). **f**, Representative tissue section of COMET Hyperplex visualizing AKR1B10 protein (aquamarine) to indicate regions of ballooned hepatocytes, *IL32* RNA probe (red) and selected macrophage markers in end-stage MASH (top) and MASH F0 (bottom). **g**, Schematic representation of the experimental setup. THP-1 cells were differentiated into macrophages with or without IL32, retinol, IL4 and LPS, followed by lipid loading with PA and OA (top). High-resolution confocal imaging of GPNMB (green), MSR1 (red) and DAPI (blue) in cells treated with IL32 and retinol or PMA (control), after lipid loading with 200 µM PA and 400 µM OA or treated with 1% BSA ($n = 3$ per condition). **h**, Lipid content was quantified in differentiated THP-1 cells conditioned with IL32, retinol, IL4 and/or LPS, followed by 24 lipid loading (200 µM PA and 400 µM OA) or BSA control ($n = 3$ per condition). Absorbance of Oil Red O staining was measured at 492 nm. Data are presented as mean ± s.e.m. (one-way ANOVA with Tukey's multiple comparisons test). **i**, Endo-lysosomal degradation was assessed in 10 ng ml^{-1} PMA-treated THP-1 cells exposed to 50 or 100 ng ml^{-1} IL32 for 24 h, followed by incubation with fluorescence-labeled silica beads. In real time, the excitation and emission wavelength were measured, and the RFUs were determined. **j**, Relative gene expression in THP-1 cells treated with 1% BSA or 1 mM PA, and conditioned with or without IL32, measured through RT-qPCR ($n = 3$ per condition). Data are presented as mean ± s.e.m. (one-way ANOVA with Tukey's multiple comparisons test). OA, oleic acid; RFUs, relative fluorescence units; PMA, phorbol 12-myristate-13-acetate; H&E, hematoxylin and eosin. Schematics in **a** and **g** created in BioRender; Govaere, O. https://BioRender.com/msfl43p (2025).

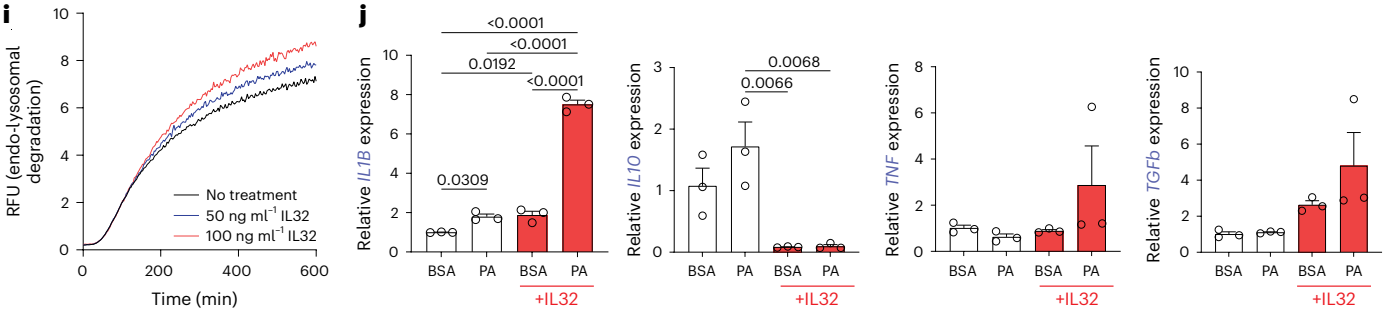
https://doi.org/10.1038/s41588-026-02600-3

IL32-producing hepatocytes drive phenotype of GPNMB$^+$ macrophages

We then used precision-cut liver slices (PCLS) as an ex vivo model to explore the function of GPNMB$^+$ macrophages. Human liver tissue was used to generate PCLS, which were cultured overnight and subsequently exposed to either lipids or BSA (control) for 24 h ($n = 3$). At baseline ($t = 0$ h), histology showed steatohepatitis with scattered GPNMB$^+$ cells (Fig. 5a). After lipid treatment, GPNMB$^+$ cells formed aggregates within congested hepatocyte regions, appeared enlarged compared with control or baseline and sequential hematoxylin and eosin staining showed increased ceroid macrophage formation, indicative of enhanced clearance (Fig. 5a). Lipid loading also induced early







lipogranuloma formation (Fig. 5b). To assess phagocytic activity, we performed GPNMB/cathepsin D (CATD) costaining in end-stage MASLD patient Formalin-Fixed Paraffin-Embedded (FFPE) samples. High-resolution imaging revealed extensive colocalization of GPNMB$^+$ and CATD$^+$ vesicles, consistent with active phagocytosis (Fig. 5c). To investigate how the microenvironment drives macrophage function in MASH, we focused on IL32, which was identified in both our snRNA-seq and GeoMx datasets as differentially expressed in hepatocytes from MASH samples. Notably, elevated hepatic IL32 levels have previously been reported in individuals with severe MASLD[22,28]. CosMx spatial profiling confirmed robust *IL32* expression in hepatocytes from end-stage MASLD (Fig. 5d). These spatial associations were validated on consecutive sections using COMET multiplexing, combining IL32 mRNA probes with GPNMB antibody staining (Fig. 5e). Across biopsy samples, we noted a progressive increase in *IL32*$^+$ hepatocytes and GPNMB$^+$ macrophages with disease advancement, especially within portal tracts and periportal zones in early stages. In addition, *IL32* expression was elevated in areas with abundant ballooned hepatocytes, as marked by AKR1B10, a retinol metabolism enzyme (Fig. 5f). To validate these findings in vitro, we conditioned differentiated THP-1 myeloid cells with interleukin (IL) 4, IL32, retinol or lipopolysaccharide, alongside lipid loading with oleic and palmitic acids (PAs) for 24 h (Fig. 5g). Notably, IL32-treated cells exhibited enhanced phagocytic uptake of free lipid particles after 24 h, compared to the other conditions (Fig. 5h), accompanied by increased/enlarged cytoplasmic MSR1 and GPNMB vesicle formation (Fig. 5g). Moreover, IL32 treatment stimulated dose-dependent endo-lysosomal degradation activity (Fig. 5i). Particularly PA has previously been reported to induce a pro-inflammatory response in hepatic macrophages[29]. Our in vitro data showed that conditioning THP-1 cells with IL32 enhanced PA-induced pro-inflammatory *IL1B* expression while reducing the anti-inflammatory marker *IL10* (Fig. 5j)[2]. Collectively, these results underscore a pivotal role for hepatocyte-derived IL32 in sustaining the phagocytic capacity of GPNMB$^+$ macrophages and modulating their pro-inflammatory response, linking hepatocyte signaling to macrophage function in MASH progression.

Implications of GPNMB$^+$ macrophages in clinical cohorts

To understand the clinical implications of macrophage heterogeneity, we conducted an extensive analysis of marker expression across various omic datasets—transcriptomics data from a well-characterized cohort of 206 MASLD patients and 10 controls[22], serum SomaLogic proteomics data from 247 biopsy-proven MASLD patients[6,22] and phenome-wide association study (PheWAS) analysis based on OLINK proteomics data from a large UK Biobank (UKB) cohort (53,030 participants)[30]. When we examined the expression of selected genes relevant to our spatial transcriptomics, a clear segmentation was seen between MASL and MASH for genes including *GPNMB*, *LYZ*, *LPL*, *HLA-DRA*, *HLA-DRB1*, *FABP5* and *LSP1* (Fig. 6a and Extended Data Fig. 7a). Markers associated with the portal macrophage phenotype (for example, *CCL19* and *CLEC10A*) were more highly expressed in advanced MASLD, while *CD163* expression decreased. Hepatocytic markers AKR1B10 and IL32 showed a gradual increase with disease progression (Extended Data Fig. 7a). Using our snRNA-seq data to deconvolve and estimate the proportion of myeloid cells in the bulk RNA-seq data, we revealed that the proportion of MetMacs and monocytes significantly increased with disease progression, whereas KC proportion gradually decreased (Fig. 6b)[31]. Similar trends were observed when assessing the fibrosis stage and the histological MASLD activity score (the sum of steatosis, hepatocyte ballooning and lobular inflammation scoring; Extended Data Fig. 7b). To explore the utility of macrophage markers in predicting disease activity, we performed logistic regression analysis. The composite model in the transcriptomics cohort identified the clinical variable 'age' and the expression levels of *ACP5*, *CD163*, *LPL* and *MSR1* as independent variables to classify patients with a high disease activity (Extended Data Fig. 7c). In this model, *LPL* expression levels displayed a strong contribution. Functionally, *LPL* CRISPR–Cas9 knockout using the THP-1 cell line resulted in a foamy macrophage phenotype after 24 h of lipid loading, when compared to the scramble guides (Extended Data Fig. 7d). Next, we investigated whether hepatic changes correlate with systemic alterations in circulating proteins. Using serum proteomics data from biopsy-proven MASLD patients, we stratified individuals based on 'at-risk' MASH, advanced fibrosis, the presence of MASH or high disease activity (NAS ≥ 4). The strongest significant changes in circulating markers were observed for soluble CD163, GPNMB and MSR1 (Fig. 6c). Our PheWAS analysis, encompassing a large cohort of 53,030 participants, revealed a strong association between soluble GPNMB and features of the metabolic syndrome, specifically diabetes mellitus and hypoglycemia, while CD163 was significantly associated with diabetes and chronic liver diseases (Extended Data Fig. 8). Taken together, our comprehensive analysis sheds light on the intricate interplay among macrophage heterogeneity, disease progression and circulating protein dynamics in MASLD (Extended Data Fig. 9).

Discussion

In this study, we examined hepatic macrophage heterogeneity during MASLD progression, with a focus on the microenvironmental influence. We identified substantial diversity within phagocytic GPNMB$^+$ macrophages, particularly the emergence of a metabolically active subset (MetMacs) that shifts dynamically from MASL to MASH. Notably, LAM-associated signatures and markers were detected only in a fraction of this MetMac, underscoring their distinction from previously described LAM/SAM populations[8,13]. In addition, our ex vivo and in vitro data indicate a pro-inflammatory, phagocytic role for GPNMB$^+$ macrophages in lipid handling, in contrast to TREM2$^+$ macrophages, which have been involved in MASH and fibrosis resolution[11,12]. This distinction underscores that hepatic macrophages in MASLD are more heterogeneous than previously appreciated, and that existing nomenclature does not adequately capture their complexity. Furthermore, our findings highlight the importance of spatial localization and hepatocyte-derived signaling, specifically IL32, in modulating macrophage activation states.

Within MetMacs, we identified two major populations characterized by high expression of *LPL* and *HS3ST2*. Furthermore, *LPL*, a triglyceride hydrolase essential for lipoprotein processing[32], marked macrophages enriched in steatohepatitis and was required to prevent foamy macrophage formation, as confirmed by our CRISPR–Cas9 perturbation. These findings are consistent with previously described mouse models of atherosclerosis[33,34]. Notably, both *LPL* and *MSR1* independently predicted high disease activity in our larger MASLD cohort. Recently, we have demonstrated that MSR1 promotes lipid droplet uptake and lipogranuloma formation in MASH, highlighting adaptive lipid-clearing programs within these areas of steatohepatitis[29]. *HS3ST2*, a regulator of heparan sulfate biosynthesis, was selectively upregulated in MetMacs and has previously been associated with a pro-inflammatory phenotype in atherosclerosis, implicating its role in lipid-rich, inflamed niches[35]. Supporting this adaptive axis, hepatocyte-derived IL32 enhanced macrophage phagocytic capacity and the pro-inflammatory response to PA (for example, *IL1B* expression), thereby linking metabolic stress to immune reprogramming. IL32 is a multifunctional cytokine associated with disease progression in patients with MASLD[22,28], and has a direct role in immune cell differentiation[36]. Moreover, IL32 has been implicated in modulating inflammatory and fibrogenic responses in vitro[37,38].

GPNMB, often used as a lineage or annotation marker[15], also has functional roles in phagocytosis and debris clearance, with its cleaved form exerting immunomodulatory effects on myeloid and lymphoid cells[39,40]. Its expression increases in macrophages under lysosomal stress, particularly in lipotoxic environments, and has been used to identify MetMacs in MASLD adipose tissue[41,42]. In our study, GPNMB$^+$

https://doi.org/10.1038/s41588-026-02600-3

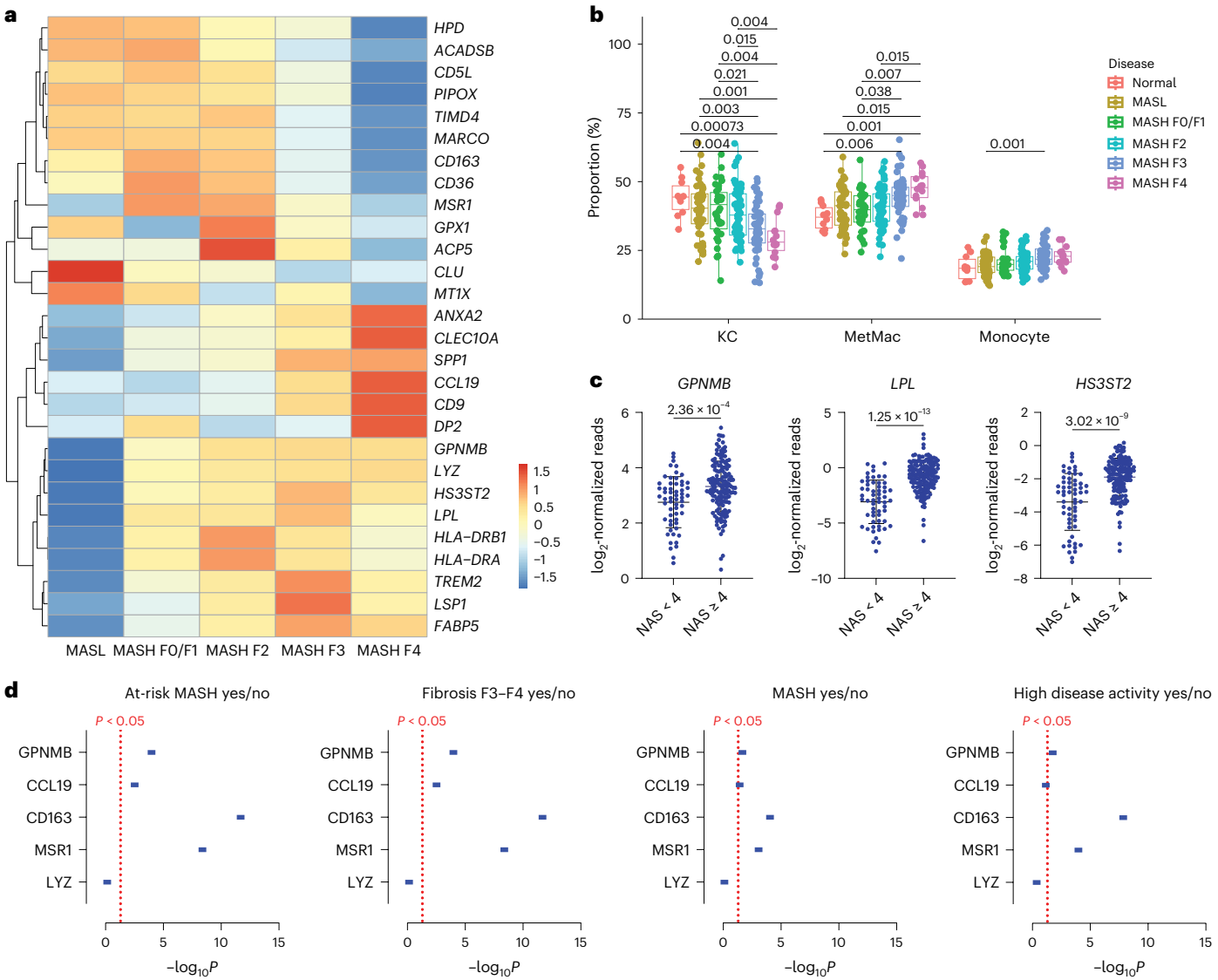


Fig. 6 | Clinical implications of macrophages in MASLD based on hepatic transcriptomics and serum proteomics. a, Heatmap of 28 genes, which were identified by the spatial transcriptomics, indicating the expression from MASL to the different MASH stages using RNA-seq data of human liver biopsies ($n = 206$). **b**, Estimating the proportion of KC, MetMacs and monocytes in the RNA-seq data with the CIBERSORTx deconvolution algorithm using the snRNA-seq data, grouped based on disease and fibrosis stage (MASLD, $n = 206$; controls, $n = 10$). Dots indicate individual patients as biological replicates, Holm–Bonferroni adjusted P values of two-sided t test and median line indicated, the boxes extend from the 25th to 75th percentiles and the whiskers extend to the minimum and maximum values. **c**, $\log_2$-normalized mRNA reads from NAS < 4 or ≥4 samples using RNA-seq data of human liver biopsies ($n = 206$). Data are presented as mean ± s.e.m. P values of two-sided Mann–Whitney U test indicated. **d**, Serum SomaLogic proteomics data from MASLD patients with histologically-proven MASLD ($n = 247$), highlighting the significant ($P < 0.05$, dashed red line) correlation of identified macrophage markers and clinical features. Serum samples were measured with the SomaScan Assay (v4.0). RFUs were $\log_{10}$ transformed. Patients were stratified based on the histological presence of MASH, advanced fibrosis stage 3 or 4, high disease activity or 'at-risk' MASH. High disease activity was defined as a NAS ≥ 4; this is the sum of the histological scores for steatosis, hepatocyte ballooning and lobular inflammation. Data are presented as P value of two-sided Mann–Whitney U test for each marker.

macrophages were observed to cluster within the portal tracts of MASLD livers, suggesting a localized role in modulating inflammation, particularly in relation to infiltrating monocytes and lymphocytes. Furthermore, we observed a depletion of resident KCs with a parallel expansion of inflammatory GPNMB$^+$ HLA-DR$^+$ macrophages. This mirrors mouse studies showing that loss of periportal MARCO$^+$ IL10$^-$ producing KCs, critical for pathogen sequestration and immune regulation, exacerbates steatohepatitis[43]. Consistent with this, we found a progressive decline of MARCO$^+$ CD5L$^+$ TIMD4$^+$ KCs, alongside fragmented populations expressing only individual markers. These findings highlight the adaptive remodeling of macrophage subsets in response to lipotoxic stress and may connect innate and adaptive immunity signaling.

In addition, we found that soluble GPNMB and CD163 levels were elevated in patients with advanced MASH, and our PheWAS analysis, based on 53,030 participants from the UKB, indicated associations with features of the metabolic syndrome. This could suggest a systemic effect, which would be in line with previous mouse studies. Gpnmb secreted from the liver has been reported to promote lipogenesis in white adipose tissue and to aggravate obesity and insulin resistance[44]. Aging, a risk factor for advanced MASLD, has been described to increase the expression of Gpnmb in various organs in mice, including the liver[45].

This study has several limitations that should be acknowledged. Here snRNA-seq is subject to technical limitations, especially a higher incidence of technical dropouts due to the lower abundance of mRNA within isolated nuclei compared to whole cells. Moreover, the high

heterogeneity within the macrophage populations can make it difficult to precisely define cluster boundaries, limiting the clustering accuracy. This must be considered when drawing comparisons to scRNA-seq data. The GeoMx spatial transcriptomics platform uses an enrichment approach based on fluorescence, which, given the complex anatomical structure of the liver, may inadvertently include neighboring cells upon cell segmentation. The COMET multiplexing platform, when used with our probe panel, produced focal nuclear background staining, likely due to binding to the DNA, and should be interpreted carefully. In addition, the control samples, although showing histologically normal liver tissue, were still derived from patients rather than from truly healthy individuals. Some omics analyses were conducted on relatively small patient cohorts, meaning that validations in larger, independent cohorts are warranted. Although our findings highlight heterogeneity among GPNMB$^+$ macrophages, further studies are needed to elucidate the distinct cell states and adaptive responses to lipids, particularly in relation to the *LPL* and *HS3ST2* subsets.

In conclusion, our study sheds new light on the diversity of macrophage identities and their metabolic phenotypes during the progression of MASLD. We showed that spatial and temporal changes occur in mature macrophage markers, which adds complexity to the old dogma of macrophage-monocytes in chronic liver disease.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-026-02600-3.

References

1. Anstee, Q. M., Reeves, H. L., Kotsiliti, E., Govaere, O. & Heikenwalder, M. From NASH to HCC: current concepts and future challenges. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 411–428 (2019).
2. Hammerich, L. & Tacke, F. Hepatic inflammatory responses in liver fibrosis. *Nat. Rev. Gastroenterol. Hepatol.* **20**, 633–646 (2023).
3. Barreby, E., Chen, P. & Aouadi, M. Macrophage functional diversity in NAFLD—more than inflammation. *Nat. Rev. Endocrinol.* **18**, 461–472 (2022).
4. Rinella, M. E. et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *J. Hepatol.* **79**, 1542–1556 (2023).
5. Wong, V. W., Ekstedt, M., Wong, G. L. & Hagstrom, H. Changing epidemiology, global trends and implications for outcomes of NAFLD. *J. Hepatol.* **79**, 842–852 (2023).
6. Govaere, O. et al. A proteo-transcriptomic map of non-alcoholic fatty liver disease signatures. *Nat. Metab.* **5**, 572–578 (2023).
7. Tran, S. et al. Impaired Kupffer cell self-renewal alters the liver response to lipid overload during non-alcoholic steatohepatitis. *Immunity* **53**, 627–640 (2020).
8. Remmerie, A. et al. Osteopontin expression identifies a subset of recruited macrophages distinct from Kupffer cells in the fatty liver. *Immunity* **53**, 641–657 (2020).
9. Jaitin, D. A. et al. Lipid-associated macrophages control metabolic homeostasis in a Trem2-dependent manner. *Cell* **178**, 686–698 (2019).
10. Han, H. et al. Macrophage-derived osteopontin (SPP1) protects from nonalcoholic steatohepatitis. *Gastroenterology* **165**, 201–217 (2023).
11. Hendrikx, T. et al. Soluble TREM2 levels reflect the recruitment and expansion of TREM2$^+$ macrophages that localize to fibrotic areas and limit NASH. *J. Hepatol.* **77**, 1373–1385 (2022).
12. Ganguly, S. et al. Lipid-associated macrophages' promotion of fibrosis resolution during MASH regression requires TREM2. *Proc. Natl Acad. Sci. USA* **121**, e2405746121 (2024).
13. Ramachandran, P. et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**, 512–518 (2019).
14. Fabre, T. et al. Identification of a broadly fibrogenic macrophage subset induced by type 3 inflammation. *Sci. Immunol.* **8**, eadd8945 (2023).
15. Guilliams, M. et al. Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches. *Cell* **185**, 379–396 (2022).
16. De Ponti, F. F., Liu, Z. & Scott, C. L. Understanding the complex macrophage landscape in MASLD. *JHEP Rep.* **6**, 101196 (2024).
17. Gallage, S. et al. A researcher's guide to preclinical mouse NASH models. *Nat. Metab.* **4**, 1632–1649 (2022).
18. Kotsiliti, E. et al. Intestinal B cells license metabolic T-cell activation in NASH microbiota/antigen-independently and contribute to fibrosis by IgA-FcR signalling. *J. Hepatol.* **79**, 296–313 (2023).
19. Guillot, A. et al. Mapping the hepatic immune landscape identifies monocytic macrophages as key drivers of steatohepatitis and cholangiopathy progression. *Hepatology* **78**, 150–166 (2023).
20. Brunt, E. M. et al. Portal chronic inflammation in nonalcoholic fatty liver disease (NAFLD): a histologic marker of advanced NAFLD-clinicopathologic correlations from the nonalcoholic steatohepatitis clinical research network. *Hepatology* **49**, 809–820 (2009).
21. Angulo, P. et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* **149**, 389–397 (2015).
22. Govaere, O. et al. Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. *Sci. Transl. Med.* **12**, eaba4448 (2020).
23. Kleiner, D. E. et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–1321 (2005).
24. Gribben, C. et al. Acquisition of epithelial plasticity in human chronic liver disease. *Nature* **630**, 166–173 (2024).
25. He, S. et al. High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat. Biotechnol.* **40**, 1794–1806 (2022).
26. Bolognesi, M. M. et al. Multiplex staining by sequential immunostaining and antibody removal on routine tissue sections. *J. Histochem. Cytochem.* **65**, 431–444 (2017).
27. Antoranz, A. et al. Mapping the immune landscape in metastatic melanoma reveals localized cell-cell interactions that predict immunotherapy response. *Cancer Res.* **82**, 3275–3290 (2022).
28. Baselli, G. A. et al. Liver transcriptomics highlights interleukin-32 as novel NAFLD-related cytokine and candidate biomarker. *Gut* **69**, 1855–1866 (2020).
29. Govaere, O. et al. Macrophage scavenger receptor 1 mediates lipid-induced inflammation in non-alcoholic fatty liver disease. *J. Hepatol.* **76**, 1001–1012 (2022).
30. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
31. Newman, A. M. et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).
32. Pingitore, P. et al. Identification and characterization of two novel mutations in the LPL gene causing type I hyperlipoproteinemia. *J. Clin. Lipidol.* **10**, 816–823 (2016).
33. Babaev, V. R., Patel, M. B., Semenkovich, C. F., Fazio, S. & Linton, M. F. Macrophage lipoprotein lipase promotes foam cell formation and atherosclerosis in low density lipoprotein receptor-deficient mice. *J. Biol. Chem.* **275**, 26293–26299 (2000).

34. Takahashi, M. et al. Macrophage lipoprotein lipase modulates the development of atherosclerosis but not adiposity. *J. Lipid Res.* **54**, 1124–1134 (2013).
35. Tan, J. et al. Single-cell transcriptomics reveals crucial cell subsets and functional heterogeneity associated with carotid atherosclerosis and cerebrovascular events. *Arterioscler. Thromb. Vasc. Biol.* **43**, 2312–2332 (2023).
36. Netea, M. G. et al. Interleukin-32 induces the differentiation of monocytes into macrophage-like cells. *Proc. Natl Acad. Sci. USA* **105**, 3515–3520 (2008).
37. Park, J. S. et al. A1AT dysregulation of metabolically stressed hepatocytes by Kupffer cells drives MASH and fibrosis. *Exp. Mol. Med.* **57**, 450–465 (2025).
38. Sasidharan, K. et al. IL32 downregulation lowers triglycerides and type I collagen in di-lineage human primary liver organoids. *Cell Rep. Med.* **5**, 101352 (2024).
39. Ramachandran, P. et al. Differential Ly-6C expression identifies the recruited macrophage phenotype, which orchestrates the regression of murine liver fibrosis. *Proc. Natl Acad. Sci. USA* **109**, E3186–E3195 (2012).
40. Lazaratos, A. M., Annis, M. G. & Siegel, P. M. GPNMB: a potent inducer of immunosuppression in cancer. *Oncogene* **41**, 4573–4590 (2022).
41. Gabriel, T. L. et al. Lysosomal stress in obese adipose tissue macrophages contributes to MITF-dependent Gpnmb induction. *Diabetes* **63**, 3310–3323 (2014).
42. Boesch, M. et al. Adipose tissue macrophage dysfunction is associated with a breach of vascular integrity in NASH. *J. Hepatol.* **80**, 397–408 (2024).
43. Miyamoto, Y. et al. Periportal macrophages protect against commensal-driven liver inflammation. *Nature* **629**, 901–909 (2024).
44. Gong, X. M. et al. Gpnmb secreted from liver promotes lipogenesis in white adipose tissue and aggravates obesity and insulin resistance. *Nat. Metab.* **1**, 570–583 (2019).
45. Schaum, N. et al. Ageing hallmarks exhibit organ-specific temporal signatures. *Nature* **583**, 596–602 (2020).

Markus Boesch[1,2,3,4], Seray Anak[2,29], Dania El Abyad[2,29], Tessa Ostyn[2], Asier Antoranz[2,3], Trieu My Van[5], Daniel Newhouse[5], Caitlyn Myers[6], Lukas Van Melkebeke[1,7], Jeremy Palmer[8], Najmeh Saffarzadeh[6], Jonathan Sai-Hong Chui[2,3], Rita Feio-Azevedo[1], Lena Smets[1], Gautam Shankar[2,3], Nikolina Dubroja Lakić[2,3], Paula Longás Calvo[3,4], Thierry Voet[3,4], James Clark[8], Simon Cockell[8], Matthias Lannoo[9], Ellen Deleus[9], Joris Jaekers[9], Halit Topal[9], Baki Topal[9], Kai Markus Schneider[10,11,12,13], Michele Vacca[14,15], Michael Allison[16], Mattias Ekstedt[17], Jerome Boursier[18], Elisabetta Bugianesi[19], Vlad Ratziu[20], Frederik De Smet[2,3], Bram Boeckx[21,22], Ann K. Daly[8], Jörn M. Schattenberg[23], Jef Verbeek[1,7], Francesca Bosisio[2,3], Diether Lambrechts[21,22], Tania Roskams[2], Carolin V. Schneider[10], Sven Francque[24,25], Hannelie Korf[1], Anetta Härtlova[6,26], Dina Tiniakos[8,27], Quentin M. Anstee[8,28,30] ✉, Schalk van der Merwe[1,7,30] ✉ & Olivier Govaere[2,3,8,30] ✉

[1]Laboratory of Liver Immunology, Leuven Center of Molecular Liver Sciences, Hepatology Research Unit, CHROMETA Department, KU Leuven, Leuven, Belgium. [2]Department of Imaging and Pathology, Translational Cell and Tissue Research, KU Leuven and University Hospitals Leuven, Leuven, Belgium. [3]KU Leuven Institute for Single Cell Omics (LISCO), KU Leuven, Leuven, Belgium. [4]Laboratory of Reproductive Genomics, Department of Human Genetics, KU Leuven, Leuven, Belgium. [5]Bruker Spatial Biology, Seattle, WA, USA. [6]Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Gothenburg, Sweden. [7]Department of Gastroenterology and Hepatology, UZ Leuven, Leuven, Belgium. [8]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK. [9]Department of Abdominal Surgery, UZ Leuven, Leuven, Belgium. [10]Medical Clinic III, Gastroenterology, Metabolic diseases and Intensive Care, University Hospital RWTH Aachen, Aachen, Germany. [11]Department of Medicine 1, University Hospital Carl Gustav Carus Dresden, TUD Dresden University of Technology, Dresden, Germany. [12]Else Kroener Fresenius Center for Digital Health, Medical Faculty Carl Gustav Carus, TUD Dresden University of Technology, Dresden, Germany. [13]Center for Regenerative Therapies Dresden (CRTD), TUD Dresden University of Technology, Dresden, Germany. [14]Roger Williams Institute of Liver Studies, School of Immunology & Microbial Sciences, Faculty of Life Sciences and Medicine, King's College London, Foundation for Liver Research and King's College Hospital, London, UK. [15]Clinica Medica 'C. Frugoni' and Department of Interdisciplinary Medicine, Aldo Moro University of Bari, Bari, Italy. [16]Cambridge Liver Unit, Cambridge University Hospitals NHS Foundation Trust, Cambridge NIHR Biomedical Research Centre, Cambridge, UK. [17]Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden. [18]Hepatology Department, Angers University Hospital, Angers, France. [19]Department of Medical Sciences, Division of Gastro-Hepatology, A.O. Città della Salute e della Scienza di Torino, University of Turin, Turin, Italy. [20]Assistance Publique-Hôpitaux de Paris, Hôpital Pitié Salpêtrière, Sorbonne University, Institute of Cardiometabolism and Nutrition (ICAN), Paris, France. [21]Laboratory for Translational Genetics, Department of Human Genetics, KU Leuven, Leuven, Belgium. [22]VIB Center for Cancer Biology, Leuven, Belgium. [23]Department of Internal Medicine II, Saarland University Medical Center, Homburg, Germany. [24]Department of Gastroenterology and Hepatology, Antwerp University Hospital, Antwerp, Belgium. [25]Translational Research in Inflammation and Immunology (TWI2N), Laboratory of Experimental Medicine and Paediatrics, Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium. [26]Institute of Medical Microbiology, Medical centre, Faculty of Medicine, University of Freiburg, Freiburg, Germany. [27]Department of Pathology, Aretaieion Hospital, Medical School, National & Kapodistrian University of Athens, Athens, Greece. [28]Newcastle NIHR Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Trust, Newcastle upon Tyne, UK. [29]These authors contributed equally: Seray Anak, Dania El Abyad. [30]These authors jointly supervised this work: Quentin M. Anstee,Schalk van der Merwe, Olivier Govaere. ✉e-mail: quentin.anstee@newcastle.ac.uk; schalk.vandermerwe@uzleuven.be; olivier.govaere@kuleuven.be

Methods

Patient cohorts and sample collection

This study included 289 liver tissues and 247 serum samples and data from 53,030 participants enrolled in the UKB. Biopsies for the snRNA-seq ($n = 18$) were collected from patients undergoing bariatric surgery or cholecystectomy at the University Hospital Leuven (UZ/KU Leuven, EC S63593) and the University Hospital Antwerp (EC UZA/UAntwerpen, B300201524515). MASH biopsies for the GeoMx spatial transcriptomics ($n = 8$) were collected at the specialized Freeman Hospital hepatology center (UREC, 18/NE/0182). Additional biopsies were obtained at the University Hospital Leuven (S67418) for immunohistochemistry/fluorescence ($n = 23$), CosMx spatial molecular imaging ($n = 10$), COMET hyperplex ($n = 6$) and MILAN proteome multiplexing ($n = 8$). Samples for the bulk transcriptomics (MASLD, $n = 206$; controls, $n = 10$) and serum proteomics ($n = 247$) have been drawn from the European MASLD Registry (NCT04442334)[46]. The registry includes samples from multiple centers, including Angers and Paris, Mainz, Turin, Linköping and Newcastle upon Tyne. All liver samples were scored by an expert liver pathologist according to the semiquantitative NASH-CRN Scoring System. Fibrosis stage ranged from F0 to F4 (cirrhosis) and the NAS was defined as the sum of the scores for steatosis, hepatocyte ballooning and lobular inflammation. At-risk MASH was defined as NAS ≥ 4, with at least 1 point deriving from each NAS component, as well as F ≥2 fibrosis. Patients with a BMI ≤ 25 kg m^{-2} with normal range liver function tests and no signs of steatosis after ultrasound examination were classified as lean controls. Patients with excessive alcohol consumption, autoimmune liver diseases, viral hepatitis and the use of steatogenic medications were excluded. Ethical approval was obtained from the relevant Ethical Committees at participating centers, and all patients provided written informed consent.

Liver droplet-based snRNA-seq

Fresh liver biopsies ($n = 18$) obtained through a 14G liver biopsy needle were immediately snap frozen and stored in liquid nitrogen. To obtain a single-nuclei suspension of the liver, samples were cut into small pieces in nuclease-free water with 96 mM NaCl, 10 mM Tris–HCl pH 7.5, 1 mM $CaCl_2$, 21 mM $MgCl_2$, 0.03% of Tween-20, 0.01% of BSA and 0.1 U μl^{-1} of Rnasin-Plus (Thermo Fisher Scientific). This was followed by a mechanical homogenization through a Dounce homogenizer and filtering through a 40-µm cell strainer (Falcon). The filter was washed using nuclease-free water with 96 mM NaCl, 10 mM Tris–HCl pH 7.5, 1 mM $CaCl_2$, 21 mM $MgCl_2$ and 0.1 U μl^{-1} of Rnasin-Plus (Thermo Fisher Scientific). This step was sequentially repeated on a 20-µm, 10-µm and 5-µm cell strainer (Falcon). A single-nuclei suspension was obtained through centrifugation (500*g* for 5 min at 4 °C) and resuspension in 300 µl Ca^{2+}-free and Mg^{2+}-free PBS (Gibco, 14190144) containing 1% BSA (Sigma, A9418).

snRNA-seq was performed using the 10x Genomics 3-prime (v3) dual index assay. A total of 20,000 nuclei, counted using a LUNA-FL dual fluorescence cell counter, were loaded per lane. Afterward, the nuclei were emulsified and amplified with 3′ adaptors and sample index attachment. Sequencing was performed using the Illumina NovaSeq 6000. The sequencing reads were aligned and annotated with the human reference genome (GRCh38/hg38) and demultiplexed using Cell Ranger (v3.0.2). Ambient RNA contamination was removed with default parameters of SoupX (v1.6.2)[47]. For analysis, the obtained gene-cell matrix was used in Seurat (v4.1.1)[48]. Genes expressed in less than three cells were excluded and filtered for cells expressing 700–9,000 genes and <30% mitochondrial genes of the total UMI counts. Doublets were removed using scDblFinder (v1.8.0)[49]. Samples were log-normalized with a scale factor of 10,000 and anchor integrated with 50 dimensions. The variation in UMI counts and mitochondrial gene content was regressed out.

Dimensionality reduction and clustering snRNA-seq

Unsupervised clustering was performed in the R package Seurat (v4.1.1). Clustering was performed using shared nearest neighbor with 50 principal components. Louvain clustering with a resolution of 1.5 and UMAP reduction with 50 dimensions was used for clustering. Clusters were combined and labeled based on previously published canonical markers[13,50]. Cells were subclustered for further analysis. They underwent new Louvain clustering (resolution = 1–2) and UMAP reduction with 15–50 dimensions. Cell subtypes were manually annotated based on the canonical markers identified through the FindAllMarkers function. To subcluster the *GPNMB*$^+$ myeloid cells, myeloid cells with a *GPNMB* UMI count of ≥2 were extracted, followed by Louvain clustering (resolution = 0.3) and UMAP reduction with five dimensions. Seurat together with ggplot2 (v3.3.5), pheatmap (v1.0.12) and ggpubr (v0.6.0) were used to generate heatmaps, UMAP visualization, violin plots, feature plots, bar plots and dot plots. Differential gene analysis was performed with the Wilcoxon rank-sum test with genes only present in at least 10–25% cells. To calculate the cluster composition ratio, the number of cells in each cluster from each patient was divided by the total number of cells per patient. EnhancedVolcano (v1.18.0) was used to visualize differential gene expression results from a logistic regression with UMI counts and mitochondrial gene content as latent variables. The LAM signature score was determined using the AddModuleScore function of Seurat, using a previously described transcriptional signature (*TREM2*, *LIPA*, *LPL*, *CTSB*, *FABP4*, *FABP5*, *LGALS1*, *LGALS3*, *CD9* and *CD36*)[9].

Differential abundance testing, pathway and cell–cell interaction analysis

The miloR package (v1.5.1)[51] was used for differential abundance testing. A *k*NN graph ($k = 30$, $d = 30$) was generated and cell neighborhoods were defined (prop = 0.2). The neighborhoods were quantified using countCells after calculating their distance with calcNhoodDistance ($d = 15–50$). TestNhood was used to evaluate the abundance across the conditions. Results were visualized using buildNhoodGraph and plotDAbeeswarm. Pathways were evaluated using the R package gene set variation analysis (v1.42.0) for single-sample gene set enrichment analysis (ssGSEA). Hallmark, KEGG, Reactome and Gene Ontology gene sets were exported from the MSigDB (v7.5.1) database through GSEABase (v1.56.0). A linear model was fitted with limma (v3.58.1) to identify significantly enriched gene sets across calculated gene set scores. Plots were generated with pheatmap using the normalized average gene set score per subset. To investigate cellular communication in the liver, the subsets of interest were extracted and analyzed using CellPhoneDB (v3.1.0), with genes expressed by 20% of cells, and the analysis was run for 1,000 iterations[52].

Myeloid cells fractioning

To obtain the fraction of myeloid cells positive for certain markers, the UMI counts were extracted from the snRNA-seq data and all possible combinations with a marker UMI count of ≥2 were counted and the proportion calculated per condition. This was also done for publicly available data from ref. 24 (GSE202379). The macrophages were extracted and the MASH with or without cirrhosis was combined before calculating the macrophage fraction for the markers of interest.

Spatial GeoMx WTA transcriptomics and CosMx molecular profiling

GeoMx Human Whole Transcriptome Atlas profiling (Bruker Spatial Biology) was performed on eight FFPE biopsy sections from patients with MASH fibrosis stage F3. Five-micrometer-thick sections were mounted on VWR Superfrost Plus Micro slides (48311-703) that were baked for 4 h at 37 °C and stored at 4 °C. Samples were processed 2 days later using the Leica Bond RXm system (Leica) for staining and probe incubation according to GeoMx guidelines. Regions of interest were selected based on (1) the presence of portal infiltration, (2) steatosis with lobular inflammation and/or lipogranuloma formation and (3) parenchyma without steatosis. Fluorescence CD68, CD45 and panCK

markers were used to segment the different macrophage, immune cell and epithelial cell populations from each region. A minimum of 200 nucleated cells was recommended per region of interest to process for segmentation. A total of 80 segments were processed and probes were quantified by RNA-seq (Illumina). We first calculated the limit of quantification (LOQ) for each segment to assess signal to noise, where LOQ is equal to the geometric mean multiplied by two standard deviations. We retained segments with at least 10% probes detected above LOQ, resulting in 77 segments for downstream analyses. We further filtered the dataset to 9,572 genes detected above LOQ in at least 10% segments. Data then underwent Q3 normalization. We performed pairwise differential expression comparing tissue regions in the macrophage (CD68 enriched), immune cell (CD45 enriched) and epithelial cell (panCK enriched) populations. We fit linear mixed-effect models with the GeoMxTools Bioconductor R package (v3.0.1). Tissue region was set as the fixed effect and patient ID as a random intercept term to account for nonindependence of multiple segments per tissue. Genes were considered differentially expressed with absolute $\log_2$ fold change >0.5 and false discovery rate <0.05. After differential expression, we performed ssGSEA with Gene Set Variation Analysis (v1.44.0) to assess pathway activity in each segment. ssGSEA scores were calculated for each segment using the KEGG Brite database. ssGSEA scores were then used as input for differential pathway enrichment using the linear mixed-effect modeling approach described above[53].

CosMx spatial molecular imaging (Bruker Spatial Biology) was performed on eight biopsy samples and two explant liver samples from patients with end-stage MASH using the CosMx Human Universal Cell Characterization RNA Panel (1000-plex) supplemented with 14 custom genes of interest. The following cell markers were used for morphology visualization: B2M/CD298, panCK, CD45 and CD68, combined with DAPI. Five-micrometer-thick FFPE sections were mounted on VWR Superfrost Plus Micro slides (48311-703) that were baked for 4 h at 37 °C and stored at 4 °C. Samples were processed 7 days later using the Leica Bond RXm system (Leica) for staining and probe incubation according to CosMx guidelines. RNA target readout was performed on the CosMx SMI instrument as described previously[25]. Cloud-based AtoMx Spatial Informatics Platform (Bruker Spatial Biology) was used to export data and images for Seurat (v5.1.0)[54]. The data were merged, SCT normalized, followed by principal component analysis and UMAP calculation using 30 dimensions with Seurat default parameters for CosMx data. Louvain clustering with a resolution of 0.3 was used and clusters were combined and labeled based on canonical markers. Seurat was used for visualization with default functions, and areas of interest were captured with the Crop function. To obtain the myeloid cells positive for *GPNMB*, *HLA-DRA* and/or *LPL*, the probe counts were extracted from the data and all possible combinations with a marker probe count of >0 were counted and labeled accordingly.

Spatial profiling using the COMET system and HORIZON software

Sequential multiplex immunofluorescence staining was carried out using the COMET system (Lunaphore Technologies). Markers of interest were bound with their respective primary antibody (Supplementary Table 12) or RNA probe and visualized in a cyclic manner using secondary Alexa Fluor Plus antibodies (Thermo Fisher Scientific). Simultaneous spatial detection of *IL32* RNA probe (ACD-Bio, Bio-Techne, 541437-T3) and protein was performed on the same tissue section using RNAscope HiPlex Pro (ACDBio, Bio-Techne) and sequential immunofluorescence (seqIF Lunaphore Technologies). High-resolution imaging acquisition was followed by data processing of raw images in the HORIZON software suite (Lunaphore Technologies). The analysis workflow included automated image registration and stitching, background subtraction and normalization of fluorescence intensity. Cell annotation and quantification in regions of interest were performed on the raw images with QuPath (v0.5.1). Regions of interest were manually selected in the portal tract and parenchyma. Cell detection was performed based on the DAPI channel. For each marker of interest, classifiers were developed to identify positive staining using either manual thresholding or machine learning approaches. CD163 alone or CD14/CD68/CD163 classifiers were used to define the myeloid cell population. Individual classifiers were then combined to categorize cells based on the expression of multiple markers, enabling comprehensive phenotypic profiling and fraction analysis. For additional visualization, automated image registration, background subtraction and normalization of fluorescence intensity through the HORIZON software suite (Lunaphore Technologies; v2.3.0.0) were used.

Spatial profiling through MILAN

Immunofluorescence multiplexing staining was performed using a panel of primary antibodies as listed in Supplementary Table 12, on eight FFPE MASLD biopsy sections following the previously published protocol[27]. Immunofluorescence images were obtained using the AxioScan Z1 slide scanner (Zeiss). Stainings and the following image acquisition were done simultaneously. Serial cycles were registered using imreg (https://github.com/cgohlke/imreg). Tissue autofluorescence was subtracted using a baseline and cell segmentation was applied to the DAPI channel using StarDist (https://arxiv.org/abs/1806.03535). For every cell, topological features (x and y coordinates), morphological features (nuclear size) and molecular features (mean fluorescence intensity) were extracted. Mean fluorescence intensities were normalized with z scores for each marker and sample. Different cell populations were identified using manual annotation. Macrophages/monocytes were identified by CD68, CD14 or CD163 expression ($z > 1$); hepatocytes by ARG1 ($z > 1$); cholangiocytes by panCK ($z > 1$); and endothelial cells by CD31 ($z > 1$). The macrophage/monocyte population was further evaluated using UMAP dimensionality reduction. In this population, different subcategories were identified using manual gating—GPNMB$^+$ ($z > 1$), lysozyme$^+$ ($z > 1$), HLA-DR$^+$ ($z > 1$), SPP1 ($z > 1$) and CD163$^+$ ($z > 1$). Cell distances were calculated as the Euclidean distance between GPNMB$^+$ macrophages and the closest cholangiocyte. Continuous distances were binned in chunks of 10 µm and the fraction of cells was estimated in each chunk.

Immunohistochemistry, immunofluorescence and functional models

Detailed methodologies for immunohistochemistry, immunofluorescence and CRISPR–Cas9 knockout in THP-1 cells with lipid loading, phagocytosis and conditioning, as well as for human ex vivo PCLS preparation, are provided in the Supplementary Note.

SomaScan serum proteomics

The SomaScan Platform (SomaLogic) was used to analyze 247 human serum samples (20 µl each). These samples were sourced from the European MASLD Registry, encompassing multiple centers—Angers and Paris, Mainz, Turin, Linköping and Newcastle Upon Tyne. All patients included had biopsy-confirmed MASLD, with liver biopsies centrally evaluated by an expert pathologist using the semiquantitative NASH-CRN Scoring System. Details regarding patient recruitment and inclusion criteria have been previously described[46]. For proteomic profiling, serum samples were diluted 1 of 20 and analyzed using the SomaScan Assay (v4.0) (SomaLogic), which uses SOMAmer (Slow Off-rate Modified Aptamer) reagents to enable highly multiplexed, quantitative protein detection. In brief, each diluted sample was incubated with a mixture of SOMAmer reagents, thereby forming specific SOMAmer–protein complexes. These complexes were captured on streptavidin-coated beads, and unbound proteins were removed by washing. A photocleavable linker was used to release the SOMAmer–protein complexes into solution and a two-step immobilization process was used to reduce nonspecific binding. SOMAmer reagents were then

eluted and hybridized to a custom DNA microarray for quantification. The relative abundance of each protein was determined by measuring the fluorescence intensity of each SOMAmer, reported as relative fluorescent units, which directly correlates with protein concentration in the original sample. All steps were performed by SomaLogic according to the manufacturer's protocols and quality control procedures. Reads were $\log_{10}$ transformed and correlated with histopathological parameters. Patients were stratified based on the histological presence of MASH, advanced fibrosis stage 3 or 4, high disease activity or 'at-risk' MASH. High disease activity was defined as NAS ≥4; this is the sum of the histological scores for steatosis, hepatocyte ballooning and lobular inflammation. Patients were categorized as 'at-risk' MASH when the biopsy was scored as MASH combined with a fibrosis stage of at least 2 and a NAS of ≥4. All biopsies were centrally scored by an expert pathologist.

Bulk transcriptomics

Bulk RNA-seq data from 206 MASLD patients and 10 controls (GSE135251) were processed as previously reported[22]. Quality was assessed using FastQC (v0.11.5) and MultiQC (v1.2dev), and aligned to the GRCh38 genome. Gene-level count tables were generated using HTSeq; counts were normalized using the trimmed mean of *M*-values method and transformed with limma's voom. Batch and sex effects were removed using limma's removeBatchEffects function. Heatmaps were generated using the R package pheatmap. Binary logistic regression analysis was performed in SPSS (IBM) using a backward stepwise likelihood ratio model by adding clinical and mRNA gene expression variables.

Deconvolution

SpatialDecon (v1.10.0) was used to deconvolve the GeoMx dataset. Here the snRNA-seq data of myeloid cells or the whole dataset were used as a reference dataset to estimate the cell abundance, using default parameters. To estimate the macrophage abundance, we used CIBERSORTx to deconvolute the bulk RNA-seq data. In short, a signature gene profile from the snRNA-seq data of our myeloid cells was generated using default parameters, filtering for nonhematopoietic cells and genes expressed in at least 25% of the population. The liver RNA-seq data were then deconvoluted using the obtained signature gene profile after batch correction across 100 permutations.

PheWAS

PheWAS was conducted using data from the UKB, a large-scale biobank containing genetic and health information from approximately 500,000 UK participants. Our study included a subset of 53,030 participants for whom OLINK proteomics data were available. The UKB has approval from the North West Multi-centre Research Ethics Committee, all participants provided informed consent and these data were analyzed using UKB access 71300. The proteomic data were obtained using the OLINK Proteomics platform, which quantifies the expression levels of targeted proteins. This platform uses a proximity extension assay technology, allowing for the high-throughput analysis of multiple proteins simultaneously. The Olink Explore 3072 platform features 2,941 immunoassays targeting 2,925 proteins, including GPNMB. Data are reported in Normalized Protein eXpression units, measured on a $\log_2$ scale. The UKB plasma samples were processed at Olink's Uppsala facilities, using three NovaSeq 6000 sequencing systems. The coding for clinical diagnoses in our dataset followed the World Health Organization's International Classification of Diseases, Tenth Revision (ICD-10) coding systems. Phenotypes were used from hospital inpatient records until January 2023, including a wide range of health outcomes. We adjusted our analyses for age, sex and BMI at baseline. Age was treated as a continuous variable, while sex and BMI were categorized according to standard classifications. PheWAS analyses were performed using R (v4.0.2). We converted the ICD-10 codes to 1253 associated Phecodes using PheWAS package. With the PheWAS R package, a series of case–control tests were performed—(1) the case group was generated by including patients with the tested Phecode; and (2) participants were assigned to the control group based on the absence of the tested Phecode.

Statistics and reproducibility

One-way analysis of variance with Tukey post hoc test, Kruskal–Wallis test with Bonferroni correction and Mann–Whitney *U* test were performed in IBM SPSS (v29) or GraphPad Prism (v9). Holm–Bonferroni post hoc test was done in R. Binary logistic regression analysis was performed in IBM SPSS (v.29) using a backward stepwise likelihood ratio model. A P value <0.05 was considered significant. No statistical method was used to predetermine sample size. No statistical method was used to predetermine sample size. No data were excluded from the analyses, the experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. Graphs were produced using GraphPad Prism (v9) or the respective R packages.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The snRNA-seq dataset generated in this study is available from the European Genome-Phenome Archive under accession EGAS50000000768. Burker Spatial CosMx and WTA GeoMx data are available from GEO under accession GSE312698 and via Zenodo at https://doi.org/10.5281/zenodo.18153344 (ref. 55). Bulk RNA-seq data from human MASLD samples are available from GEO under accession GSE135251, as previously published[22]. Data used for Extended Data Fig. 4g were derived from GEO accession GSE202379 (ref. 24). Lunaphore COMET data are available from the KU Leuven Research Data Repository at https://doi.org/10.48804/ERGUWX (ref. 56). MILAN multiplex digital reconstructions and overlay images are accessible at figshare (https://doi.org/10.6084/m9.figshare.31986669.v1 (ref. 57)), and original scans are available from the corresponding author upon reasonable request. SomaLogic plasma proteomics data analyzed here were originally generated as previously published[6]. OLINK proteomics data are available to approved researchers from UKB (www.ukbiobank.ac.uk). Immunohistological figures are available at figshare https://doi.org/10.6084/m9.figshare.28881296.v1 (ref. 58) and https://doi.org/10.6084/m9.figshare.28881638.v2 (ref. 59). Source data are provided with this paper.

Code availability

All data analysis scripts use existing methods and are provided in a GitHub repository provided with this paper (https://github.com/govaerelab/OG_MB-MASLD-LIVER-2025 and https://github.com/MarkusBoesch93/OG_MB-MASLD-LIVER-2025) and via Zenodo at https://doi.org/10.5281/zenodo.17980902 (ref. 60).

References

46. Hardy, T. et al. The European NAFLD registry: a real-world longitudinal cohort study of nonalcoholic fatty liver disease. *Contemp. Clin. Trials* **98**, 106175 (2020).
47. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, giaa151 (2020).
48. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
49. Germain, P. L., Lun, A., Garcia Meixide, C., Macnair, W. & Robinson, M. D. Doublet identification in single-cell sequencing data using scDblFinder. *F1000Res* **10**, 979 (2021).

50. MacParland, S. A. et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat. Commun.* **9**, 4383 (2018).
51. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022).
52. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
53. Danaher, P. et al. Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. *Nat. Commun.* **13**, 385 (2022).
54. Hao, Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
55. Govaere, O. GeoMx Human Whole Transcriptome Atlas profiling on at-risk MASH F3 patients [data set]. *Zenodo* https://doi.org/10.5281/zenodo.18153344 (2026).
56. Bösch, M., Anak, S., Ostyn, T., El Abyad, D. & Govaere, O. Metabolic dysfunction-associated steatohepatitis (MASH) GeoMx and COMET dataset (version V1) [dataset]. *KU Leuven RDR* https://doi.org/10.48804/ERGUWX (2025).
57. Govaere, O. MILAN multiplex proteomics stainings of human MASLD biopsies: digital reconstructions and overlay images. figshare. *figure* https://doi.org/10.6084/m9.figshare.31986669.v1 (2026).
58. Govaere, O. DAB images of liver biopsies with GPNMB staining. *figshare* https://doi.org/10.6084/m9.figshare.28881296.v1 (2025).
59. Govaere, O. DAB stainings of human liver samples with the markers used in immunofluorescence multiplexing. *figshare* https://doi.org/10.6084/m9.figshare.28881638.v2 (2025).
60. Boesch, M. Integrated multi-omics of metabolic dysfunction-associated steatohepatitis. *Zenodo* https://doi.org/10.5281/zenodo.17980902 (2025).

Acknowledgements

This study was supported by the Research Foundation—Flanders (FWO; project G032324N to O.G.) and the NanoString/Bruker WTA grant UK/Ireland (to O.G.), and in part by the LITMUS project, which received funding from the Innovative Medicines Initiative 2 Joint Undertaking (grant agreement 777377 to Q.M.A.). This joint undertaking receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA). This communication reflects the view of the authors, and neither IMI nor the European Union and EFPIA is liable for any use that may be made of the information contained herein. Q.M.A. is an NIHR Senior Investigator and is supported by the Newcastle NIHR Biomedical Research Centre. A.A. received funding from the FWO (senior postdoctoral grant 12ATN24N). The NanoString/Bruker WTA grant (UK/Ireland) was used to generate the GeoMx data at the Amsterdam CX Lab (NanoString/Bruker). The other funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The authors thank L.B. Lisboa, M. Leon, R. van Eijsden, A. Mason, M. Wallays, R. Schepers, T. van Brussel and G. Philips for the excellent technical assistance and expertise. The authors are indebted to patients and their families for participating in the research. This research has been conducted using the UKB Resource (application 71300). UKB data were accessed by C.V.S. and K.M.S. Copyright 2024, NHS England. The data were reused with the permission of NHS England and/or UKB, and all rights are reserved. This study used data provided by patients and collected by the NHS as part of their care and support.

Author contributions

O.G., Q.M.A. and S.v.d.M. conceptualized the study, and directed and supervised the research. M.B. and O.G. have designed the study and drafted the paper. D.T., T.R. and O.G. did the histopathological scoring. S.A., J.P., C.M., N.S., A.H. and O.G. performed the in vitro experiments. S.A., D.E.A., T.O., J.S.-H.C. and N.D.L. did the sample preparation and acquisition for the various spatial proteome multiplexing. S.A., D.E.A., T.O. and O.G. performed the immunohistochemistry and proteome multiplex quantifications. D.L. and B.B. supervised the snRNA-seq sample preparation and snRNA-seq data preprocessing. M.B., D.E.A., A.A., G.S., T.M.V., D.N. and S.C. performed the bioinformatics analysis. K.M.S. and C.V.S. performed the UKB analysis. L.V.M., M.L., E.D., J.J., H.T., B.T., M.V., M.A., M.E., J.B., E.B., V.R., A.K.D., J.M.S., J.V., S.F., Q.M.A. and S.v.d.M. recruited patients and collected the data and specimens. R.F.-A., L.S., P.L.C., T.V., J.C., F.D.S., F.B. and H.K. helped with data acquisition and analysis.

Competing interests

O.G. received project funding from Bruker Spatial Biology to run the GeoMx experiment. D.M. and T.M.V. are employees of Bruker Spatial Biology. M.A. has research collaborations with GSK and AstraZeneca. J.M.S. declares consultant honorary from Akero, Alentis, Alexion, Altimmune, AstraZeneca, 89Bio, Bionorica, Boehringer Ingelheim, Boston Pharmaceuticals, Gilead Sciences, GSK, HistoIndex, Ipsen, Inventiva Pharma, Madrigal Pharmaceuticals, PRO.MED.CS Praha a.s., Kríya Therapeutics, Eli Lilly, MSD Sharp & Dohme GmbH, Novartis, Novo Nordisk, Pfizer, Roche, Sanofi and Siemens Healthineers; and speaker honorarium from AbbVie, Boehringer Ingelheim, Gilead Sciences, Ipsen, Eli Lilly, Novo Nordisk, Madrigal Pharmaceuticals, Stockholder options—Hepta Bio. D.T. is a consultant for Inventiva SA and ICON Plc. Q.M.A. received project funding from AstraZeneca, Boehringer Ingelheim and Interecept; is a consultant for Abbvie, Alimentiv, Akero, AstraZeneca, Axcella, 89Bio, Boehringer Ingelheim, Boston Pharaceticals, Corcept Therapeutics, Echosens, Eli Lilly, Gilead, GlaxoSmithKline, HistoIndex, Intercept, Inventiva, Madrigal, Medpace, Merck, Metadeq, Novo Nordisk, PathAI, Pfizer, Pharmanest, Prosciento, Regeneron, Resolution Therapeutics, Roche and Terns; obtained extra funding from Avalere, Catalyst Medical Education, Integritas Communications, Medscape, Novo Nordisk, Springer Healthcare and Elsevier; and is on the advisory board from Medpace and Gilead. The other authors declare no competing interests.

Additional information

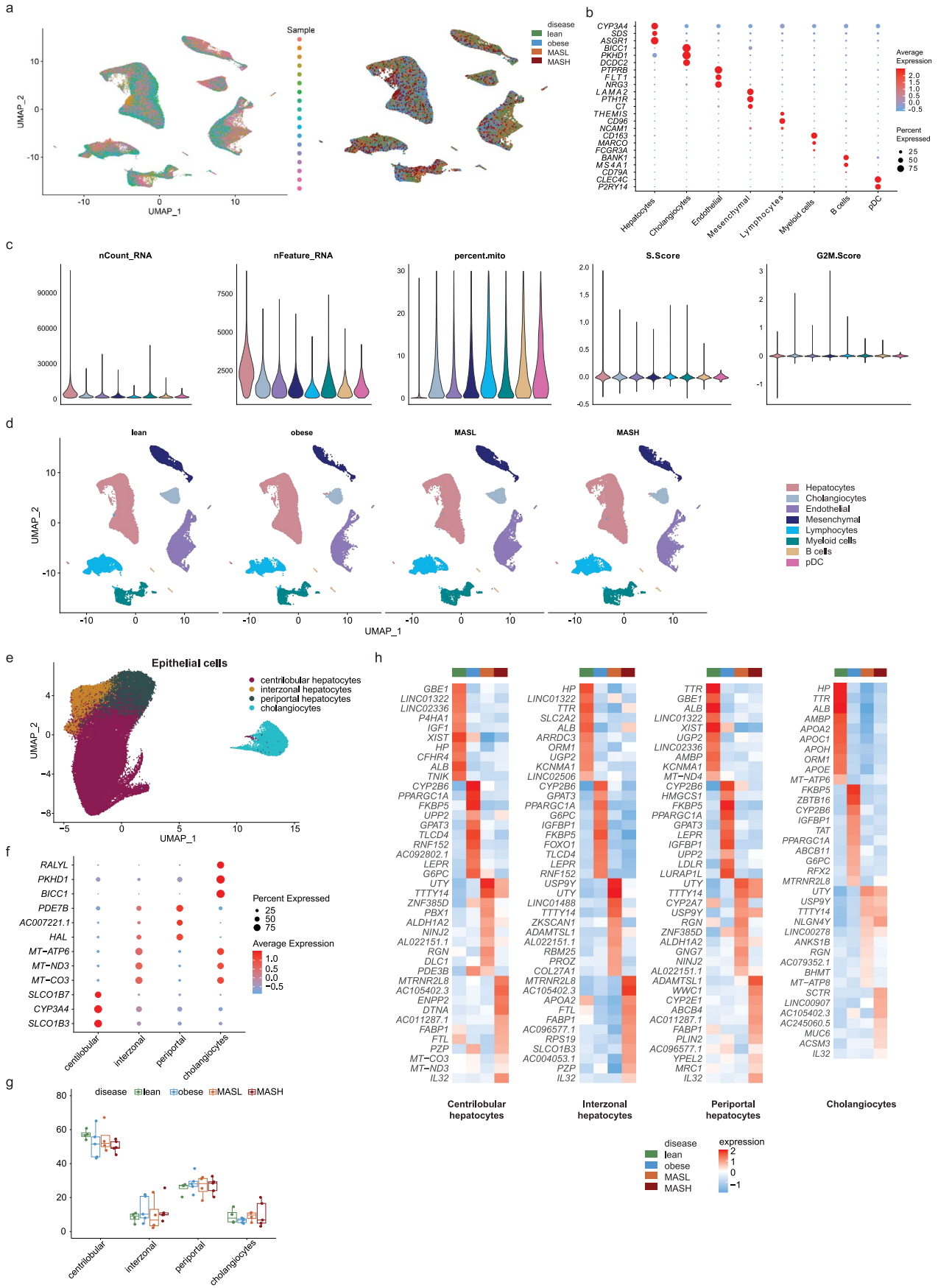
Extended data is available for this paper at https://doi.org/10.1038/s41588-026-02600-3.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-026-02600-3.

Correspondence and requests for materials should be addressed to Quentin M. Anstee, Schalk van der Merwe or Olivier Govaere.

Peer review information *Nature Genetics* thanks Ping Chen, Prakash Ramachandran, and the other anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

https://doi.org/10.1038/s41588-026-02600-3

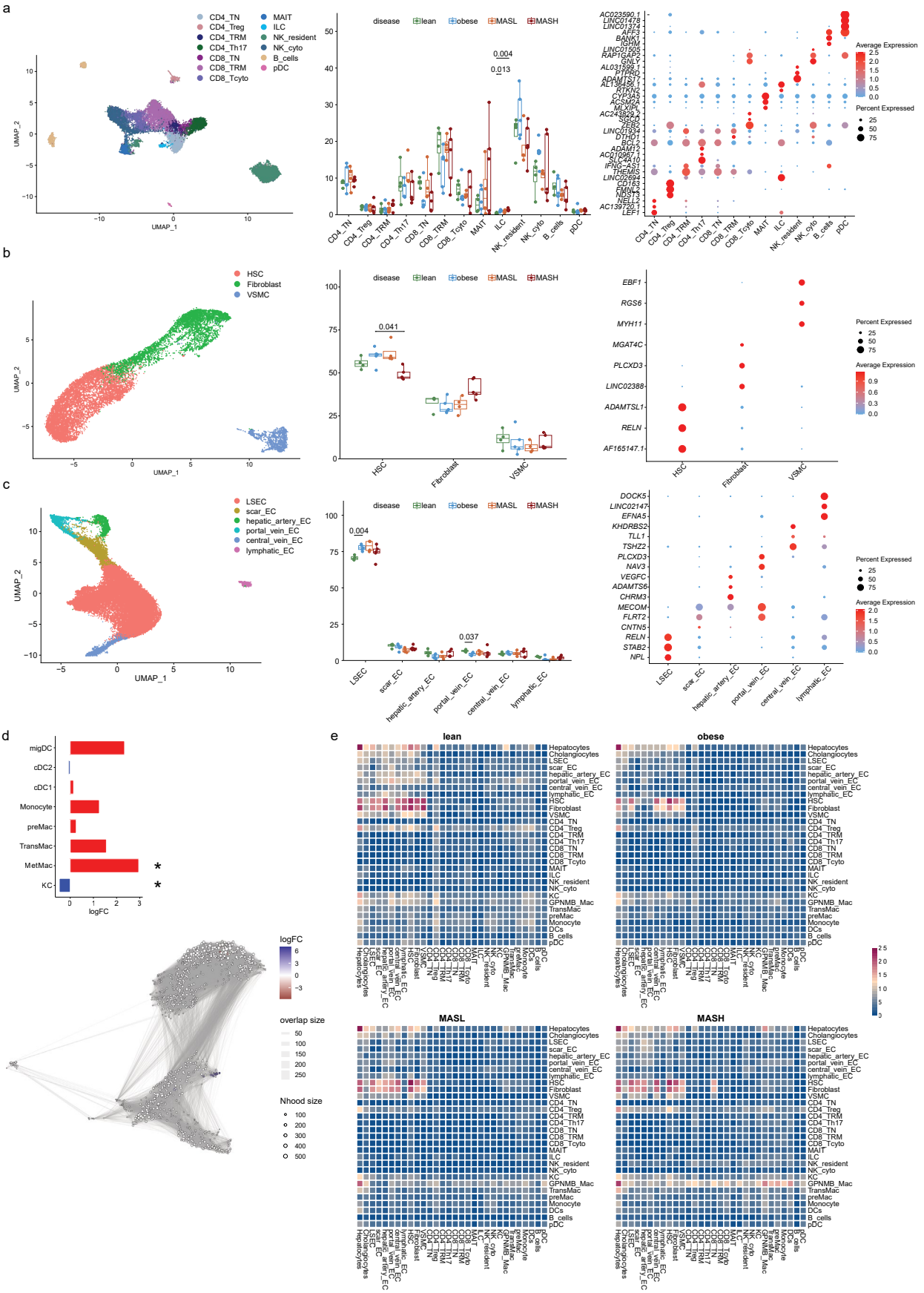


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41588-026-02600-3

Extended Data Fig. 1 | Integration of hepatic single-nucleus RNA sequencing. **a**, UMAP plot of hepatic cells of individual patients (left) and UMAP plot of hepatic cells indicating the patient sample classification as lean (green), obese (blue), MASL (orange) and MASH (red; right). **b**, Dot plot of selected differentially expressed genes. The expression is indicated by color saturation and the size indicates the percentage of expressing cells in the respective cluster. **c**, Violin plots of various cell quality metrics, the cell type is marked by individual coloring. **d**, Individual UMAP plots of hepatic cells per disease. **e**, UMAP plot of 109,659 epithelial cells. **f**, Dot plot of the top 3 differentially expressed genes. The expression is indicated by color saturation and the size indicates the percentage of expressing cells in the respective cluster. **g**, Bar plot showing the ratio (%) of the epithelial subtypes derived from each group. Dots indicate individual patients as biological replicate, Holm–Bonferroni adjusted *p* values of two-sided t-test and median line indicated, the boxes extend from the 25th to 75th percentiles, and the whiskers extend to the minimum and maximum values. **h**, Heatmap illustrating the expression of the top 10 and selective marker genes identified in hepatocyte and cholangiocyte subtypes. The expression is indicated by color saturation.

https://doi.org/10.1038/s41588-026-02600-3

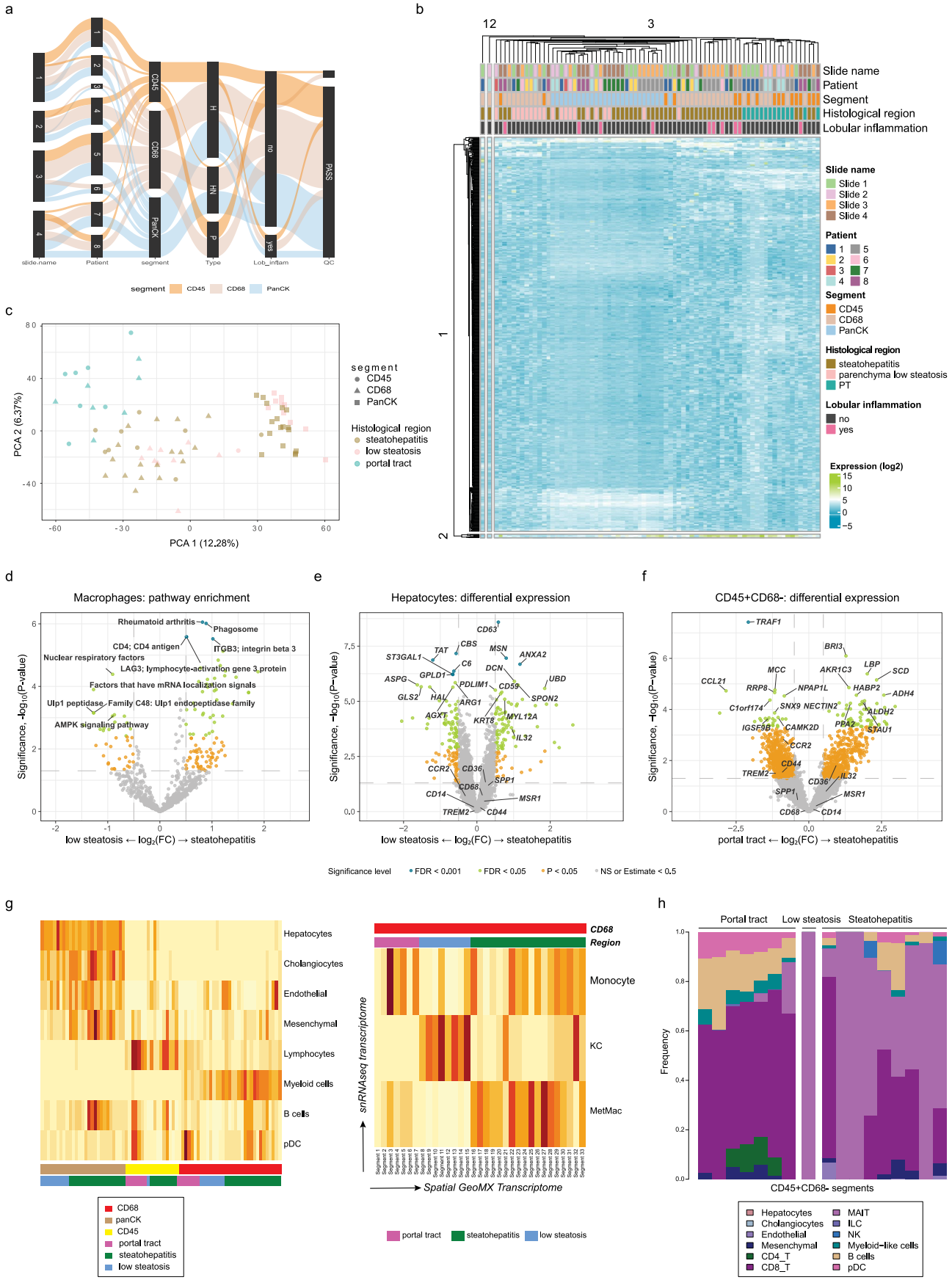


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41588-026-02600-3

Extended Data Fig. 2 | Hepatic subclustering. a, UMAP plot of 16,863 lymphocytes, B cells and pDCs (left). Bar plot showing the ratio (%) of the lymphocyte, B cell and pDC subtypes derived from each group. Dots indicate individual patients as biological replicate, Holm–Bonferroni adjusted *p* values of two-sided t-test and median line indicated, the boxes extend from the 25th to 75th percentiles, and the whiskers extend to the minimum and maximum values (middle). Dot plot of top 3 differentially expressed genes. The expression is indicated by color saturation and the size indicates the percentage of expressing cells in the respective cluster (right). **b**, UMAP plot of 10,681 mesenchymal cells (left). Bar plot showing the ratio (%) of the mesenchymal subtypes derived from each group. Dots indicate individual patients as biological replicate, Holm–Bonferroni adjusted *p* values of two-sided t-test and median line indicated, the boxes extend from the 25th to 75th percentiles, and the whiskers extend to the minimum and maximum values (middle). Dot plot of top 3 differentially expressed genes. The expression is indicated by color saturation and the size indicates the percentage of expressing cells in the respective cluster (right). **c**, UMAP plot of 26,113 endothelial cells (left). Bar plot showing the ratio (%) of the endothelial subtypes derived from each group. Dots indicate individual patients as biological replicate, Holm–Bonferroni adjusted *p* values of two-sided t-test and median line indicated, the boxes extend from the 25th to 75th percentiles, and the whiskers extend to the minimum and maximum values (middle). Dot plot of top 3 differentially expressed genes. The expression is indicated by color saturation and the size indicates the percentage of expressing cells in the respective cluster (right). **d**, MiloR calculated differential neighborhoods abundance log fold change between the lean, obese, MASL and MASH conditions of myeloid populations. Differential abundant neighborhoods after multiple hypothesis testing at false discovery rate (FDR) 5% are colored. Blue coloration indicates significantly increased and red decreased abundance in the MASLD condition, based on the FDR value. Results embedded onto the UMAP (bottom) or the average log fold change of all neighborhoods depicted as bar plot (top). **e**, Heatmap showing the number of potential ligand-receptor pairs between cell groups per disease predicted by CellPhoneDB. The number of unique interactions is indicated by color saturation, from high (red) to low (blue). $*p < 0.05$.

https://doi.org/10.1038/s41588-026-02600-3

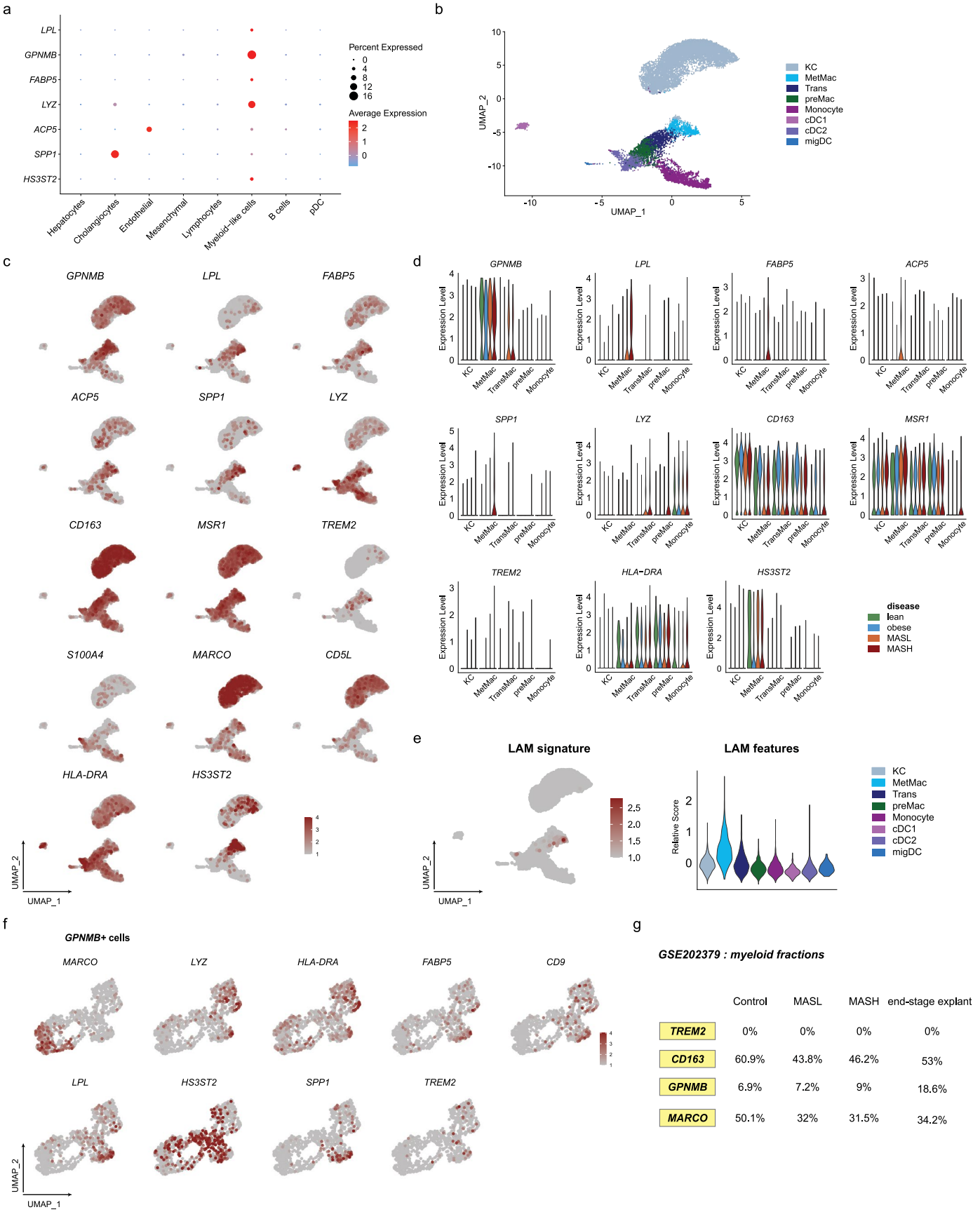


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41588-026-02600-3

Extended Data Fig. 3 | GeoMx spatial profiling. a, Riverplot showing the distribution of patient, segment, type, lobular inflammation (Lob_inflam) and quality control (QC) per slide. **b**, Hierarchical clustering heatmap showing the top variable genes to visualize unsupervised gene expression patterns in each region of interest. The expression is indicated by color saturation—the slide name, patient, segment, type and lobular inflammation (Lob_inflam) are marked by individual coloring. **c**, PCA plot of the various regions of interest and cell segments of all biopsies. **d**, Volcano plot showing the differential pathways in $CD68^+$ macrophage segments between the regions of low steatosis and steatohepatitis (right). **e**, Volcano plot showing the differential expressed genes in hepatocyte segments (panCK) between the regions of low steatosis and steatohepatitis (left). **f**, Volcano plot showing the differential expressed genes in $CD45^+$ $CD68^-$ segments between the regions of the portal tract and steatohepatitis (right). **g**, Heatmap of the estimated cell abundancies across the various segments estimated with SpatialDecon based on the snRNA-seq hepatic cell data (left) or across the CD68 segments based on the snRNA-seq myeloid cell data (right). Red indicates a high and white a low estimated cell abundance. **h**, The mean cell frequencies in the $CD45^+$ $CD68^-$ segments across the various regions of interest estimated with SpatialDecon based on the snRNA-seq hepatic cell data, estimating primarily lymphocytes to be present in those segments.

https://doi.org/10.1038/s41588-026-02600-3



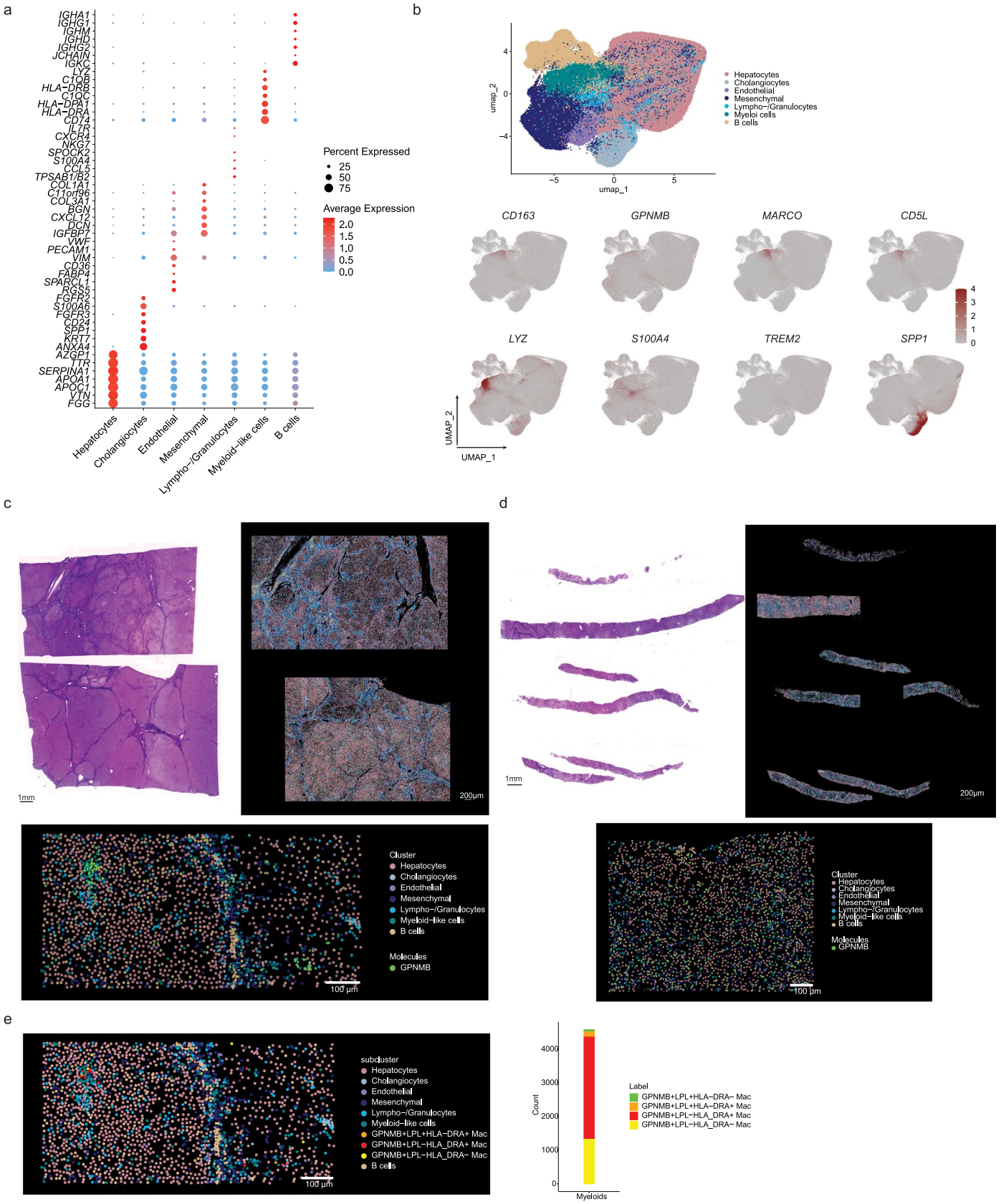
	Control	MASL	MASH	end-stage explant
TREM2	0%	0%	0%	0%
CD163	60.9%	43.8%	46.2%	53%
GPNMB	6.9%	7.2%	9%	18.6%
MARCO	50.1%	32%	31.5%	34.2%

Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41588-026-02600-3

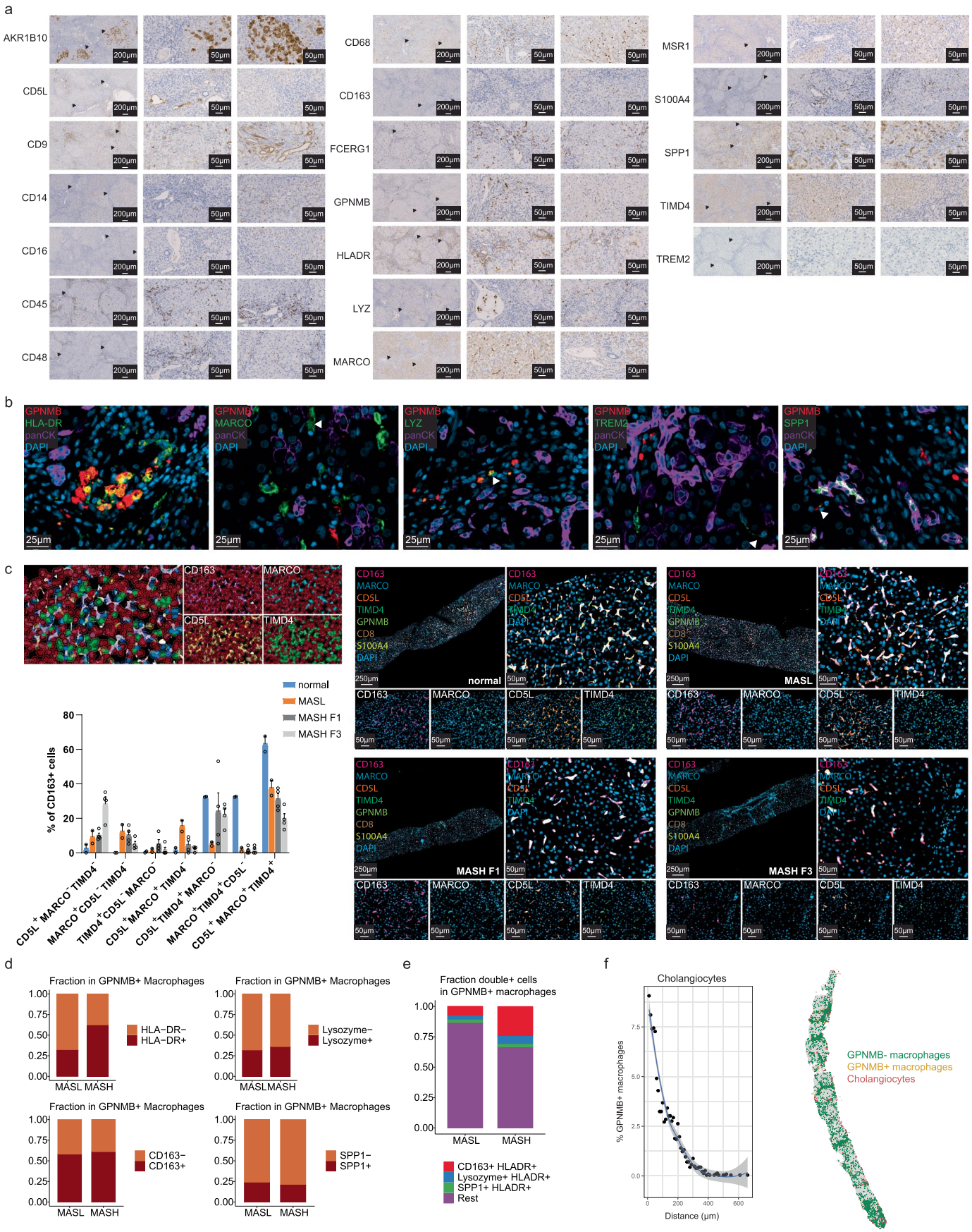
Extended Data Fig. 4 | Hepatic gene expression on single-nucleus level. **a**, Dot plot of selected differentially expressed genes in the snRNA-seq data. Expression is indicated by color saturation and size indicates the percentage of expressing cells in the respective cluster. **b**, UMAP plot of 12,834 myeloid cells from the snRNA-seq data. **c**, UMAP showing the expression of selected genes in myeloid cells of the snRNA-seq data. Expression is indicated by color saturation (low (gray) to red (high)). **d**, Violin plots of selective marker genes in myeloid cells, the disease is marked by individual coloring. **e**, UMAP showing the LAM signature score (*TREM2*, *LIPA*, *LPL*, *CTSB*, *FABP4*, *FABP5*, *LGALS1*, *LGALS3*, *CD9*, *CD36*) in myeloid cells of the snRNA-seq data. Expression is indicated by color saturation (low (gray) to red (high)) (left). And violin plots of the LAM signature score in myeloid cells of the snRNA-seq data, the myeloid subcluster is marked by individual coloring (right). **f**, UMAP showing the expression of genes in 944 *GPNMB*$^{+}$ myeloid cells (*GPNMB* UMI count ≥ 2) of the snRNA-seq data. Expression is indicated by color saturation (low (gray) to red (high)). **g**, Fractions of 3,014 macrophages from publicly available hepatic snRNA-seq data mentioned in ref. 24 (GSE202379) positive for *TREM2*, *CD163*, *GPNMB* or *MARCO* based on their gene expression (UMI count ≥ 2) across healthy controls (n = 4), MASL (n = 7), MASH (n = 31) and end-stage explants (n = 5) in percentages.

https://doi.org/10.1038/s41588-026-02600-3



Extended Data Fig. 5 | CosMx spatial overview. a, Dot plot of top 7 differentially expressed genes in the CosMx data. The expression is indicated by color saturation and the size indicates the percentage of expressing cells in the respective cluster. **b**, UMAP plot of 552,676 hepatic cells (top). UMAP showing the expression of certain genes in the hepatic cells. Expression is indicated by color saturation (low (gray) to red (high)). **c**,**d**, H&E staining and CosMx molecular profiling highlighting different hepatic cell populations (top) and zoom-in with GPNMB probes (bottom) for end-stage liver disease samples (n = 2) and representative MASL and MASH biopsies. **e**, Representative tissue section of CosMx molecular profiling in end-stage liver disease, highlighting different hepatic cell populations and different *GPNMB*, *LPL* and/or *HLA-DRA* positive/negative macrophages. Bar plot showing the number of different *GPNMB*, *LPL* and/or *HLA-DRA* positive/negative macrophages.

https://doi.org/10.1038/s41588-026-02600-3

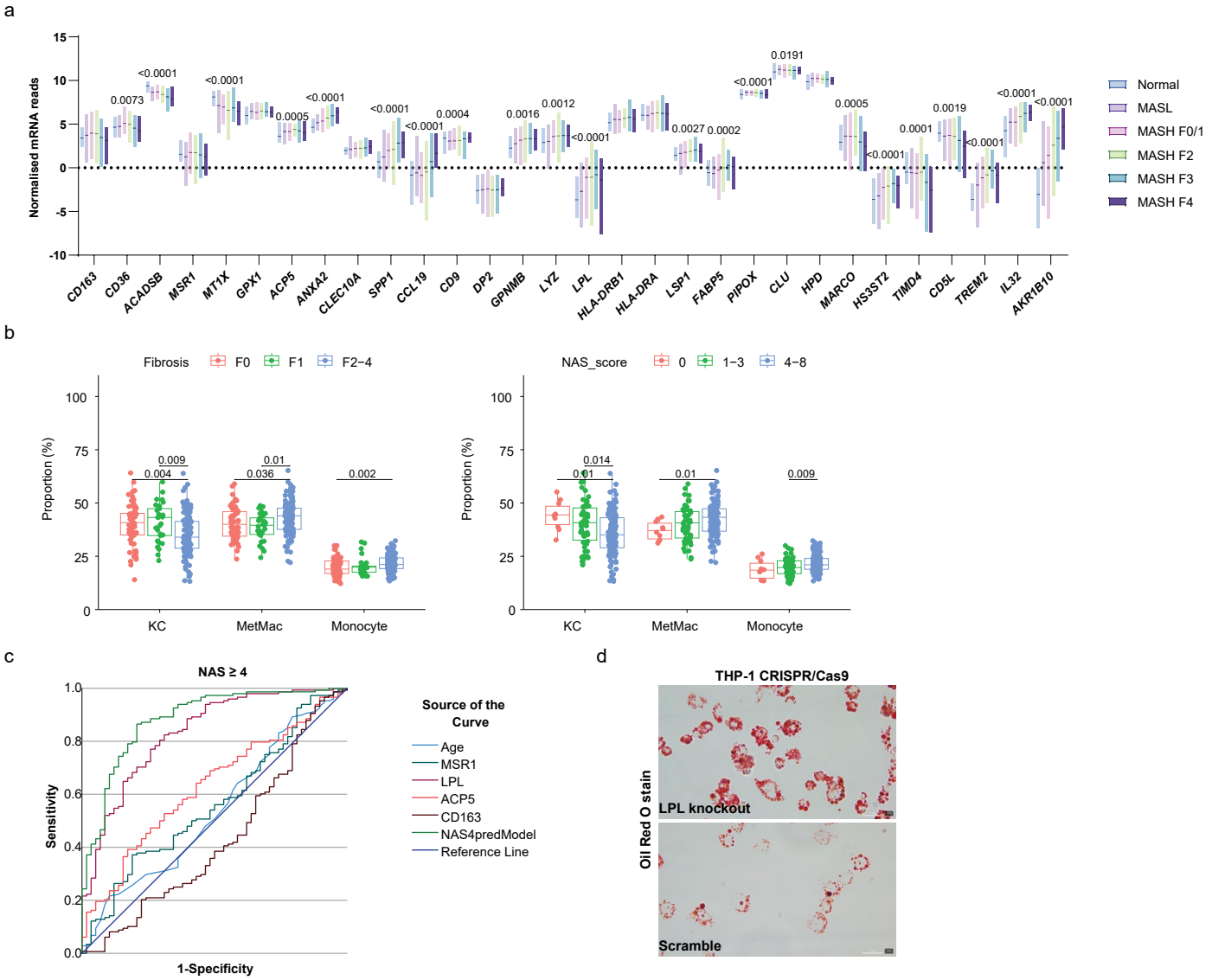


Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s41588-026-02600-3

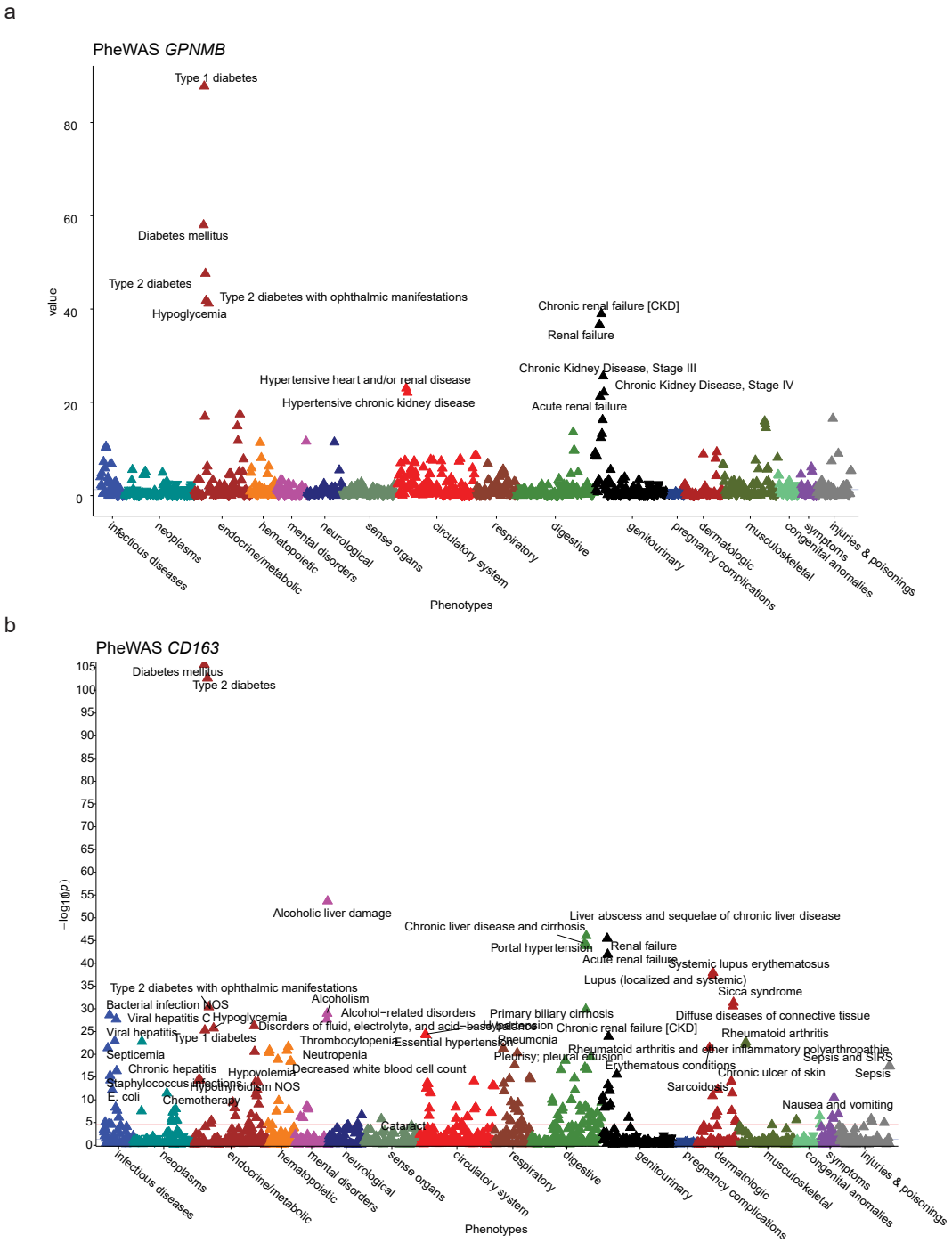
Extended Data Fig. 6 | Validation of MILAN and COMET Hyperplex. **a**, Representative immunohistochemistry images of various markers validated for MILAN and COMET Hyperplex in human liver biopsies. Regions of interest are indicated via arrows and zoomed in. Scale bars = 200 μM and 50 μM. **b**, COMET Hyperplex images illustrating various myeloid markers in end-stage MASLD. **c**, Digital reconstruction of a COMET Hyperplex region based on the trained classifier, showing alone and as composite cells positive for CD163 (pink), MARCO (blue), CD5L (yellow), TIMD4 (green) or none (red) (top left). Bar plot showing the abundance of $CD163^+$ cells positive for one or multiple markers across various disease stages (2–4 regions quantified per biopsy/disease stage, presented as mean ± s.e.m.). Representative spatial COMET Hyperplex overlayed multicolor fluorescent images of immune cells markers with zoom-in of KC markers (CD163, MARCO, CD5L, TIMD4) as composite or alone for normal (top middle), MASL (top right), MASH F1 (bottom middle) and MASH F3 (bottom right). **d**, Bar plots of the fraction of HLA-DR-, lysozyme-, CD163- and SPP1-positive and -negative cells in the $GPNMB^+$ macrophages in MASL and MASH identified via MILAN. **e**, Bar plot indicating the fraction of double-positive cells in $GPNMB^+$ macrophages between MASL and MASH identified via MILAN. **f**, Line plot indicating the abundance of $GPNMB^+$ macrophages at a distance (μm) to cholangiocytes identified via MILAN. The line indicates the spline curve fit to the data points. Digital reconstruction of a representative human liver sample visualizes colocalization of cholangiocytes with GPNMB- and $GPNMB^+$ macrophages.

https://doi.org/10.1038/s41588-026-02600-3



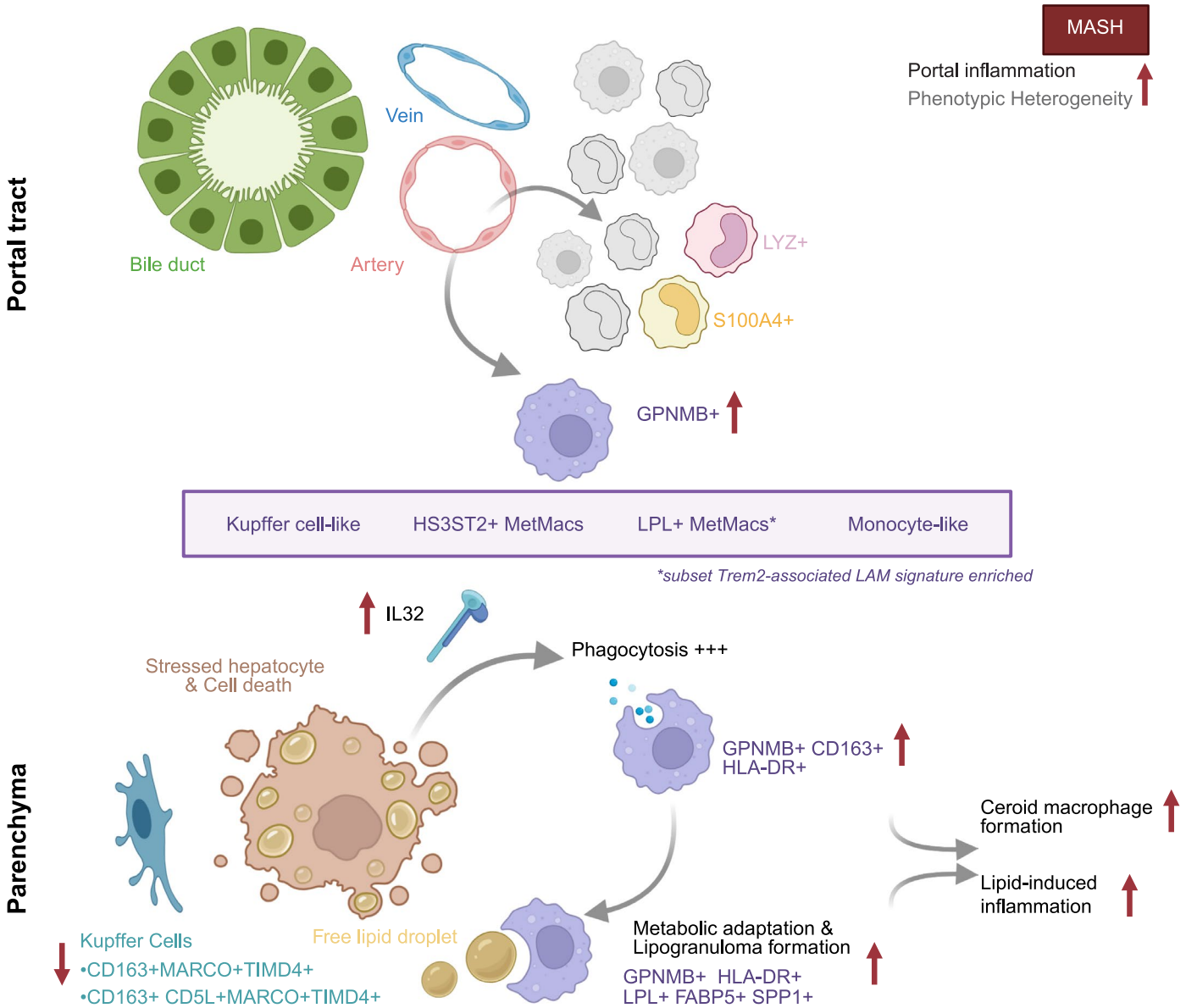
Extended Data Fig. 7 | Relevance of macrophages in hepatic transcriptome. **a**, Normalized mRNA reads from normal, MASL and different MASH stages using RNA-seq data of human liver biopsies (n = 216). *P* values of Kruskal–Wallis test indicated. **b**, Estimating the proportion of KC, MetMacs and monocytes in the RNA-seq data with the CIBERSORTx deconvolution algorithm using the snRNA-seq data, grouped based on fibrosis stage (left) and MAS score (right). Dots indicate individual patients as biological replicate, Holm–Bonferroni adjusted *p* values of two-sided t-test and median line indicated, the boxes extend from the 25th to 75th percentiles, and the whiskers extend to the minimum and maximum values. **c**, AUROC curves of a logistic regression analysis of the RNA-seq cohort using backward elimination of variables. AUROC curves showing the combined predictive model with the clinical and gene expression covariates for NAS ≥ 4. **d**, Representative Oil red O-stained images (n = 3) of THP-1 cells transfected with CRISPR–Cas9 targeting LPL or scramble control after lipid loading with 2000 µM oleic acid for 24 h.

https://doi.org/10.1038/s41588-026-02600-3



Extended Data Fig. 8 | GPNMB and CD163 disease association in a PheWAS study cohort. a,**b**, Manhattan plot of phenome-wide association study (PheWAS) analysis based on OLINK proteomics data, indicating the association between soluble GPNMB and CD163 with disease phenotypes (n = 53,030).

https://doi.org/10.1038/s41588-026-02600-3



Extended Data Fig. 9 | Proposed interactions and effects of stressed hepatocytes on macrophage phenotypes in MASH. Graphical abstract showing summary conclusions. Figure created in BioRender; Govaere, O. BioRender.com/936ff0n (2025).

nature research

Corresponding author(s): Olivier Govaere

Last updated by author(s): Dec 19, 2025

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	single nucleus RNA sequencing, nanoString GeoMx transcriptomics, nanoString CosMx, Lunaphore Comet, bulk RNA transcriptomics, SomaLogic proteomics
Data analysis	R, Python, Graphpad 9 and IBM SPSS vs29. Custom code and scripts were made available on https://github.com/govaerelab/OG_MB-MASLD-LIVER-2025 and https://github.com/MarkusBoesch93/OG_MB-MASLD-LIVER-2025 and deposited in Zenodo at 10.5281/zenodo.17988759 and 10.5281/zenodo.17980902 respectively.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The snRNA dataset is accessible from the European Genome Archive (EGA) under accession number EGAS50000000768. Burker Spatial CosMx data is accessible from GEO under number GSE312698, bulk RNA-seq data from human MASLD samples under GSE135251.
Lunaphore COMET and Bruker Spatial Biology WTA GeoMx output is accessible through the KU Leuven Research Data Repository: https://doi.org/10.48804/ERGUWX . The data presented in Fig.S4G was derived from GSE202379.

Immunohistological figures are available on figshare.com: https://doi.org/10.6084/m9.figshare.28881296.v1 https://doi.org/10.6084/m9.figshare.28881638.v1 . All other data is available on reasonable request from the corresponding authors.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	This study included 289 liver tissue and 247 serum samples and data from 53,030 participants enrolled in the UKBiobank. Biopsies for the snRNA-seq (n=18) were collected from patients undergoing bariatric surgery or cholecystectomy at the University Hospital Leuven (Belgium) and the University Hospital Antwerp (Belgium). MASH biopsies for the GeoMx spatial transcriptomics (n=8) were collected at the specialised hepatology centre Newcastle-upon-Tyne (UK); biopsies for immunohistochemistry/fluorescence (n=23), CosMx spatial molecular imaging (n=10), Comet Hyperplex (n=6) and MILAN proteome multiplexing (n=8) at the University Hospital Leuven (Belgium). Samples for the bulk transcriptomics (MASLD n=206, controls=10) and serum proteomics (n=247) have been drawn from the European NAFLD Registry (NCT04442334). Functional work was done using THP-1 cell line and human precision-cut tissue slices with a technical replicate of n=3. No statistical method was used to predetermine sample size, but a biological/technical replicate of n>=3 was deemed sufficient. All liver samples were scored by an expert liver pathologist according to the semi-quantitative NASH-CRN Scoring System. Fibrosis stage ranged from F0 to F4 (cirrhosis) and the MASLD Activity Score (MAS) was defined as the sum of the scores for steatosis, hepatocyte ballooning and lobular inflammation. At-risk MASH was defined as MAS≥4, with at least 1 point deriving from each MAS component, plus F≥2 fibrosis. Patients with a BMI ≤ 25 kg/m² with normal range liver function tests and no signs of steatosis following ultrasound examination were classified as lean controls.
Data exclusions	No data was excluded from the analyses. Patients with excessive alcohol consumption, autoimmune liver diseases, viral hepatitis, and the use of steatogenic medications were excluded.
Replication	CosMx, Comet and MILAN results were successfully used to replicate GeoMx and snRNA-seq results.
Randomization	NA
Blinding	Pathologist were blinded during liver disease scoring, but researchers were not blinded during analysis.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*
Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access & import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

CD68 916104 KP1 Mouse IgG1 Biolegend
SPP1 AF1433 Goat R&D systems
GPNMB ab227695 SP299 Rabbit Abcam
HLA-DR sc-56545 SPM289 Mouse IgG2b Santa Cruz Biotechnology
CD163 ab182422 EPR19518 Rabbit Abcam
panCK AB_1834350 AE1/AE3 Mouse IgG1 Invitrogen
CD16 NCL-CD16 2H7 Mouse IgG2a Novacastra
Lysozyme A0099 Rabbit Dako
MSR1/CD204 LS-C11521 aa325-342 Goat LSBio
ARG1 AMAB90545 CL0186 Mouse IgG1 Merck
CD31 LSC-C173974 OTI2C6 Mouse IgG2a LSBio
CD14 75181S D7A2T Rabbit Cell Signaling Technology
GPNMB CL488-66926 2B10B8 Mouse IgG1 Proteintech
Cathepsin D ab75852 EPR3057Y Rabbit Abcam
GPNMB CL594-66926 2B10B8 Mouse IgG1 Proteintech
Fc ε RIγ sc-390222 E-12 Mouse IgG2a Santa Cruz Biotechnology
CD3 ab17143 F7.2.38 Mouse IgG1 Abcam
CD8 sc-53212 C8/144B Mouse IgG1 Santa Cruz Biotechnology
CK19 M0888 RCK108 Mouse IgG1 Dako
panCK NB120-11213 C-11, PCK-26, CY-90, KS-1A3, M20, and A5 Mouse IgG1 Novus Biologicals
CK18 66187-1-Ig 1G11C4 Mouse IgG1 Proteintech
αSMA M0851 1A4 Mouse IgG2b Dako
CD20 M075501-2 L26 Mouse IgG2a Dako
MARCO PA5-64134 Rabbit ThermoFisher Scientific
TREM2 ab318262 EPR26209-22 Rabbit Abcam
TIMD4 HPA015625 Rabbit Sigma Aldrich
CD5L HPA065686 Rabbit Sigma Aldrich
CD9 ab236630 EPR23105-121 Rabbit Abcam
S100A4 ab124805 EPR2761 Rabbit Abcam
CD48 ab134049 EPR4108 Rabbit Abcam
AKR1B10 ab192865 EPR14421 Rabbit Abcam
MSR1 17858-1-AP Rabbit Proteintech
CD4 ab133616 EPR6855 Rabbit Abcam
CD45 ab10558 Rabbit Abcam
CCL19 HPA067758 Rabbit R&D systems
CD36 HPA002018 Rabbit Sigma Aldrich
AF488-labeled anti-mouse IgG1 115-545-205 Goat 1/300 Jackson ImmunoResearch
AF488-labeled anti-goat 705-545-147 Donkey 1/300 Jackson ImmunoResearch
AF555-labeled anti-mouse IgG2b A21147 Goat 1/400 Invitrogen
AF555-labeled anti-mouse IgG2a A21137 Goat 1/400 Invitrogen
AF555-labeled anti-rabbit A31572 Donkey 1/400 Invitrogen
AF555-labeled anti-rat A21434 Goat 1/400 Invitrogen
AF647-labeled anti-rabbit A-31573 Donkey 1/400 Invitrogen
AF647-labeled anti-rabbit 711-605-152 Donkey 1/300 Jackson ImmunoResearch

Validation

Validation statements on manufacture's website. Primary antibodies were validated by standard immunohistochemsitry at the Pathology Department Leuven and reviewed by an expert pathologist. Representative images are included in the paper.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

THP-1 - TIB-202, ATCC

Authentication

Authentication performed and stated by provider ATCC. No further authentication performed.

Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination

Commonly misidentified lines
(See ICLAC register)

N/A

Palaeontology and Archaeology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where

Dating methods	*they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	*For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	This study included 289 liver tissue and 247 serum samples. Patient demographics are described in Supplementary Table 1.
Recruitment	Patients have been treated and diagnosed for MASLD based on histology at specialised centres. See manuscript for more details.
Ethics oversight	This study has been approved by the relevant Ethical Committees in the participating centres and all patients having provided informed consent.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT04442334
Study protocol	Hardy, T., et al. The European NAFLD Registry: A real-world longitudinal cohort study of nonalcoholic fatty liver disease. Contemp Clin Trials 98, 106175 (2020).
Data collection	The cohort comprised 224 liver biopsies and 247 serum samples from patients with histological-proven MASLD. All liver samples were scored according to the semi-quantitative NASH-CRN Scoring System by an expert liver pathologist.
Outcomes	Histological diagnosis.

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016)

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.