Mapping vascular plasticity in liver fibrogenesis identifies novel fibrosis-associated endothelial cells in early-stage liver disease

Authors: Christina Gkantsinikoudi[1], Joshua P Dignam[1], Raju Kumar[2], Elliot Jokl[3], Wenhao Li[2], Maryna Samus[1], Stephanie Landi[4], Varinder S Athwal[4], Timothy J Kendall[5], Antal Rot[1], Jonathan A Fallowfield[5], Karen Piper Hanley[3], William Alazawi[2], and Neil P Dufton[1*]

Affiliations:

[1]William Harvey Research Institute, Queen Mary University London, London, UK.

[2]Blizard Institute, Queen Mary University London, London, UK

[3] University of Manchester, Manchester, UK.

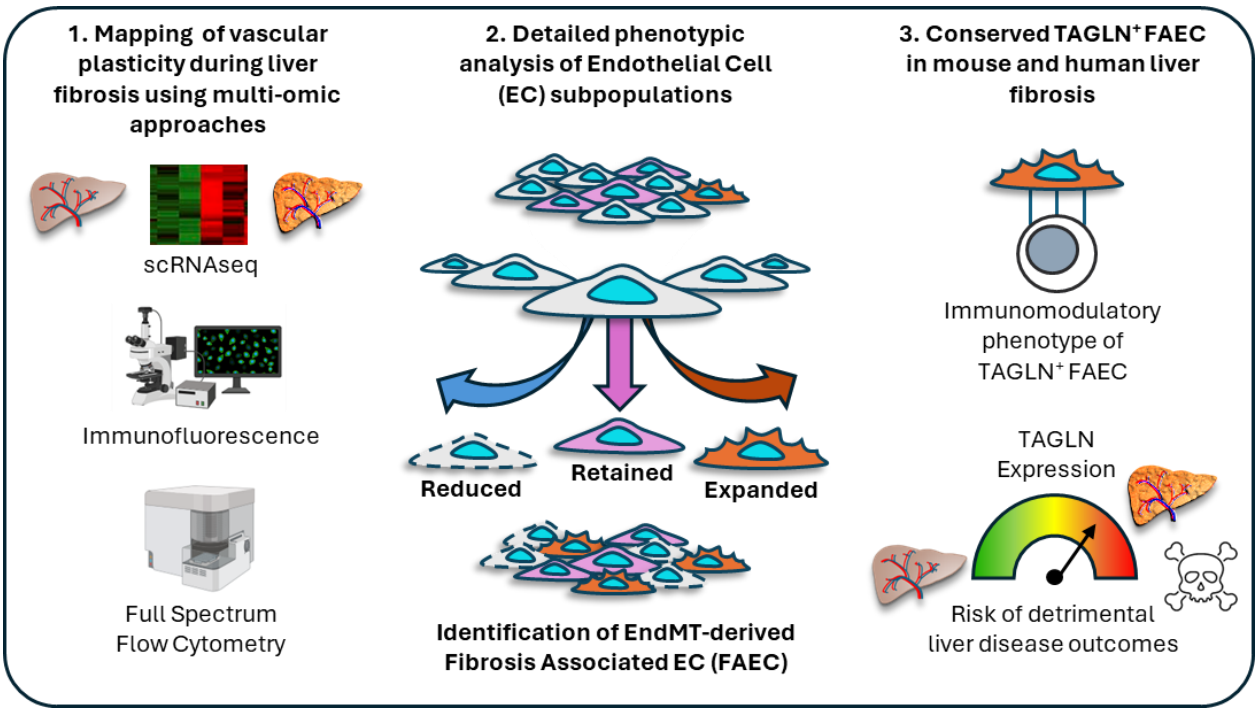
[4] Manchester University NHS Foundation Trust, Manchester, England, UK

[5] Centre for Inflammation Research, Institute for Regeneration & Repair, University of Edinburgh, Edinburgh

*Corresponding author. Email n.dufton@qmul.ac.uk

One Sentence Summary: Phenotyping fibrosis-associated endothelial cells reveals pro-fibrotic and immunomodulatory subpopulations at early stages of liver disease.

Graphical Abstract



Abstract:

Vascular plasticity is a crucial biological asset enabling our bodies to rapidly adapt to infections and acute inflammation. However, repeated insult during chronic disease can result in these vascular adaptations becoming irreversible, thereby driving disease progression and fibrosis. This study aimed to understand if phenotypic changes in endothelial cell (EC) identity could be indicative of progressive fibrosis and thereby offer new diagnostic and therapeutic opportunities for patients with metabolic dysfunction-associated steatotic liver disease (MASLD). Previous research has documented that a significant shift in EC transcriptomic signature occurs during liver fibrosis in both pre-clinical models and patients. However, the protein expression profile, phenotype and functional role of these new EC subpopulations that are induced during fibrogenesis is unclear. In this study, we integrate high-resolution imaging, proteomic and transcriptomic analysis which collectively highlight a central role for endothelial-to-mesenchymal transition (EndMT)-induced EC plasticity in the derivation of 'fibrosis-associated' EC (FAEC). We demonstrate that: 1) full spectrum flow cytometry can provide new opportunities to categorize and phenotype EC subpopulations, 2) two distinct EndMT-derived FAEC subpopulations expand during fibrogenesis; THY1.2^{+}ICAM1^{+} and TAGLN^{+}MCAM^{+} EC that display unique immunomodulatory and metabolic phenotypes, 3) TAGLN^{+} FAEC are a conserved, pro-fibrotic cell type arising at early stages of MASLD, and 4) increased hepatic expression of TAGLN is significantly associated with detrimental patient outcomes at all stages of liver disease. This study will pave the way for the development of FAEC-specific diagnostic and therapeutic approaches to tackle progressive fibrotic disease.

INTRODUCTION

Fibrosis is the principal predictor of major adverse liver outcomes and of mortality in metabolic dysfunction-associated steatotic liver disease (MASLD). MASLD is the most common cause of chronic liver disease worldwide and its prevalence has risen in parallel with rising levels of obesity and type II diabetes*(1)*. Despite major advances, including two drugs licensed in the United States*(2, 3)*, the two main unmet clinical challenges in MASLD are sensitive and accurate diagnostics and treatments with clinically significant effect sizes, both of which rely on improved understanding of pathology including fibrogenesis.

Zonation of liver lobules containing specialised subpopulations of hepatic endothelial cells (EC), hepatocytes and immune cells, is critical for vascular, metabolic and immune homeostasis in the liver *(4–6)*. Vascular zonation is dramatically impacted by chronic disease, with several studies documenting the changing EC landscape and the derivation of novel endothelial cell clusters between healthy and fibrotic liver tissue in mice *(7, 8)* and patients *(9, 10)*. Endothelial-to-mesenchymal transition (EndMT) is a central endothelial plasticity pathway responsible for phenotypic vascular changes *(11–13).* It is defined as the process by which EC lose expression of endothelial lineage markers, such as CD31, and upregulate mesenchymal markers such as transgelin (TAGLN)*(14)*, THY1.2*(15)* (also termed CD90.2) or platelet-derived growth factor receptor α (PDGFRα)*(16)*. This process is predominantly driven through canonical transforming growth factor β (TGFβ) signalling and induces phenotypic changes such as enhanced expression of cytoskeletal and matrix-associated factors, including vimentin and collagen I and III *(8, 17)*. Importantly, EndMT has been demonstrated to occur in both pre-clinical models of liver fibrosis *(11, 18)* and end-stage liver disease in patients *(11, 12)*. The role of EndMT as a pathogenic driver, rather than a consequence of fibrogenesis, and its predictive value regarding disease progression

remains ill-defined. We hypothesise that FAEC are EndMT-derived subpopulations that instigate both pro-fibrotic and pro-inflammatory pathways, driving the early transition from tissue injury to progressive liver fibrosis.

To address our hypothesis, we developed a multi-omic approach integrating single cell RNA sequencing (scRNAseq), immunofluorescent microscopy (IF) and full spectrum flow cytometry (FSFC) to identify and phenotype EC subpopulations during different stages of carbon tetrachloride (CCl_4)- and Western diet (WD)-induced mouse models of liver fibrosis. We characterise the induction of two distinct EndMT-derived FAEC populations in our pre-clinical studies and translate our findings to three cohorts of MASLD patients. We show $TAGLN^+$ FAEC are a common immunomodulatory EC subpopulation that share transcriptional characteristics, arising in both mouse models and patient samples, that are significantly elevated at early-stage disease prior to clinical appearance of fibrosis. Furthermore, the elevated expression *TAGLN* is a significant risk factor for liver decompensation events, even in the absence of fibrosis, and confers a significant additive effect for detrimental patient outcomes.

Taken together, we believe this study will pave the way for developing FAEC-specific diagnostic and therapeutic approaches to tackle progressive fibrotic liver disease.

RESULTS

Mapping EC plasticity in healthy and fibrotic liver tissue

We analyzed a publicly available dataset (GSE145086) enriched for hepatic immune cells, hepatic stellate cells (HSC) and EC in control C57BL/6 mice or following 2 and 4 weeks of CCl_4-induced liver fibrosis*(19)* (Fig. 1A). We identified four hepatic EC clusters which expressed combinations of genes known to be characteristic for each zone *(4, 6)*; Zone 1 EC (Cluster 8), Zone 2 EC (Clusters 0 and 1) and Zone 3 EC (Cluster 6; Fig. 1B). Each cluster also displayed distinct expression profiles for five EC-identity genes that are expressed as cell surface proteins: *Pecam1* (CD31), *Mcam* (MCAM), *Thbd* (TM), *Emcn* (EMCN) and *Lyve1* (LYVE1); (Fig. 1B-C). Immunofluorescence intensity profile for each of these markers, across the 200 µm distance between a portal and central vein confirmed this zonal expression in healthy liver tissue (Fig. 1D-E).

After 4 weeks' treatment with CCl_4 to induce fibrosis, there was a marked shift in the abundance of mid-lobular EC populations during fibrogenesis (Fig. 1F and G), with pseudotime analysis showing reduction in the Zone 2 (Cluster 0) seen in health and expansion of what we have termed fibrosis-associated endothelial cells (FAEC) phenotype (Cluster 1) that transition towards Zone 1 and Zone 3 clusters (Fig. 1H). These expanding FAECs display enriched expression of a panel of EndMT-associated, TGFβ-driven, genes after 4 weeks' CCl_4 treatment (Fig. 1I). Notably, the expression of the EC identity genes also changes dynamically in response to CCl_4 treatment (Fig. 1I), highlighting the plasticity of hepatic EC during fibrogenesis.

To study these shifts in EC identity and phenotype, we treated male C57BL/6 mice with peanut oil vehicle alone (sham), CCl_4 for 4 weeks (4w; early fibrosis), 8 weeks (8w; established bridging

fibrosis) or treated for an initial 4 weeks and then allowed to recover from tissue fibrosis for a further 4 weeks (Recovery Group; Fig. 2A-B).

Consistent with the transcriptional changes described above (Fig. 1), quantification of widefield IF images revealed significant changes in vascular activation showing the increase in co-localisation of vascular identity markers (CD31, TM and EMCN; Fig. 2C and Fig S1A-D) together with TAGLN (Fig. 2C & Fig S1A & E) resulting in a robust induction of $CD31^{+}TAGLN^{+}$ EndMT (Fig. 2D). EndMT and vascular activation were preferentially localized in Zone 3 peri-central regions (Fig. 2E). Taken together, we show that endothelial plasticity pathways are activated in distinct zones during liver fibrogenesis and tissue recovery, validating the suitability of this model to assess the role of EndMT in specific EC subpopulations leading to the derivation of FAEC phenotypes.

To observe phenotypic changes in multiple EC subpopulations simultaneously, a 14-marker panel was developed for FSFC*(20)* enabling us to profile the expression of 6 EC identity (CD31, MCAM, TM, EMCN, LYVE1 and PDPN), activation (CD34, ICAM1 and LY6A) and EndMT (TAGLN, THY1.2 and PDGFA) markers (Fig. S2A). All four treatment groups were concatenated and $CD31^{+}CD45^{-}$ ECs (Fig. S2B) were clustered based on expression of the identity markers, producing twelve populations (Concatenated data; Fig. 2F and split by treatment group Fig. 2G) that were subcategorized each based on change in abundance. **Reduced**: Significantly reduced abundance following CCl_4 treatment compared to Sham, **Retained**: no significant change following CCl_4 treatment compared to Sham and **Expanded**: significantly increased abundance compared to Sham (quantified in Fig. 2H). We show a striking alignment between the transcriptomic (Fig. 1G) and proteomic (Fig. 2H) plasticity data of hepatic EC but also highlight

the considerable proteomic heterogeneity, which offers further insight into the phenotypic attributes of EC subpopulations during fibrogenesis and recovery.

Phenotyping EC subpopulations during fibrogenesis and tissue recovery

'Reduced' ECs were the predominant subcategory of EC in healthy liver, comprising three clusters (1-3; Fig. 3A), which become less abundant following 4 and/or 8w CCl_4 and returned towards basal levels in the Recovery group (Fig. 3B-D). All 3 Reduced EC clusters are $LYVE1^+$ $ICAM1^-$ and can be distinguished based on expression of TM and EMCN (Fig. 3E), suggesting that the majority of these EC are situated in Zones 2 and 3 based on the expression profiles in Fig.1. These ECs had limited activation profiles, with LY6A induced in cluster 2 only (Fig. 3F), and no EndMT markers in response to CCl_4 (Fig. S3A-C). The key phenotypic characteristics of the clusters in this category are summarised in Fig. 3G.

Five 'Retained' clusters (4-8; Fig. 3H) did not change in abundance between healthy and CCl_4-treated mice (Fig. 3I-M) although some clusters did express markers of activation and EndMT: $MCAM^+LY6A^+TAGLN^+$ cluster 4 and $PDGFR\alpha^+$ cluster 6 (Fig. S3 D-H). Based on their expression profiles for $Ly6A^+$ $ICAM1^-$ and differential expression of Lyve1 and EMCN (Fig 3N-P) these EC do not appear to be associated to a specific zone.

Four Expanded EC clusters (9-12; Fig. 4A) significantly increased in abundance between healthy and CCl_4-treated mice (Fig. 4B-E). We collectively define these clusters as FAEC as they are instigated in response to tissue fibrosis while returning to basal levels in the Recovery group. All expanded clusters expressed elevated markers of activation and EndMT (Fig. S3I-L). Notably, Clusters 9 and 10 expressed 'classical' EndMT markers, downregulating the expression of the EC identity markers EMCN and TM, respectively, while inducing the expression of

THY1.2/PDGFRα, with increased inflammatory activation via ICAM1 (Fig. S3I&J). The most prevalent Expanded clusters, 11 and 12, displayed 'non-classical' EndMT phenotypes exhibiting both increased expression of EC identity markers, including CD31 and TM, together with enhanced TAGLN and LY6A expression (Fig. S3K&L). The idiosyncrasies of these 4 clusters are further emphasised by the differential contour plots (Fig. 4F) and histograms between the Sham and 8w CCl_4 groups (Fig. 4G) and are summarised in schematics (Fig. 4H&I).

In 8w CCl_4- and Sham-treated mouse liver sections, THY1.2$^+$CD31$^+$ and TAGLN$^+$CD31$^+$ FAEC are mutually exclusive cells within Zone 2 and Zone 3 regions of the liver, respectively (Fig. 4J&K). To determine whether TGFβ-induced EndMT can give rise to both FAEC populations, isolated healthy mouse liver EC stimulated *in vitro* with 10 ng/ml TGFβ2 for 48 hours also showed differential expression of THY1.2 and TAGLN (Fig. 4L&M).

Together, these data demonstrate EC heterogeneity and plasticity during fibrosis and tissue recovery and identify two distinct types of FAEC. While THY1.2$^+$ and TAGLN$^+$ FAEC may arise from a common TGFβ-driven EndMT pathway, they may also originate from separate healthy EC subpopulations. To test this assertion, we addressed the potential for Reduced clusters to transition towards Expanded FAEC clusters using Wishbone trajectory analysis.

Derivation of FAEC using Wishbone trajectory mapping

To determine whether EC in Reduced clusters transition towards Expanded FAEC, we first assessed their overlapping traits. Reduced clusters 1 and 3 and Expanded clusters 9 and 10 overlapped in expression of LYVE1 (Fig. 5A) and LY6A (Fig. 5B) in Sham animals and displayed marked plasticity following treatment. Notably, a subset of cluster 10 became THY1.2$^+$ (Fig. 5B) while there was a complete transition for cluster 9 towards a THY1.2$^+$ phenotype in response to

CCl_4 treatment (Fig. S3I&J). Wishbone trajectory analysis for the four clusters at all fibrosis stages, showed that LYVE1^{high} ICAM1^{low} THY1.2^{low} EC converge at a branch point which enables them to either shift towards a FAEC phenotype of clusters 9 and 10 in response to 8w CCl_4 treatment or return to a healthy phenotype during recovery (Fig. 5C). The inverse relationship of these markers is highlighted in the trajectory heatmaps showing that LYVE1 and THY1.2 appear to be mutually exclusive, while ICAM1 expression is enhanced at a later stage of transition to FAEC (Fig. 5D). The putative transition pathway is summarised schematically (Fig. 5E).

Similarly, Reduced Cluster 2 and the two TAGLN$^+$ FAEC clusters 11 and 12 displayed overlapping expression of MCAM and LY6A (Fig. 5F) and TAGLN (Fig. 5G), which were all elevated following 4 and 8 weeks' CCl_4 treatment. Incremental acquisition of the TAGLN$^+$MCAM$^+$ FAEC phenotype during fibrosis is only partially restored after recovery, suggesting that these FAEC may persist in the tissue even after the insult is removed (Fig. 5H). The trajectory heatmap highlights the persistence of TAGLN, LY6A and MCAM in both the fibrogenic and recovery branches (Fig. 5I), suggesting these FAEC have decreased plasticity, impairing their capacity to recover a healthy phenotype (Fig. 5J).

Immunomodulatory function of FAEC

To assess the influence of these two distinct FAEC subpopulations on the inflammatory response within the liver during fibrogenesis and recovery, we explored the transcriptomic signatures of EC enriched for either *Thy1* or *Tagln2* in the scRNAseq dataset and compared them with the phenotypic characteristics in our FSFC data. Abundance of these cell types was similar: *Thy1*$^+$ FAEC comprising ~1% of total EC by scRNAseq and THY1.2$^+$ FAEC ~5% of total EC by FSFC and, *Tagln*$^+$ (scRNAseq) and TAGLN$^+$ (FSFC) FAEC comprising ~20%. Gene expression in

Thy1^{+} FAEC was indicative of leukocyte interaction and activation (Fig. 6A) and THY1.2^{+} FAEC expressed increased levels of ICAM1 following CCl_4 treatment in both cluster 9 and 10, and reduced levels of the anti-inflammatory glycoprotein EMCN in cluster 9 only (Fig. 6B). *Tagln*$^{+}$ FAEC displayed more specific enrichment for metabolic and inflammatory genes (Fig. 6C) and TAGLN^{+} FAEC were characterised by upregulation of MCAM and LY6A (Fig. 6D). In line with the localisation and phenotype of FAEC CD45^{+} leukocytes (Fig. 6E) are preferentially recruited to peri-central regions (Fig. 6F) following 4 and 8 weeks' CCl_4 treatment.

TAGLN^{+} FAEC are conserved in Western diet-fed mice and MASLD patients

Having characterised the induction of FAEC subpopulations in a murine CCl_4-induced liver fibrosis, we sought to determine if FAEC were also an early characteristic of a murine dietary model of MASLD induced by feeding animals a Western Diet (WD). There was significant uptake of lipid within the liver (Fig. 7A and B) and moderate hepatic fibrosis (Fig. 7C and D). In keeping with the instigation of mild tissue fibrosis, compared to the CCl_4 model, the impact on vascular architecture (Fig. S4A and B) and EC identity markers was less pronounced, but still resulted in significantly increased EMCN, TM and TAGLN following 12w of WD (Fig. S4C).

Assessment of EC subpopulations by FSFC distinguished twelve hepatic EC clusters (Fig. 7E). Consistent with our earlier findings (Fig. 2F), we identified phenotypically distinct Reduced, Retained and Expanded subpopulations (Fig. S5A-L) albeit the differences in proportion were more subtle than CCl_4 (Fig. 7F). Reduced clusters 1 and 2 (Fig. 7G), and Expanded clusters 11 and 12 (Fig. 7H) shared overlapping characteristics when compared on the UMAP (Fig. 7I), such as ICAM1 and TAGLN expression (Fig. 7J). Wishbone analysis revealed a shared trajectory towards

a TAGLN$^+$ FAEC phenotype (Fig. 7K and L), which occurred more gradually than in response to CCl_4 (Fig. 7M).

Finally, we examined whether the induction of TAGLN expression was associated with disease severity and patient outcomes in human MASLD. In liver biopsy samples from patients who underwent bariatric surgery, there was a significant expansion of both the number of *TAGLN*$^+$ EC and expression of *TAGLN*$^+$ in patients with compared to those without fibrosis (Fig. S6A). IF staining for TAGLN$^+$CD31$^+$ EndMT cells (Fig. S6B) showed a significant increase in patients with early stages (F0/1) of disease (Fig. S6C) compared to healthy control. In keeping with the phenotype of TAGLN$^+$ FAEC in murine models, we found enrichment of *MCAM*, *TAGLN* and *THY1* (Fig. S6D) with GO analysis showing similar enrichment to the mouse data in vascular remodelling, pro-fibrotic and pro-inflammatory pathways (Fig. S6E). In spatial transcriptomic analysis of four patient biopsies (Supplementary Table 1). TAGLN$^+$ FAEC co-localised with ECM and inflammatory macrophages with regions of fibrosis (Fig. 8A). In line with patient scRNAseq data (Fig. S7A-C), the appearance of TAGLN$^+$ FAEC was apparent at early-stage disease before observable fibrosis and persisted in patients with cirrhosis (Fig. S7A). Comparison of TAGLN$^+$EC vs TAGLN$^-$ EC transcriptomic signatures revealed that pro-fibrotic (Collagen) and pro-inflammatory (ICAM) pathways were preferentially enriched in TAGLN$^+$ FAEC (Fig. 8B and Fig. S7B). Detailed CellChat analysis identified the expression of *THBS, LAMA* and *COL* as potentially unique sender/receiver partners (Fig. 8C and Fig. S7C).

To evaluate the potential clinical impact of elevated TAGLN expression in MASLD patients, we utilised the SteatoSITE*(21)* database to assess patient outcomes. High *TAGLN* expression (above 26.76 cpm) in liver biopsies revealed patients at a significantly elevated risk of all-cause mortality compared to those with lower expression levels (Fig. 8D). Furthermore, patients with high levels

of *TAGLN* but no fibrosis (Fig. 8E, F0; green line) had a significantly higher risk of hepatic decompensation than patients with cirrhosis but low *TAGLN* expression (Fig. 8E, F4; blue line), suggesting that *TAGLN* expression has the potential to be an early and fibrosis-independent risk factor of detrimental patient outcomes.

Overall, our data elucidate the derivation of two distinct FAEC subpopulations by integrating scRNAseq and IF with a novel FSFC analysis of EC plasticity. We show that $TAGLN^{+}MCAM^{+}$ EC are the predominant FAEC phenotype in two mechanistically distinct models of murine liver fibrosis and are also present in patients with MASLD. Notably, these FAEC share common pro-fibrotic and pro-inflammatory phenotypes, which have potential immunomodulatory functions by co-localising with leukocytes (summarized in Fig. 8F). Finally, we reveal that the induction of *TAGLN* expression is significantly associated with detrimental clinical outcomes in liver disease patients, independent of fibrosis severity, supporting the hypothesis that $TAGLN^{+}$ FAEC play an underappreciated role in disease pathogenesis.

DISCUSSION

The role of vascular heterogeneity in health and disease has come into sharp focus in recent years, with considerable advances in single-cell transcriptomic*(7, 22)* and proteomic*(6)* profiling revealing remarkable insight into organotypic and pathological endothelial heterogeneity. In this study, three broad categories of liver EC in liver fibrosis: 'Reduced', 'Retained' and 'Expanded' have been defined and phenotyped. We find that the Reduced clusters that express $CD31^{+}LYVE1^{+}MCAM^{neg}$ (some also expressed EMCN) localise to mid-lobular (Zone 2) and peri-central (Zone 3) regions, the same regions where the Expanded cells arise from, in keeping with known CCl_4 damage pattern*(23, 24)*. Using Wishbone trajectory analysis, we were able to show that they can change their phenotype and contribute to the development of Expanded populations. All Expanded clusters expressed EndMT traits as well as enhanced inflammatory profiles in response to CCl_4-induced liver fibrosis. These EC clusters were the most highly plastic of all the EC subpopulations assessed, dramatically changing their abundance and phenotype in response to both CCl_4 timepoints and following tissue recovery, thus we designated them as FAEC. Finally, the association of these markers with diseases severity and outcome in human disease strongly suggest translational potential of these findings*(25)*.

EC are highly plastic*(26)* and their regulation of reparative or pathogenic pathways *(27, 28)* represent new diagnostic and therapeutic potential in human disease. We and others have determined that TGFβ-induced EndMT cells are an important component of the fibrogenic hepatic vascular niche*(11, 12, 29, 30)*, however, the functional consequences of these phenotypic changes remain elusive. There is currently no consensus on how to define EC heterogeneity in the liver, and the application of multiple techniques has led to a variety of distinct subpopulations, ranging from immunohistochemistry studies*(31)* detailing type 1 and type 2 LSEC to single-cell RNAseq

identifying five to seven endothelial populations in mice*(32)* and humans*(10)*. A key advancement of our study is the complementary use of available scRNAseq data with multiple identity marker and high-throughout proteomic phenotyping through FSFC. While previous studies applied lineage tracing approaches to validate the occurrence of EndMT*(7, 12, 17, 33)* in several tissues and disease states, however, they are restricted to an individual identity marker. In contrast, our approach provides a more comprehensive analysis of EC plasticity, giving large-volume, single-cell data with multiplexing capabilities, high-dimensional reduction analysis and predictive algorithms, such as Wishbone trajectory*(20, 34)*, to trace the derivation of novel FAEC populations.

The most prominent distinguishing feature of these FAEC clusters was the expression of either THY1.2$^+$ or TAGLN$^+$. THY1.2$^+$ FAEC express higher levels of ICAM1 and do not express EMCN and LYVE1. THY1.2 is a member of the immunoglobulin supergene family and is commonly associated with smooth muscle cells. EC expressing *Thy1* are associated with inflammation-induced angiogenesis*(35)* and trans-endothelial leukocyte migration*(36)*. ICAM1 is a key regulator of leukocyte rolling and adhesion to LSEC*(37)*. Conversely, loss of EMCN and LYVE1 can also influence immune cell interactions, as both are documented to play protective roles in preventing aberrant leukocyte recruitment, respectively *(38, 39)*.

TAGLN$^+$ FAEC were not only the most abundant but were also conserved across species and displayed reduced plasticity, which may account for their persistence across all stages of liver disease assessed. TAGLN, a member of the calponin superfamily, is an actin-binding protein which can regulate the cytoskeleton*(40)* and ectopic TAGLN expression is hallmark of EndMT *(26, 41)*. The upregulation of MCAM may point towards its role as an adhesion molecule that promotes cell-ECM and cell-cell interactions*(42)*. MCAM is a well-characterised biomarker for

advanced liver disease*(43)* and it is known to induce pro-coagulant responses in metastatic tumours*(44)*. Upregulation of Ly6A is consistent with previous reports detailing the expansion of Ly6A^{+} LSEC during liver injury *(45, 46)*. Notably, TAGLN^{+} MCAM^{hi} FAEC have a conserved phenotype that is induced in early stages of liver injury in both CCl_4- and 12-week WD-induced mouse models and in patients without observable fibrosis, suggesting that the appearance of TAGLN^{+}FAEC could be a novel, sensitive marker of disease progression.

TAGLN^{+} FAEC co-localise with regions enriched for ECM and infiltrating leukocytes in both mouse and human fibrosis. In the large SteatoSite cohort of histologically confirmed MASLD *(21)*, we find that high hepatic *TAGLN* expression is significantly associated with adverse clinical outcomes, at all disease stages extending prognostic insight beyond fibrosis severity.

Taken together, we have identified distinct EndMT-derived FAEC with defined localisation and characteristics which have the capacity to reshape the vascular and immune landscape of the liver. TAGLN^{+} FAEC represent an exciting opportunity to develop new diagnostic and therapeutic approaches to tackle the growing clinical burden of chronic liver disease.

MATERIALS AND METHODS

Animals

All procedures used in this study were approved by the Animal Use and Care Committee of Queen Mary University of London and were in accordance with national and international regulations (Project license: PP0505769).

8-10 weeks old wild-type (WT) C57BL/6 male mice were purchased from Charles River Laboratories UK Ltd. (Kent, UK). For all *in vivo* experiments, power calculations were performed using preliminary data to calculate the appropriate number of animals to be included in the study.

Murine models of liver injury

Carbon tetrachloride (CCl_4; 289116-100ML, Sigma Aldrich) was diluted 1 in 5 in peanut oil (P2144-250ML, Sigma) and was administered with an intraperitoneal (IP) injection at a dose of 0.8mg/Kg in 100µl. Injections were administered twice a week for 2, 4 or 8 weeks. For the recovery experiments, mice were injected with CCl_4 for 4 weeks and were afterwards left to recover for an additional 4 weeks. As a control, no treatment or Sham injections of 100µl peanut oil were used.

To model a western diet (WD) and induce hepatic steatosis and fibrosis, mice were fed the modified Clinton-Cybulsky cholesterol diet (58R6, IPS Product Supplies) with the addition of carbohydrates (42g/L) consisting of 55% fructose (F0127-1KG, Sigma-Aldrich) and 45% sucrose (S9378-1KG, Sigma-Aldrich) in the drinking water for 12 weeks At each stated time point, animals were sacrificed, followed by cardiac perfusion with 3ml sterile deionized phosphate-buffered saline (PBS) (BR0014G, ThermoFisher Scientific) with 0.5% bovine serum albumin (BSA) (BP9702-100, Fisher Scientific) and 2mM EDTA (E7889-100ML, Sigma Aldrich) through the LV. Individual liver lobes were removed for immunohistochemistry, tissue digestion for tissue culture or flow cytometry and RNA extraction.

Immunohistochemistry

Tissues were fixed for 2 hours in 4% paraformaldehyde before being prepared for either OCT or paraffin embedding. Tissue sections were cut at 10µm or 5µm, respectively. Samples were initially blocked in 3% BSA diluted in PBS +0.1% Tween 20 (P1379-100ML, Sigma) (PBST) for 1 hour at room temperature (RT). The samples were incubated with primary antibodies (Supplementary table 2) overnight (diluted in 1% BSA). They were then washed 3 times with PBST and incubated with appropriate secondary antibodies (Supplementary table 2) for 1 hour at RT. The slides were washed and mounted with FluoroMount G (00-4958-02, Invitrogen). Images were acquired using an inverted confocal laser scanning microscope at a 10x magnification.

Picrosirius Red Staining

Paraffin-embedded sections were deparaffinised with Histoclear solution (NAT1330, Scientific Laboratory Supplies) and rehydrated. For staining, the Abcam Picrosirius red stain kit was used (ab245887, Abcam). The slides were then incubated in phosphomolybdic acid for 8 minutes and were transferred to picrosirius red stain for 60 minutes. They were then washed 3 times with acetic acid and were mounted using the Entellan (107960, Merck) mounting medium. The samples were visualised using the Olympus CKX53 light microscope (Olympus, Japan) and images were acquired at 4x magnification.

To analyse the data, the whole region from 2 sections per sample was quantified. The percentage (%) of picrosirius red-positive area was calculated using Image J and the colour deconvolution macro*(47)*.

Tissue Digestion

Liver samples were mechanically disrupted until homogenised using scissors. The homogenates were then further digested enzymatically in 3ml digestion buffer containing Hank's Balanced Salt

Solution (HBSS) with $CaCl_2$ and $MgCl_2$ (14025092, Gibco) with 20mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) (15630-080, Gibco), 2mg/ml collagenase (17018-029, Gibco), 8U/ml dispase II (04942078001, Roche) and 50U/ml DNase (18047-019, Invitrogen). For digestion, the homogenates were incubated at 37°C for 30 minutes on a shaker. The samples were then filtered twice through a 70µm cell sieve (431751, Scientific Laboratory Supplies). The cell pellets were then resuspended for tissue culture or full spectrum flow cytometry.

Staining for Full Spectrum Flow Cytometry (FSFC)

Single cell suspensions were transferred to a V-shaped 96-well plate (50µl/ well). The cells were washed with 200µl PBS and were resuspended in 50µl Fc blocking solution (130-092-575, Miltenyi Biotec) containing the Live/Dead™ Fixable Blue stain (L23105, Invitrogen) at a dilution of 1 in 5,000 for 30 min at 4ºC. The cells were washed twice with 200µl PBS and were then incubated with cell surface binding antibodies for 30 min at 4ºC. FSFC primary antibody details are listed for the endothelial cell panel (Supplementary table 3). The samples were then washed with PBS and centrifuged at 400g for 2 minutes at 4°C. For fixation and permeabilization, the cell pellets were resuspended in the intracellular fixation and permeabilization buffer (00-5521-00, Invitrogen) and the plate was incubated for 20 min at room temperature (RT). The samples were then washed with 200µl permeabilization wash buffer (00-8333-56, Invitrogen) and centrifuged at 400g for 2 minutes. The cell pellets were resuspended in the intracellular binding antibody solution and incubated for 30 minutes at RT. The samples were washed with permeabilization wash buffer, followed by centrifugation at 400g for 2 minutes. The cell pellets were resuspended in 200µl PBS with 2% Foetal Bovine Serum (FBS) (A4766801, Gibco) and 2mM EDTA and were stored at 4°C

for acquisition. Appropriate fluorescence minus one (FMO) controls were also prepared in parallel. For comprehensive details regarding this protocol refer to Gkantsinikoudi *et al., 2025(20)*.

FSFC data acquisition and analysis

FSFC data acquisition was performed using the Cytek® 4 laser Aurora (UV, V, B, R) spectral flow cytometer and the SpectroFlow 3.3.0 software (Cytek® Biosciences, US). For spectral unmixing, single stain reference control beads were used for all fluorescent antibodies. Heat killed cells (90ºC) were used to prepare the reference control for the Live/Dead Fixable Blue stain. Multiple autofluorescence extraction was used to account for the autofluorescence signatures of different cell populations. For this, a tissue- and treatment-specific unstained sample was used.

All data was analysed using the OMIQ Software (Dotmatics, US). Cellular debris was excluded from the analysis by appropriate gating using the FSC-A vs SSC-A plot, followed by exclusion of duplets using the FSC-A vs FSC-H plot. Live cells were selected based on the Live/Dead Fixable Blue stain. Then, in the CD45 vs CD31 plot, the $CD45^+$ and $CD45^-$ populations were analysed separately. In the $CD45^-$ gate, EC populations were analysed, while in the $CD45^+$ gate, analysis of immune cell populations was performed. Both classical gating analysis and dimensionality reduction analysis was used.

For dimensionality reduction, the first step of the analysis pipeline involved data quality control, which was performed using the FlowAI algorithm*(48)*. Then, manual compensation was performed where necessary. For unsupervised clustering, the FlowSOM algorithm*(49)* and the cells were clustered based on the endothelial markers CD31, TM, EMCN, LYVE1, MCAM and PDPN to generate endothelial clusters with different identities. For cluster visualisation, uniform manifold approximation and projection for dimension reduction (UMAP) plots*(50)* were used. UMAP were generated using the Euclidean metric composed of the same six EC markers with a

minimum distance of 0.1, nearest neighbours of 15. Contour plots, histograms and Wishbone trajectory analysis were generated using OMIQ software as detailed*(20)*.

Image analysis and profile plot generation

For image analysis, the Volocity 3D Image analysis software (PerkinElmer, US) was used. The total area for different markers was calculated and normalised using the total tissue area or total number of nuclei. For EndMT quantification, the area of colocalization of an endothelial and a mesenchymal marker was calculated.

To visualise and quantify the distribution of different markers across the three zones of the liver tissue, profile plots were generated using Image J and the Macro plot line profile multicolour macro written by Kees Straatman, University of Leicester (Leicester, UK*(51)*), which was modified by Dr Paul Imbert. Regions of interest were selected that spanned from a portal vein (zone 1) led to a central vein (zone 3; Fig S1A). A 200 µm line was used to measure the fluorescence intensity between these two structures. The distance between the portal and central veins was expressed as a percentage and segmented into zone 1 (0-33% of line), zone 2 (34-66% of line) and zone 3 (67-100% of line) for normalisation across all samples. Area under the curve was used to calculate the fluorescence signal under the curve within each zone.

Single cell RNA sequencing Data Acquisition (mouse carbon tetrachloride study)

Raw 10x Genomics single cell RNA sequencing data from the GSE145086 dataset were obtained from the Gene Expression Omnibus and processed in R. For each sample, the standard 10x output files (gene-cell count matrix, barcode file, and gene annotation file) were imported, and sparse matrices were constructed alongside their corresponding barcodes and gene identifiers. Gene symbols were inspected for redundancy and made unique where necessary, and each cell was assigned a unique sample specific identifier to enable accurate provenance tracking. Per cell

metadata were generated for every sample, and individual datasets were converted into Seurat objects. Following quality checks to ensure alignment of barcodes, metadata, and gene features, all samples were merged into a single unified Seurat object.

For trajectory reconstruction, Seurat objects were converted into Monocle3 cell data sets using the as.cell_data_set function from SeuratWrappers. Cluster and partition information from Seurat were retained. The principal graph was learned using learn_graph, and pseudotime was estimated using order_cells, setting the endothelial cell cluster representing the quiescent state as the root. Pseudotime values were extracted and used to infer lineage progression among endothelial subpopulations. Only Monocle3's internal pseudotime calculations were used for subsequent analysis; graphical visualization was performed in R but excluded from final interpretation.

Genes significantly upregulated in selected comparisons were converted to Entrez IDs using the bitr function in the clusterProfiler (v4.10.0) package (OrgDb = org.Mm.eg.db). Gene Ontology (GO) enrichment analysis for Biological Processes (BP) was performed using enrichGO(with the Benjamini–Hochberg adjusted p-value threshold of 0.05. The most significantly enriched biological processes were used for interpretation of functional differences among endothelial subpopulations.

All analyses were conducted in R (v4.3.3) on Apocrita HPC system (http://doi.org/10.5281/zenodo.438045). The following packages were used: Seurat (v5.0.3), Monocle3 (v1.3.7), SeuratWrappers (v0.3.1), clusterProfiler (v4.10.0), org.Mm.eg.db (v3.17.0), ggplot2 (v3.5.1), and dplyr (v1.1.4).

Marker Gene Identification and Differential Expression

Cluster-specific marker genes were identified using Seurat's FindAllMarkers function, retaining genes expressed in at least 25% of cells within a cluster and exhibiting a log fold change greater

than 0.25. Differential expression analyses between experimental conditions were performed using FindMarkers with test.use = "MAST", and significant genes were defined by adjusted $p < 0.05$ and $\log_2 FC > 0.25$.

All results were exported as CSV files for reproducibility and integration with pathway enrichment analysis.

CiteSeq sequencing Bariatric Surgery Patients

We performed antibody-augmented single cell RNA for cell populations sampled from liver tissue sampled during bariatric surgery (liver disease) or gallbladder surgery (healthy). Histological assessment of MASLD and fibrosis was conducted by a specialist pathologist according to the NASH CRN. Log-normalized TAGLN expression values were exported from the Seurat object and plotted in GraphPad Prism to compare expression between healthy and fibrotic liver samples*(52).*

High-definition spatial transcriptomics data generation and analysis

5µm patient biopsy FFPE sections were prepared and stained according to the Visium HD FFPE Tissue Preparation protocols (GC000684, Revision D) and processed following the standard Visium HD Spatial Gene Expression Reagents Kits User Guide (GC000685, Revision D) as outlined by 10X Genomics. Prepared libraries were sequenced on the Illumina NovaSeq6000 SP lane with 100 cycles. The sequencing data was pre-processed using Space Ranger (v3.1.3, 10X Genomics), including the generation of the CLOUPE output at 8µmx8µm binning resolution. On a per-biopsy basis, endothelial bins were identified within Loupe Browser (v.8.0.0, 10X Genomics) by a Log2 Feature Sum of endothelial marker genes (PECAM1, THBD, EMCN, MCAM, LYVE1) and split into TAGLN+ and TAGLN-sub-populations (endo_TAGLN+ and endo_TAGLN-). Non-endothelial bins were defined as “rest_of_tissue”. Differential gene

expression between TAGLN+ and TAGLN-endothelium-associated bins was identified within Loupe Browser using the 'Run Differential Expression' tool, with significance adjusted for multiple comparisons using Benjamini-Hochberg correction. To define potential receptor-ligand analyses, barcode annotations were exported from Loupe Browser. The Visium HD raw counts data matrix was loaded into a SpatialData object, subset to the relevant barcodes prior to annotation transfer, normalization and log transformation of the expression data and h5ad conversion. CellChat (v2.2.0) was used to generate a spatial CellChat object for each biopsy using the full human interaction database (CellChatDB.human) and to generate visualizations. Compute CommunProb was performed using parameters "type = truncated Mean" and "trim=0.01".

SteatoSITE ethics statement

Anonymised tissue was supplied with approval from the National Health Service Research Scotland (NRS) Biorepository Network (Ref: SR1032; 2 Aug 2018). Unified transparent approval for unconsented data inclusion was granted by the West of Scotland Research Ethics Committee 4 (Ref: 25/WS/0004; 16 Jan 2025), the Public Benefit and Privacy Panel for Health and Social Care (Ref: 2425-0221; 20 May 2025), institutional R&D departments, and Caldicott Guardians.

SteatoSITE analysis

SteatoSITE v2 with follow up to end December 2023 was used. Only biopsy cases with RNAseq were used for analysis using R4.4.1. For overall survival analysis (n=505), surv_cutpoint() from the 'survminer' (0.5.1) with a minimum proportion 0.25 was used to define the optimal cutpoint of TAGLN expression (normalised cpm) using the maximally selected rank statistics from the 'maxstat' (0.7-26) package. For time-to-liver related event analysis, cases with no liver event coding prior to the biopsy date (n=435) were fit with a time-to-event proportional hazards regression model using NASH-CRN fibrosis stage, age, gender, and TAGLN expression as

predictive variables using coxph from 'survival' (3.6-4), and the cumulative incidence function of the model at t=4000 days stratified for NASH-CRN fibrosis stage plotted as a function of TAGLN expression using a plot_surv_at_t from 'contsurvplot' (0.2.2).

Statistical Analysis

All statistical analysis was completed using the GraphPad Prism 10.0.0 software (GraphPad, US). Statistical tests were performed to identify statistically significant results. The Shapiro-Wilk normality test was first used to identify whether the data are distributed normally. After that, a t-test or a one-way analysis of variance (ANOVA) was used for normally distributed data and the equivalent non-parametric test was used for data that did not pass the normality test. Appropriate multiple comparisons tests were also applied to all datasets. The exact test used for each dataset is detailed in the corresponding figure legend of each Figure. Statistical significance was defined as $P \leq 0.05$ (ns: $p > 0.05$, *: $p \leq 0.05$, **: ≤ 0.01, ***: ≤ 0.001, ****: ≤ 0.0001). All data are presented as mean ± standard error of the mean (SEM).

List of Supplementary Materials

Supplementary Figures 1-7

Supplementary Tables 1-3

References

1. Z. M. Younossi, P. Golabi, J. M. Paik, A. Henry, C. Van Dongen, L. Henry, The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. *Hepatology* **77** (2023), doi:10.1097/HEP.0000000000000004.

2. A. J. Sanyal, P. N. Newsome, I. Kliers, L. H. Østergaard, M. T. Long, M. S. Kjær, A. M. G. Cali, E. Bugianesi, M. E. Rinella, M. Roden, V. Ratziu, Phase 3 Trial of Semaglutide in Metabolic Dysfunction–Associated Steatohepatitis. *New England Journal of Medicine* **392** (2025), doi:10.1056/nejmoa2413258.

3. S. A. Harrison, P. Bedossa, C. D. Guy, J. M. Schattenberg, R. Loomba, R. Taub, D. Labriola, S. E. Moussa, G. W. Neff, M. E. Rinella, Q. M. Anstee, M. F. Abdelmalek, Z. Younossi, S. J. Baum, S. Francque, M. R. Charlton, P. N. Newsome, N. Lanthier, I. Schiefke, A. Mangia, J. M. Pericàs, R. Patil, A. J. Sanyal, M. Noureddin, M. B. Bansal, N. Alkhouri, L. Castera, M. Rudraraju, V. Ratziu, A Phase 3, Randomized, Controlled Trial of Resmetirom in NASH with Liver Fibrosis. *New England Journal of Medicine* **390** (2024), doi:10.1056/nejmoa2309000.

4. K. B. Halpern, R. Shenhav, H. Massalha, B. Toth, A. Egozi, E. E. Massasa, C. Medgalia, E. David, A. Giladi, A. E. Moor, Z. Porat, I. Amit, S. Itzkovitz, Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. *Nat Biotechnol* **36**, 962–970 (2018).

5. A. Gola, M. G. Dorrington, E. Speranza, C. Sala, R. M. Shih, A. J. Radtke, H. S. Wong, A. P. Baptista, J. M. Hernandez, G. Castellani, I. D. C. Fraser, R. N. Germain, Commensal-driven immune zonation of the liver promotes host defence. *Nature* **589**, 131–136 (2021).

6. D. Inverso, J. Shi, K. H. Lee, M. Jakab, S. Ben-Moshe, S. R. Kulkarni, M. Schneider, G. Wang, M. Komeili, P. A. Velez, M. Riedel, C. Spegg, T. Ruppert, C. Schaeffer-Reiss, D. Helm, I. Singh, M. Boutros, S. Chintharlapalli, M. Heikenwalder, S. Itzkovitz, H. G. Augustin, A spatial vascular transcriptomic, proteomic, and phosphoproteomic atlas unveils an angiocrine Tie-Wnt signaling axis in the liver. *Dev Cell* **56**, 1677-1693 e10 (2021).

7. B. Ruan, J. L. Duan, H. Xu, K. S. Tao, H. Han, G. R. Dou, L. Wang, Capillarized Liver Sinusoidal Endothelial Cells Undergo Partial Endothelial-Mesenchymal Transition to Actively Deposit Sinusoidal ECM in Liver Fibrosis. *Front Cell Dev Biol* **9**, 671081 (2021).

8. M. Winkler, T. Staniczek, S. W. Kürschner, C. D. Schmid, H. Schönhaber, J. Cordero, L. Kessler, A. Mathes, C. Sticht, M. Neßling, A. Uvarovskii, S. Anders, X. jun Zhang, G. von Figura, D. Hartmann, C. Mogler, G. Dobreva, K. Schledzewski, C. Géraud, P. S. Koch, S. Goerdt, Endothelial GATA4 controls liver fibrosis and regeneration by preventing a pathogenic switch in angiocrine signaling. *J. Hepatol.* **74** (2021), doi:10.1016/j.jhep.2020.08.033.

9. F. Schaffner, H. Popper, Capillarization of Hepatic Sinusoids in Man. *Gastroenterology* **44** (1963), doi:10.1016/S0016-5085(63)80130-4.

10. P. Ramachandran, R. Dobie, J. R. Wilson-Kanamori, E. F. Dora, B. E. P. Henderson, N. T. Luu, J. R. Portman, K. P. Matchett, M. Brice, J. A. Marwick, R. S. Taylor, M. Efremova, R. Vento-Tormo, N. O. Carragher, T. J. Kendall, J. A. Fallowfield, E. M. Harrison, D. J. Mole, S. J. Wigmore, P. N. Newsome, C. J. Weston, J. P. Iredale, F. Tacke, J. W. Pollard, C. P. Ponting, J. C. Marioni, S. A. Teichmann, N. C. Henderson, Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**, 512–518 (2019).

11. N. P. Dufton, C. R. Peghaire, L. Osuna-Almagro, C. Raimondi, V. Kalna, A. Chauhan, G. Webb, Y. Yang, G. M. Birdsey, P. Lalor, J. C. Mason, D. H. Adams, A. M. Randi, Dynamic regulation of canonical TGFbeta signalling by endothelial transcription factor ERG protects from liver fibrogenesis. *Nat Commun* **8**, 895 (2017).

12. J. Ribera, M. Pauta, P. Melgar-Lesmes, B. Cordoba, A. Bosch, M. Calvo, D. Rodrigo-Torres, P. Sancho-Bru, A. Mira, W. Jimenez, M. Morales-Ruiz, A small population of liver endothelial cells undergoes endothelial-to-mesenchymal transition in response to chronic liver injury. *Am J Physiol Gastrointest Liver Physiol* **313**, G492–G504 (2017).

13. I. Coelho, N. Duarte, A. Barros, M. P. Macedo, C. Penha-Goncalves, Trem-2 Promotes Emergence of Restorative Macrophages and Endothelial Cells During Recovery From Hepatic Tissue Damage. *Front Immunol* **11**, 616044 (2020).

14. J. Ma, G. van der Zon, M. A. F. V. Gonçalves, M. van Dinther, M. Thorikay, G. Sanchez-Duffhues, P. ten Dijke, TGF-β-Induced Endothelial to Mesenchymal Transition Is Determined by a Balance Between SNAIL and ID Factors. *Front. Cell Dev. Biol.* **9** (2021), doi:10.3389/fcell.2021.616610.

15. Q. Zhou, X. Zhang, X. Tong, C. Tang, X. Chen, L. Peng, X. Xia, L. Zhang, Single Cell Meta-Analysis of Endothelial to Mesenchymal Transition (EndMT) in Glucose Metabolism of the Digestive Diseases. *Front. Mol. Biosci.* **9** (2022), doi:10.3389/fmolb.2022.866408.

16. S. Piera-Velazquez, S. A. Jimenez, Endothelial to Mesenchymal Transition: Role in Physiology and in the Pathogenesis of Human Diseases. *Physiol Rev* **99**, 1281–1324 (2019).

17. E. M. Zeisberg, O. Tarnavski, M. Zeisberg, A. L. Dorfman, J. R. McMullen, E. Gustafsson, A. Chandraker, X. Yuan, W. T. Pu, A. B. Roberts, E. G. Neilson, M. H. Sayegh,

S. Izumo, R. Kalluri, Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. *Nat Med* **13**, 952–961 (2007).

18. Z. Li, B. Chen, W. Dong, M. Kong, Z. Fan, L. Yu, D. Wu, J. Lu, Y. Xu, MKL1 promotes endothelial-to-mesenchymal transition and liver fibrosis by activating TWIST1 transcription. *Cell Death Dis.* **10**, 899 (2019).

19. M. K. Terkelsen, S. M. Bendixen, D. Hansen, E. A. H. Scott, A. F. Moeller, R. Nielsen, S. Mandrup, A. Schlosser, T. L. Andersen, G. L. Sorensen, A. Krag, K. N. Natarajan, S. Detlefsen, H. Dimke, K. Ravnskjaer, Transcriptional Dynamics of Hepatic Sinusoid-Associated Cells After Liver Injury. *Hepatology* **72** (2020), doi:10.1002/hep.31215.

20. C. Gkantsinikoudi, M. Terranova-Barberio, N. P. Dufton, Addressing complex autofluorescence signatures in solid tissue samples to enhance full spectrum flow cytometry of non-immune cells (2025), doi:10.64898/2025.12.19.695385.

21. T. J. Kendall, M. Jimenez-Ramos, F. Turner, P. Ramachandran, J. Minnier, M. D. McColgan, M. Alam, H. Ellis, D. R. Dunbar, G. Kohnen, P. Konanahalli, K. A. Oien, L. Bandiera, F. Menolascina, A. Juncker-Jensen, D. Alexander, C. Mayor, I. N. Guha, J. A. Fallowfield, An integrated gene-to-outcome multimodal database for metabolic dysfunction-associated steatotic liver disease. *Nat. Med.* **29** (2023), doi:10.1038/s41591-023-02602-2.

22. L. M. Becker, S. H. Chen, J. Rodor, L. P. M. H. de Rooij, A. H. Baker, P. Carmeliet, Deciphering endothelial heterogeneity in health and disease at single-cell resolution: progress and perspectives*Cardiovasc. Res.* **119** (2023), doi:10.1093/cvr/cvac018.

23. L. Zhao, Y. Jin, K. Donahue, M. Tsui, M. Fish, C. Y. Logan, B. Wang, R. Nusse, Tissue Repair in the Mouse Liver Following Acute Carbon Tetrachloride Depends on Injury-Induced Wnt/β-Catenin Signaling. *Hepatology* **69** (2019), doi:10.1002/hep.30563.

24. A. Miura, T. Hosono, T. Seki, Macrophage potentiates the recovery of liver zonation and metabolic function after acute liver injury. *Sci. Rep.* **11** (2021), doi:10.1038/s41598-021-88989-9.

25. Y. Gao, Z. S. Galis, Exploring the Role of Endothelial Cell Resilience in Cardiovascular Health and Disease*Arterioscler. Thromb. Vasc. Biol.* **41** (2021), doi:10.1161/ATVBAHA.120.314346.

26. L. S. Tombor, D. John, S. F. Glaser, G. Luxan, E. Forte, M. Furtado, N. Rosenthal, N. Baumgarten, M. H. Schulz, J. Wittig, E. M. Rogg, Y. Manavski, A. Fischer, M. Muhly-Reinholz, K. Klee, M. Looso, C. Selignow, T. Acker, S. I. Bibli, I. Fleming, R. Patrick, R. P. Harvey, W. T. Abplanalp, S. Dimmeler, Single cell sequencing reveals endothelial plasticity with transient mesenchymal activation after myocardial infarction. *Nat Commun* **12**, 681 (2021).

27. L. J. Greenspan, B. M. Weinstein, To be or not to be: endothelial cell plasticity in development, repair, and disease*Angiogenesis* **24** (2021), doi:10.1007/s10456-020-09761-7.

28. H. Zhang, K. O. Lui, B. Zhou, Endocardial cell plasticity in cardiac development, diseases and regeneration*Circ. Res.* **122** (2018), doi:10.1161/CIRCRESAHA.117.312136.

29. A. Kitao, Y. Sato, S. Sawada-Kitamura, K. Harada, M. Sasaki, H. Morikawa, S. Shiomi, M. Honda, O. Matsui, Y. Nakanuma, Endothelial to mesenchymal transition via transforming growth factor-beta1/Smad activation is associated with portal venous stenosis in idiopathic portal hypertension. *Am J Pathol* **175**, 616–626 (2009).

30. Z. Li, B. Chen, W. Dong, M. Kong, Z. Fan, L. Yu, D. Wu, J. Lu, Y. Xu, MKL1 promotes endothelial-to-mesenchymal transition and liver fibrosis by activating TWIST1 transcription. *Cell Death Dis* **10**, 899 (2019).

31. O. Strauss, A. Phillips, K. Ruggiero, A. Bartlett, P. R. Dunbar, Immunofluorescence identifies distinct subsets of endothelial cells in the human liver. *Sci. Rep.* **7** (2017), doi:10.1038/srep44356.

32. S. Hu, S. Liu, Y. Bian, M. Poddar, S. Singh, C. Cao, J. McGaughey, A. Bell, L. L. Blazer, J. J. Adams, S. S. Sidhu, S. Angers, S. P. Monga, Single-cell spatial transcriptomics reveals a dynamic control of metabolic zonation and liver regeneration by endothelial cell Wnt2 and Wnt9b. *Cell Rep. Med.* **3** (2022), doi:10.1016/j.xcrm.2022.100754.

33. B. C. Cooley, J. Nevado, J. Mellad, D. Yang, C. St Hilaire, A. Negro, F. Fang, G. Chen, H. San, A. D. Walts, R. L. Schwartzbeck, B. Taylor, J. D. Lanzer, A. Wragg, A. Elagha, L. E. Beltran, C. Berry, R. Feil, R. Virmani, E. Ladich, J. C. Kovacic, M. Boehm, TGF-beta signaling mediates endothelial-to-mesenchymal transition (EndMT) during vein graft remodeling. *Sci Transl Med* **6**, 227ra34 (2014).

34. M. Setty, M. D. Tadmor, S. Reich-Zeliger, O. Angel, T. M. Salame, P. Kathail, K. Choi, S. Bendall, N. Friedman, D. Pe'Er, Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat. Biotechnol.* **34** (2016), doi:10.1038/nbt.3569.

35. W. Sen Lee, M. K. Jain, B. M. Arkonac, D. Zhang, S. Y. Shaw, S. Kashiki, K. Maemura, S. L. Lee, N. K. Hollenberg, M. E. Lee, E. Haber, Thy-1, a novel marker for angiogenesis upregulated by inflammatory cytokines. *Circ. Res.* **82** (1998), doi:10.1161/01.RES.82.8.845.

36. T. A. Rege, J. S. Hagood, Thy-1 as a regulator of cell-cell and cell-matrix interactions in axon regeneration, apoptosis, adhesion, migration, cancer, and fibrosis. *The FASEB Journal* **20** (2006), doi:10.1096/fj.05-5460rev.

37. K. Yuki, The role of ICAM-1 isoforms in neutrophil recruitment to liver. *The Journal of Immunology* **196** (2016), doi:10.4049/jimmunol.196.supp.119.7.

38. A. Zahr, P. Alcaide, J. Yang, A. Jones, M. Gregory, N. G. Dela Paz, S. Patel-Hett, T. Nevers, A. Koirala, F. W. Luscinskas, M. Saint-Geniez, B. Ksander, P. A. D'Amore, P. Argüeso, Endomucin prevents leukocyte-endothelial cell adhesion and has a critical role under resting and inflammatory conditions. *Nat. Commun.* **7** (2016), doi:10.1038/ncomms10363.

39. A. S. Jauch, S. A. Wohlfeil, C. Weller, B. Dietsch, V. Häfele, A. Stojanovic, M. Kittel, H. Nolte, A. Cerwenka, M. Neumaier, K. Schledzewski, C. Sticht, P. S. Reiners-Koch, S. Goerdt, C. Géraud, Lyve-1 deficiency enhances the hepatic immune microenvironment entailing altered susceptibility to melanoma liver metastasis. *Cancer Cell Int.* **22** (2022), doi:10.1186/s12935-022-02800-x.

40. K. Tsuji-Tamura, S. Morino-Koga, S. Suzuki, M. Ogawa, The canonical smooth muscle cell marker TAGLN is present in endothelial cells and is involved in angiogenesis. *J. Cell Sci.* **134** (2021), doi:10.1242/JCS.254920.

41. Q. Huang, Y. Gan, Z. Yu, H. Wu, Z. Zhong, Endothelial to Mesenchymal Transition: An Insight in Atherosclerosis*Front. Cardiovasc. Med.* **8** (2021), doi:10.3389/fcvm.2021.734550.

42. Z. Wang, Q. Xu, N. Zhang, X. Du, G. Xu, X. Yan, CD146, from a melanoma cell adhesion molecule to a signaling receptor*Signal Transduct. Target. Ther.* **5** (2020), doi:10.1038/s41392-020-00259-8.

43. M. Gromak, H. Luxenburger, T. Böttler, R. Thimme, M. Hofmann, N. Roehlen, in *39. Jahrestagung der Deutschen Arbeitsgemeinschaft zum Studium der Leber*, (2023), vol. 61.

44. J. Stalin, W. Traboulsi, L. Vivancos-Stalin, M. Nollet, A. Joshkon, R. Bachelier, B. Guillet, R. Lacroix, A. Foucault-Bertaud, A. S. Leroyer, F. Dignat-George, N. Bardin, M. Blot-Chabaud, Therapeutic targeting of soluble CD146/MCAM with the M2J-1 monoclonal

antibody prevents metastasis development and procoagulant activity in CD146-positive invasive tumors. *Int. J. Cancer* **147** (2020), doi:10.1002/ijc.32909.

45. Z. Shen, B. Shen, W. Dai, C. Zhou, X. Luo, Y. Guo, J. Wang, X. Xu, Z. Sun, X. Cai, H. Dong, L. Lu, Expansion of macrophage and liver sinusoidal endothelial cell subpopulations during non-alcoholic steatohepatitis progression. *iScience* **26** (2023), doi:10.1016/j.isci.2023.106572.

46. R. I. Velliou, A. I. Legaki, P. Nikolakopoulou, N. I. Vlachogiannis, A. Chatzigeorgiou, Liver endothelial cells in NAFLD and transition to NASH and HCC*Cellular and Molecular Life Sciences* **80** (2023), doi:10.1007/s00018-023-04966-7.

47. G. Landini, G. Martinelli, F. Piccinini, Colour deconvolution: Stain unmixing in histological imaging*Bioinformatics* **37** (2021), doi:10.1093/bioinformatics/btaa847.

48. G. Monaco, H. Chen, M. Poidinger, J. Chen, J. P. De Magalhães, A. Larbi, FlowAI: Automatic and interactive anomaly discerning tools for flow cytometry data. *Bioinformatics* **32** (2016), doi:10.1093/bioinformatics/btw191.

49. S. Van Gassen, B. Callebaut, M. J. Van Helden, B. N. Lambrecht, P. Demeester, T. Dhaene, Y. Saeys, FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry Part A* **87** (2015), doi:10.1002/cyto.a.22625.

50. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3** (2018), doi:10.21105/joss.00861.

51. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis*Nat. Methods* **9** (2012), doi:10.1038/nmeth.2089.

52. R. Kumar, W. Li, J. Loy, K. Devalia, A. Goralcyzk, H. Malik, L. James, F. Marelli-Berg, P. Ramachandran, J. Boot, P. Stevens, E. Wozniak, C. Mein, R. D. Goldin, W. Alazawi, TOP-

441 CD8+ T cells are pro-inflammatory, cytotoxic, and clonally expanded in the early stages of fibrosis due to metabolic dysfunction-associated steatohepatitis (MASH). *J. Hepatol.* **82** (2025), doi:10.1016/s0168-8278(25)01569-7.

Acknowledgements:

This research was supported by Dr Paul Imbert, manager of the CMR Advanced Bioimaging Facility (CMR-ABF), and Manuela Terranova-Barberio, manager of the WHRI Flow Cytometry Facility which are both located at the William Harvey Research Institute, Charterhouse Square, Queen Mary University of London, EC1M 6BQ, United Kingdom.

Lindsay Birchall is thanked for her technical assistance in spatial transcriptomics sample preparation. The University of Manchester Genomics Technology Core Facility is thanked for their assistance with the Visium HD spatial transcriptomics protocol.

Funding:

Medical Research Council project grant MR/Y013751/1 (ND)

Bart's Charity Seed Grant MGU0514 (ND)

British Heart Foundation PhD studentship FS/4yPhD/F/21/34161 (CG)

Medical Research Council project grant MR/T031883/1 (WA)

Innovate UK project grant TS/U00457X/1 (KPH)

Innovate UK project grant TS/R017581/1 (JAF/TJK)

Guts UK DGO2019_16 (JAF/TJK)

Author contributions:

Conceptualization: CG, JPD, AW, NPD

Methodology: CG, RK, EJ, SL, TJK, NPD

Investigation: CG, JPD, RK, EJ, TJK, WL, MS, NPD

Visualization: CG, RK, EJ, TJK, NPD

Funding acquisition: KPH, AR, WA, NPD

Project administration: AR, JAF, SL, VSL, KPH, AW, NPD

Supervision: AR, JAF, KPH, AW, NPD

Writing – original draft: CG, JPD, NPD

Writing – review & editing: CG, JPD, RK, EJ, TJK, WL, MS, AR, JAF, KPH, AW, NPD

Competing interests: Authors declare that they have no competing interests. JAF would like to disclose that he serves as a consultant and/or advisory board member for Medixci, Resolution Therapeutics, Kynos Therapeutics, Gyre Therapeutics, Ipsen, River 2 Renal Corp., Stimuliver, Guidepoint and ICON plc, has received speakers' fees from HistoIndex, Resolution Therapeutics and Société internationale de développement professionnel continu Cléo, and research grant funding from GlaxoSmithKline and Genentech

Data and materials availability: Murine scRNAseq data were analyzed from 10x Genomics single-cell transcriptomic data of isolated non-parenchymal liver cells from healthy and diseased (2w and 4w CCl4 treated) mice deposited on the GEO database (GSE145086).

For access to the full SteatoSITE resource, including rich clinical outcome annotation please visit **https://regeneration-repair.ed.ac.uk/research/research-resources/steatosite**. Raw RNA-seq data (FASTQ files) and minimal metadata (fibrosis stage) is available in the European Nucleotide Archive (https://www.ebi.ac.uk/ena; study accession number: PRJEB58625).

Data from the Citeseq (Fig. S6) and Visium HD Spatial Transcriptomics analyses (Fig. 8 and Fig. S7) will be made available on the GEO databases upon acceptance.

Fig. 1

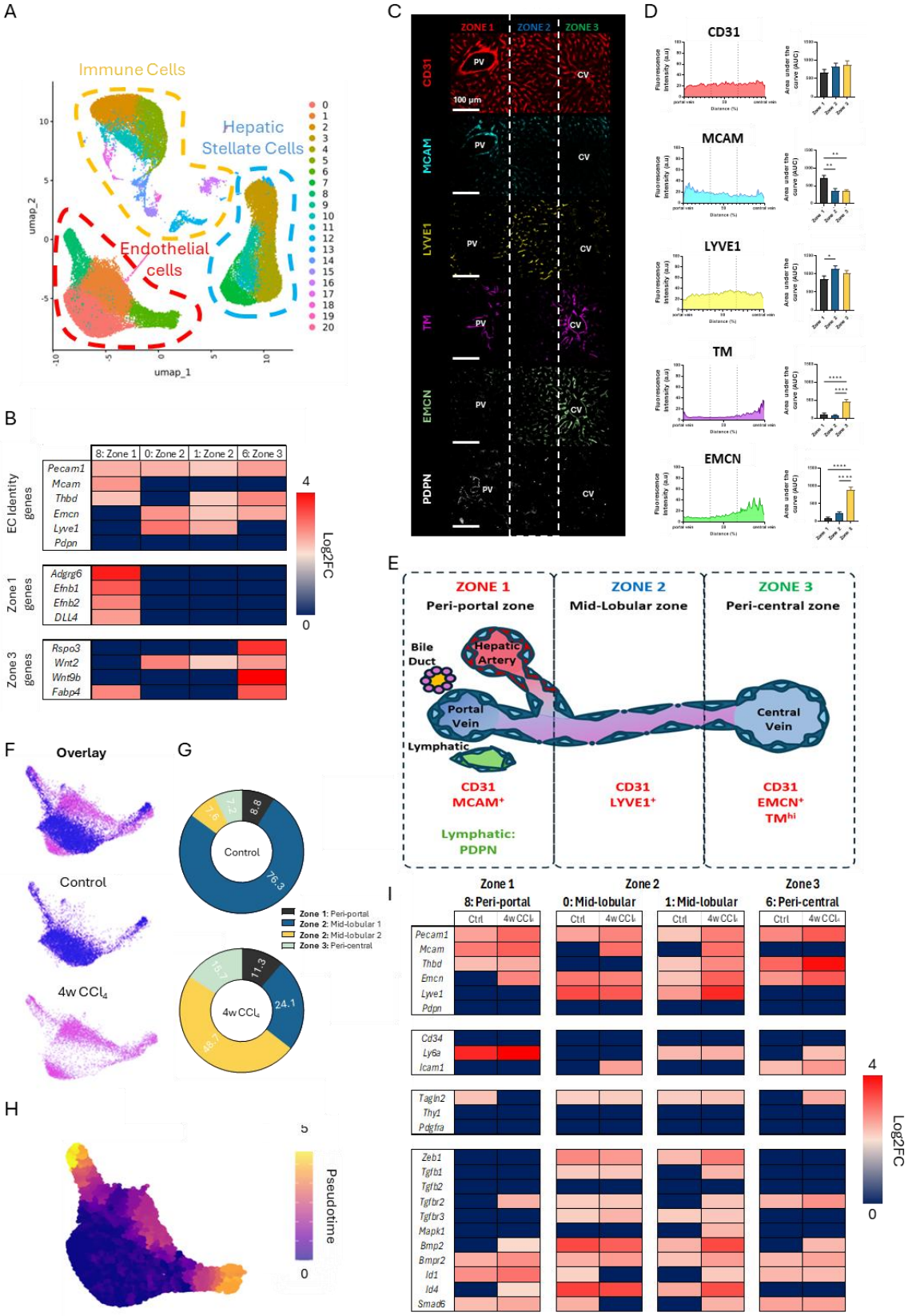

A
Immune Cells
Hepatic Stellate Cells
Endothelial cells
umap_1
umap_2
B
8: Zone 1
0: Zone 2
1: Zone 2
6: Zone 3
EC Identity genes
Pecam1
Mcam
Thbd
Emcn
Lyve1
Pdpn
Zone 1 genes
Adgrg6
Efnb1
Efnb2
DLL4
Zone 3 genes
Rspo3
Wnt2
Wnt9b
Fabp4
Log2FC
C
ZONE 1
ZONE 2
ZONE 3
CD31
MCAM
LYVE1
TM
EMCN
PDPN
PV
CV
100 µm
D
CD31
MCAM
LYVE1
TM
EMCN
E
ZONE 1
Peri-portal zone
ZONE 2
Mid-Lobular zone
ZONE 3
Peri-central zone
Bile Duct
Hepatic Artery
Portal Vein
Central Vein
Lymphatic
CD31 MCAM+
CD31 LYVE1+
CD31 EMCN+ TMhi
Lymphatic: PDPN
F
Overlay
Control
4w CCl4
G
Control
4w CCl4
Zone 1: Peri-portal
Zone 2: Mid-lobular 1
Zone 2: Mid-lobular 2
Zone 3: Peri-central
H
Pseudotime
I
Zone 1
8: Peri-portal
Zone 2
0: Mid-lobular
1: Mid-lobular
Zone 3
6: Peri-central
Ctrl
4w CCl4
Pecam1
Mcam
Thbd
Emcn
Lyve1
Pdpn
Cd34
Ly6a
Icam1
Tagln2
Thy1
Pdgfra
Zeb1
Tgfb1
Tgfb2
Tgfbr2
Tgfbr3
Mapk1
Bmp2
Bmpr2
Id1
Id4
Smad6
Log2FC

Fig. 1. **Characterisation of EC identity and plasticity markers in health and fibrogenesis.** (A) UMAP of non-parenchymal cell scRNAseq data (GSE145086) for control (untreated), 2- and 4-week CCl_4 treated male C57BL/6 mice*(19)*. (B) Heatmap of gene expression for the four EC clusters (clusters 0, 1, 6 and 8) under baseline conditions identifying the zonation of EC identity markers (top panel) based on expression of established markers for peri-portal /Zone 1 (mid-panel) and peri-central/Zone 3 (bottom panel) EC. (C) Representative confocal images of healthy liver tissue stained for the EC identity markers CD31, melanoma cell adhesion molecule (MCAM), LYVE1, thrombomodulin (TM) and endomucin (EMCN). Podoplanin (PDPN) included as lymphatic marker. (D) Immunofluorescence intensity distribution across the portal-central vein axis and bar graphs of signal quantification for each zone for CD31, MCAM, LYVE1, TM and EMCN. Scale bars= 100µm. (E) Schematic summarising the zonal distribution of hepatic EC markers. (F) Visual overlay of EC clusters from scRNAseq for control mice (blue) and 4w CCl_4 treatment (magenta). (G) Quantification of the abundance of each EC cluster in control (top panel) and 4w CCl_4 treated mice (bottom panel). (H) Pseudotime analysis of progressive fibrogenesis from healthy, 2w and 4w CCl_4-treated samples showing the transition away from cluster 0 mid-lobular EC toward clusters 1, 6 and 8. (I) Heatmap of gene expression for EC markers (top panels) and TGFβ-activated genes (bottom panels) for the four EC clusters comparing control and 4w CCl_4 treatment groups.

Fig. 2

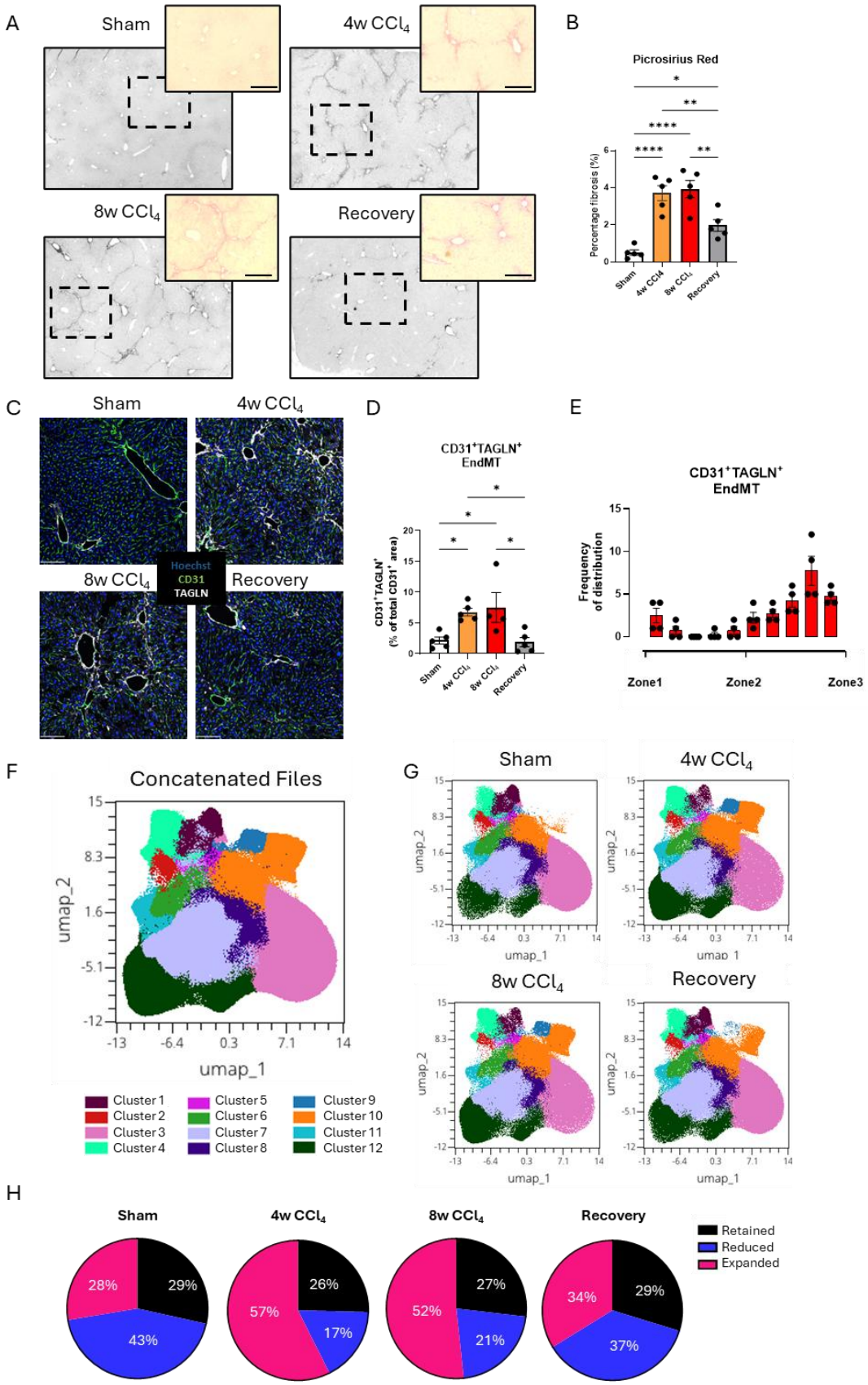

A
Sham
4w CCl4
8w CCl4
Recovery
B
Picrosirius Red
Percentage fibrosis (%)
Sham
4w CCl4
8w CCl4
Recovery
C
Sham
4w CCl4
8w CCl4
Recovery
Hoechst
CD31
TAGLN
D
CD31+TAGLN+ EndMT
CD31+TAGLN+ (% of total CD31+ area)
E
CD31+TAGLN+ EndMT
Frequency of distribution
Zone1
Zone2
Zone3
F
Concatenated Files
umap_2
umap_1
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Cluster 11
Cluster 12
G
Sham
4w CCl4
8w CCl4
Recovery
umap_2
umap_1
H
Sham
4w CCl4
8w CCl4
Recovery
Retained
Reduced
Expanded
28%
29%
43%
26%
57%
17%
27%
52%
21%
34%
29%
37%

Fig. 2. Mapping EC proteomic changes in response to CCl_4-induced liver fibrosis and tissue recovery. (A) Representative images of picrosirius red staining of liver sections from Sham, 4- and 8-week CCl_4 treated and recovery groups. Scale bar= 250µm. (B) Quantification of picrosirius red staining, expressed as a percentage (%) of total tissue area. (C) Representative confocal images of IF staining for Hoechst (blue), CD31 (green) and TAGLN (white) in Sham, CCl_4-treated and recovery groups. Scale bar= 100µm. (D) Quantification of $CD31^+TAGLN^+$ EndMT cells expressed as a percentage of the total $CD31^+$ area. (E) Bar graph of EndMT cell distribution across the portal vein to central vein axis. (F, G) UMAP of $CD45^-CD31^+$ cells from (F) control, CCl_4-treated and recovery group samples displaying 12 distinct endothelial cell clusters and (G) $CD45^-CD31^+$ cells split by treatment group. (H) Pie charts of the percentage composition of total EC characterised as 'Retained' (black), 'Reduced' (blue) and 'Expanded' (pink) for each treatment group. For statistical analysis, one-way ANOVA tests with Tukey's multiple comparisons test were used. N=5 animals/group. *; $p<0.05$, **; $p<0.01$, ****; $p<0.0001$.

Fig. 3

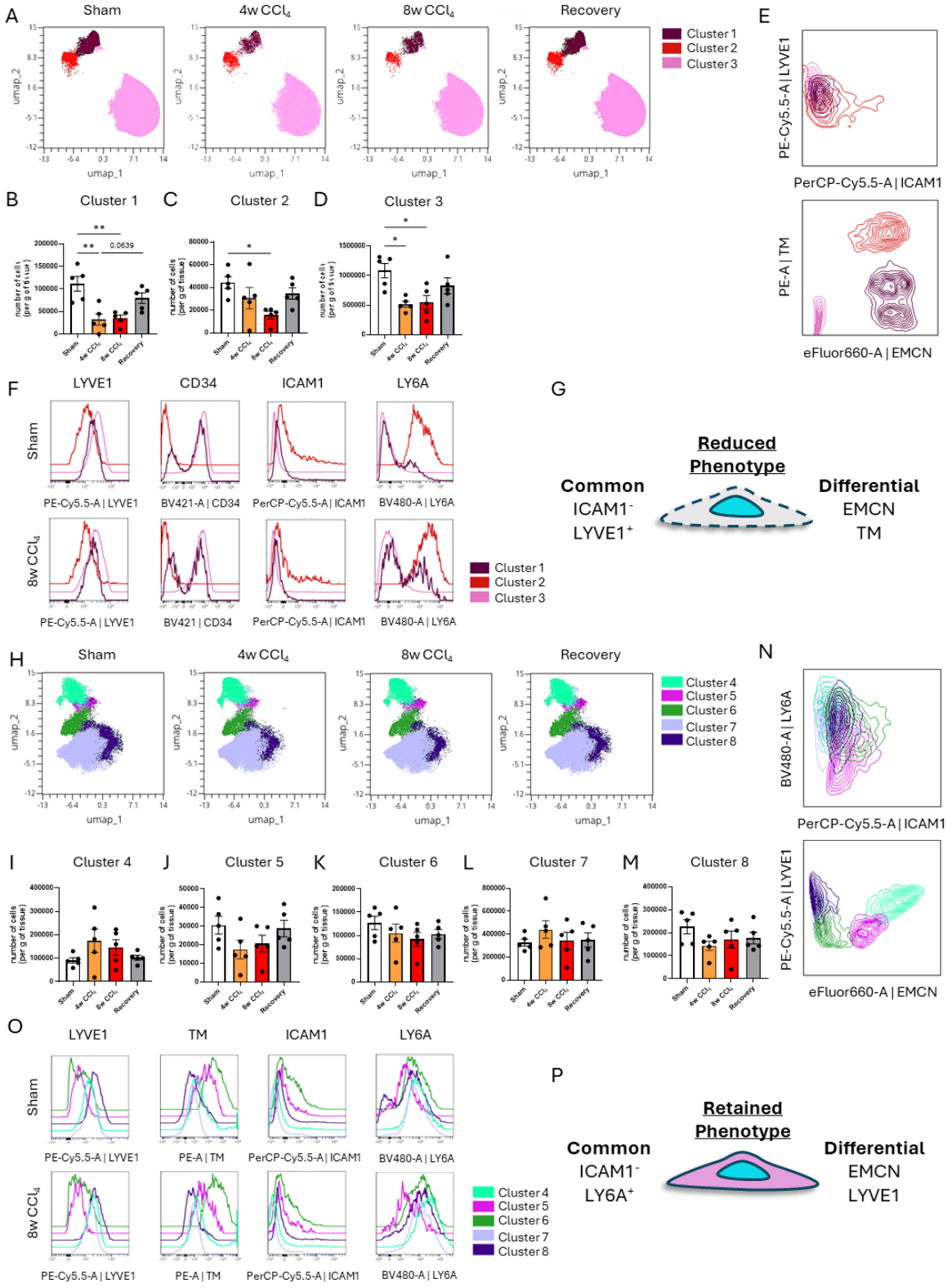

A
Sham
4w CCl4
8w CCl4
Recovery
Cluster 1
Cluster 2
Cluster 3
umap_1
umap_2
B
Cluster 1
C
Cluster 2
D
Cluster 3
number of cells (per g of tissue)
Sham
4w CCl4
8w CCl4
Recovery
E
PE-Cy5.5-A | LYVE1
PerCP-Cy5.5-A | ICAM1
PE-A | TM
eFluor660-A | EMCN
F
LYVE1
CD34
ICAM1
LY6A
PE-Cy5.5-A | LYVE1
BV421-A | CD34
PerCP-Cy5.5-A | ICAM1
BV480-A | LY6A
G
Reduced Phenotype
Common
ICAM1⁻
LYVE1⁺
Differential
EMCN
TM
H
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
N
BV480-A | LY6A
I
Cluster 4
J
Cluster 5
K
Cluster 6
L
Cluster 7
M
Cluster 8
O
LYVE1
TM
ICAM1
LY6A
PE-A | TM
P
Retained Phenotype
Common
ICAM1⁻
LY6A⁺
Differential
EMCN
LYVE1

Fig. 3. Phenotyping the characteristics of Reduced and Retained EC subpopulations by FSFC. (A) UMAP visualisation of three Reduced EC clusters; 1 (purple), 2 (red) and 3 (pink). (B-D) Quantification of the number of cells in each cluster expressed as number of cells per gram of liver tissue for (B) cluster 1, (C) cluster 2 and (D) cluster 3 per treatment group. (E) Contour plots of clusters 1, 2 and 3 for ICAM1 vs LYVE1 (top panel) and endomucin (EMCN) vs thrombomodulin (TM) (bottom panel). (F) Histograms of clusters 1, 2 and 3 comparing the median fluorescence intensity (MFI) of LYVE1, CD34, ICAM1 and LY6A in Sham (top panels) and 8w CCl_4-treated animals (bottom panels). (G) Schematic summarising the common and differential traits of the three Reduced EC clusters. (H) UMAP visualisation of five Retained EC clusters; 4 (turquoise), 5 (light purple), 6 (green), 7 (violet) and 8 (dark purple). (I-M) Quantification of the number of cells in each cluster expressed as number of cells per gram of liver tissue for (I) cluster 4, (J) cluster 5, (K) cluster 6, (L) cluster 7 and (M) cluster 8 per treatment group. (N) Contour plots of clusters 4, 5, 6, 7 and 8 for ICAM1 vs LY6A (top panel) and EMCN vs LYVE1 (bottom panel). (F) Histograms of clusters 4, 5, 6, 7 and 8 comparing the MFI of LYVE1, TM, ICAM1 and LY6A in Sham (top panels) and 8w CCl_4-tretaed animals (bottom panels). (G) Schematic summarising the common and differential traits of Retained EC clusters. For statistical analysis, one-way ANOVAs with Tukey's multiple comparisons tests were used for figures B, D and I-M and a Kruskal-Wallis test with Dunn's multiple comparisons test was used for figure C. N=5 animals/group. *; $p<0.05$, **; $p<0.01$.

Fig. 4

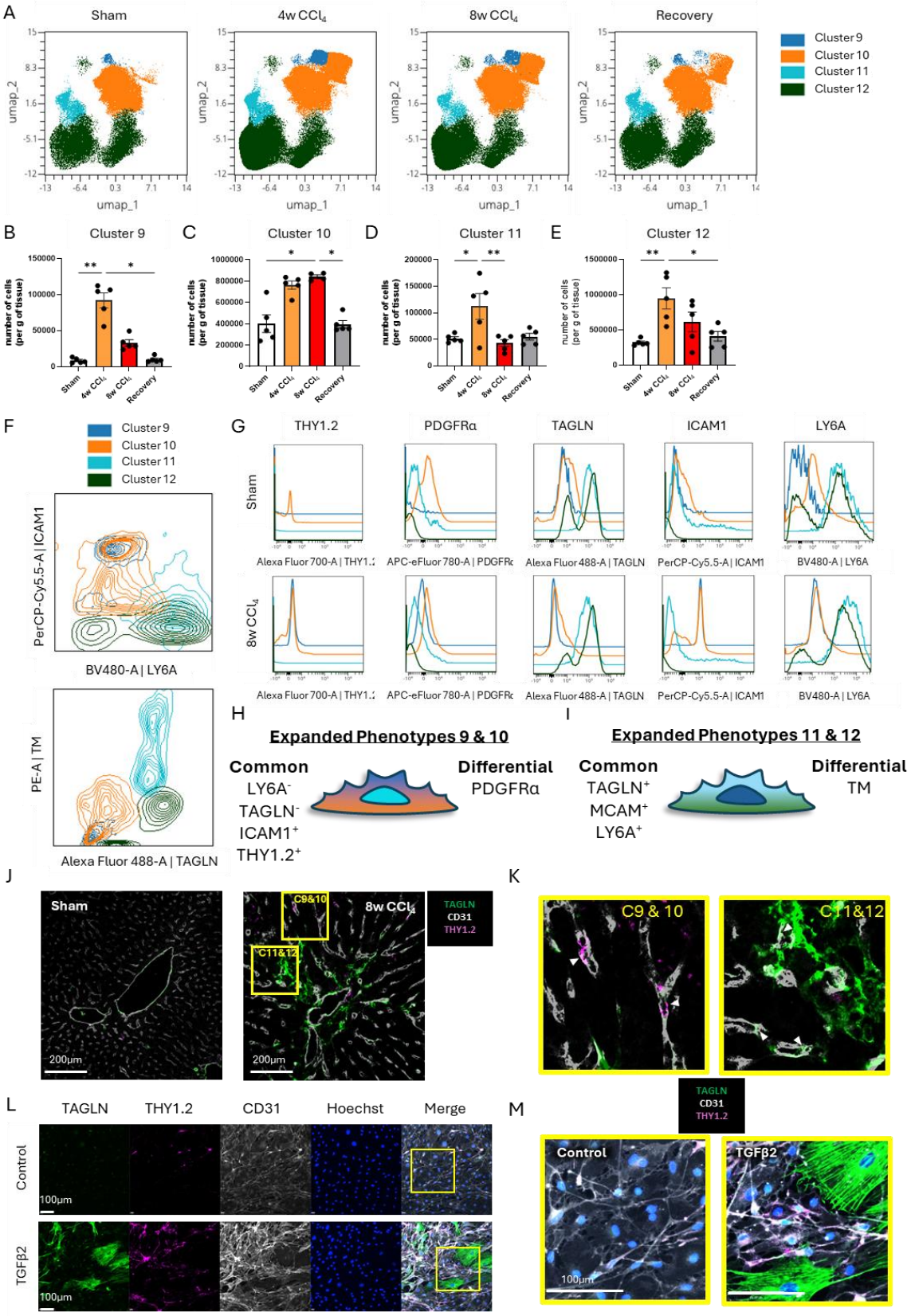

A
Sham
4w CCl4
8w CCl4
Recovery
Cluster 9
Cluster 10
Cluster 11
Cluster 12
umap_1
umap_2
B
Cluster 9
C
Cluster 10
D
Cluster 11
E
Cluster 12
number of cells (per g of tissue)
Sham
4w CCl4
8w CCl4
Recovery
F
PerCP-Cy5.5-A | ICAM1
BV480-A | LY6A
PE-A | TM
Alexa Fluor 488-A | TAGLN
G
THY1.2
PDGFRα
TAGLN
ICAM1
LY6A
Sham
8w CCl4
Alexa Fluor 700-A | THY1.2
APC-eFluor 780-A | PDGFRα
Alexa Fluor 488-A | TAGLN
PerCP-Cy5.5-A | ICAM1
BV480-A | LY6A
H
Expanded Phenotypes 9 & 10
Common
LY6A-
TAGLN-
ICAM1+
THY1.2+
Differential
PDGFRα
I
Expanded Phenotypes 11 & 12
Common
TAGLN+
MCAM+
LY6A+
Differential
TM
J
Sham
8w CCl4
C9&10
C11&12
TAGLN
CD31
THY1.2
200μm
K
C9 & 10
C11&12
L
TAGLN
THY1.2
CD31
Hoechst
Merge
Control
TGFβ2
100μm
M
Control
TGFβ2

Fig. 4. Phenotyping the characteristics of Expanded EC clusters by FSFC. (A) UMAP visualisation of four Expanded EC clusters; 9 (dark blue), 10 (orange), 11 (teal) and 12 (dark green). (B-E) Quantification of the number of cells in each cluster expressed as number of cells per gram of liver tissue for (B) cluster 9, (C) cluster 10, (D) cluster 11 and (E) cluster 12 per treatment group. (F) Contour plots of clusters 9, 10, 11 and 12 for LY6A vs ICAM1 (top panel) and transgelin (TAGLN) vs thrombomodulin (TM) (bottom panel). (G) Histograms of clusters 9, 10, 11 and 12 comparing the mean fluorescence intensity (MFI) for THY1.2, PDGFRα, TAGLN, ICAM1 and LY6A in Sham (top panels) and 8w CCl_4-treated animals (bottom panels). (H, I) Schematics summarising the common and differential traits of (H) clusters 9 and 10 and (I) clusters 11 and 12. (J) Representative confocal images of IF staining for CD31 (white), THY1.2 (magenta) and TAGLN (green) in Sham and 8w CCl_4-treated liver sections. Scale bar= 200µm (K) Magnified examples of $CD31^+THY1.2^+$ cells (left panel) and $CD31^+TAGLN^+$ cells (right panel). (L) Representative confocal images of IF staining for CD31 (white), THY1.2 (magenta), TAGLN (green) and Hoechst (blue) from isolated mouse EC cultured in the presence of either media alone or 10ng/ml TGFβ2 for 48h. Scale bar= 100µm. (M) Magnified examples of $CD31^+THY1.2^+$ cells and $CD31^+TAGLN^+$ cells in control (left panel) and TGFβ2-treated cells (right panel). Scale bar= 100µm. For statistical analysis, Kruskal-Wallis tests with Dunn's multiple comparisons test were used for figures B and C and one-way ANOVA tests with Tukey's multiple comparisons test were used for figures D and E. *; $p<0.05$, **; $p<0.01$.

Fig. 5

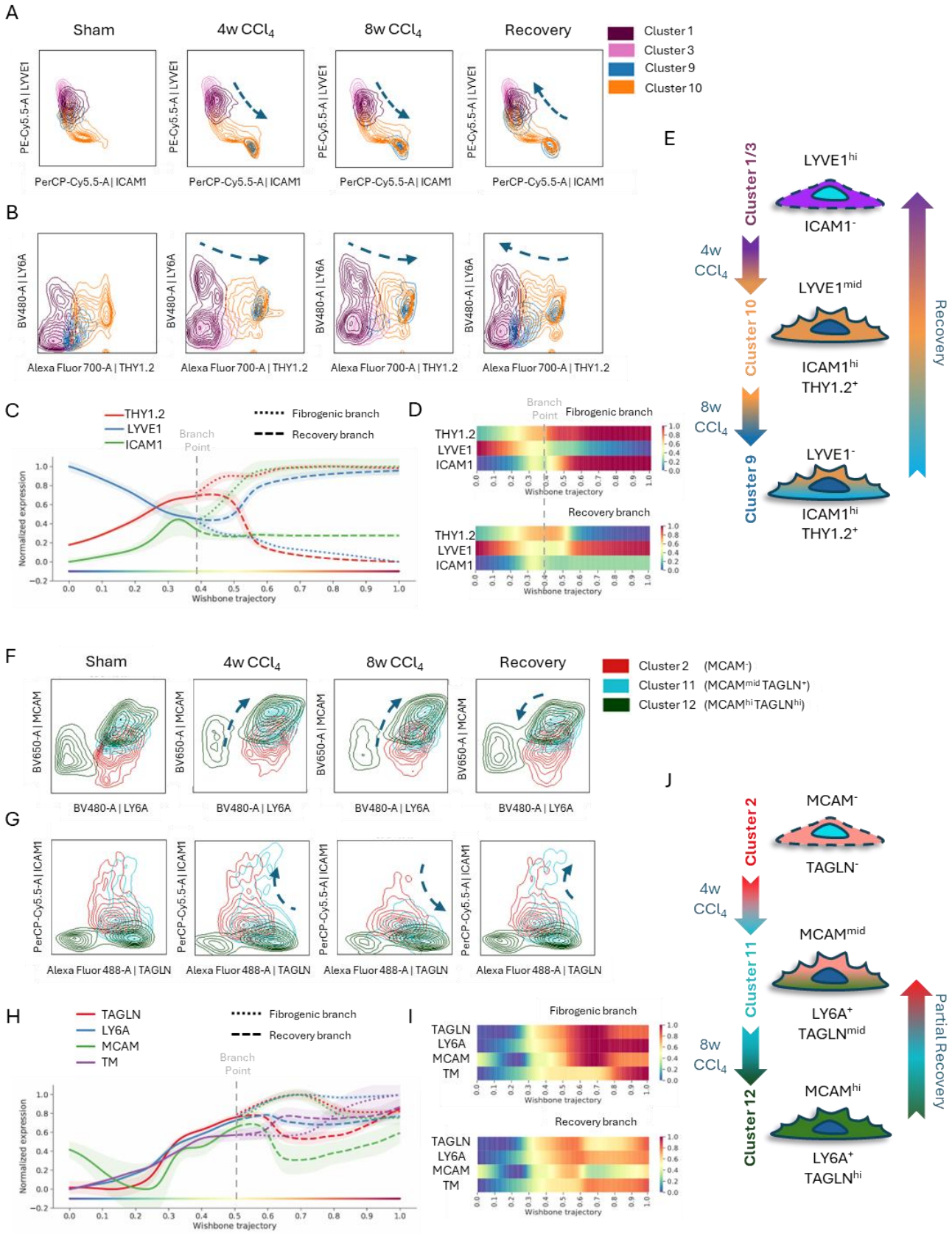

A
Sham
4w CCl4
8w CCl4
Recovery
Cluster 1
Cluster 3
Cluster 9
Cluster 10
PE-Cy5.5-A | LYVE1
PerCP-Cy5.5-A | ICAM1
B
BV480-A | LY6A
Alexa Fluor 700-A | THY1.2
C
THY1.2
LYVE1
ICAM1
Branch Point
Fibrogenic branch
Recovery branch
Normalized expression
Wishbone trajectory
D
Fibrogenic branch
Recovery branch
E
Cluster 1/3
LYVE1hi
ICAM1-
4w CCl4
Cluster 10
LYVE1mid
ICAM1hi THY1.2+
8w CCl4
Cluster 9
LYVE1-
ICAM1hi THY1.2+
Recovery
F
Cluster 2 (MCAM-)
Cluster 11 (MCAMmid TAGLN+)
Cluster 12 (MCAMhi TAGLNhi)
BV650-A | MCAM
BV480-A | LY6A
G
PerCP-Cy5.5-A | ICAM1
Alexa Fluor 488-A | TAGLN
H
TAGLN
LY6A
MCAM
TM
I
J
Cluster 2
MCAM-
TAGLN-
4w CCl4
Cluster 11
MCAMmid
LY6A+ TAGLNmid
8w CCl4
Cluster 12
MCAMhi
LY6A+ TAGLNhi
Partial Recovery

Fig. 5. Derivation of Expanded FAEC from Reduced EC clusters through acquisition of EndMT-traits. (A, B) Contour plots of clusters 1, 3, 9 and 10 for (A) ICAM1 vs LYVE1 and (B) THY1.2 vs LY6A. Arrow denotes the shifting phenotypic profile of the clusters for the different treatment groups. (C) Wishbone trajectory analysis of clusters 1, 3, 9 and 10 for THY1.2 (red), LYVE1 (blue) and ICAM1 (green) (D) Heatmap displaying the expression profile of the Fibrogenic branch (top panel) or Recovery branch (bottom panel) for THY1.2, LYVE1 and ICAM1. (E) Schematic summarising the trajectory and profile of clusters 1, 3, 9 and 10 during fibrogenesis and tissue recovery. (F, G) Contour plots of clusters 2, 11 and 12 for (F) LY6A vs MCAM and (G) transgelin (TAGLN) vs ICAM1. Arrow denotes the shifting phenotypic profile of the clusters for the different treatment groups. (H) Wishbone trajectory analysis of clusters 2, 11 and 12 for TAGLN (red), LY6A (blue), MCAM (green) and TM (purple). (I) Heatmap displaying the expression profile of the Fibrogenic branch (top panel) or Recovery branch (bottom panel) for TAGLN, LY6A, MCAM and thrombomodulin (TM). (J) Schematic summarising the trajectory and profile of clusters 2, 11 and 12 during fibrogenesis and tissue recovery.

Fig. 6

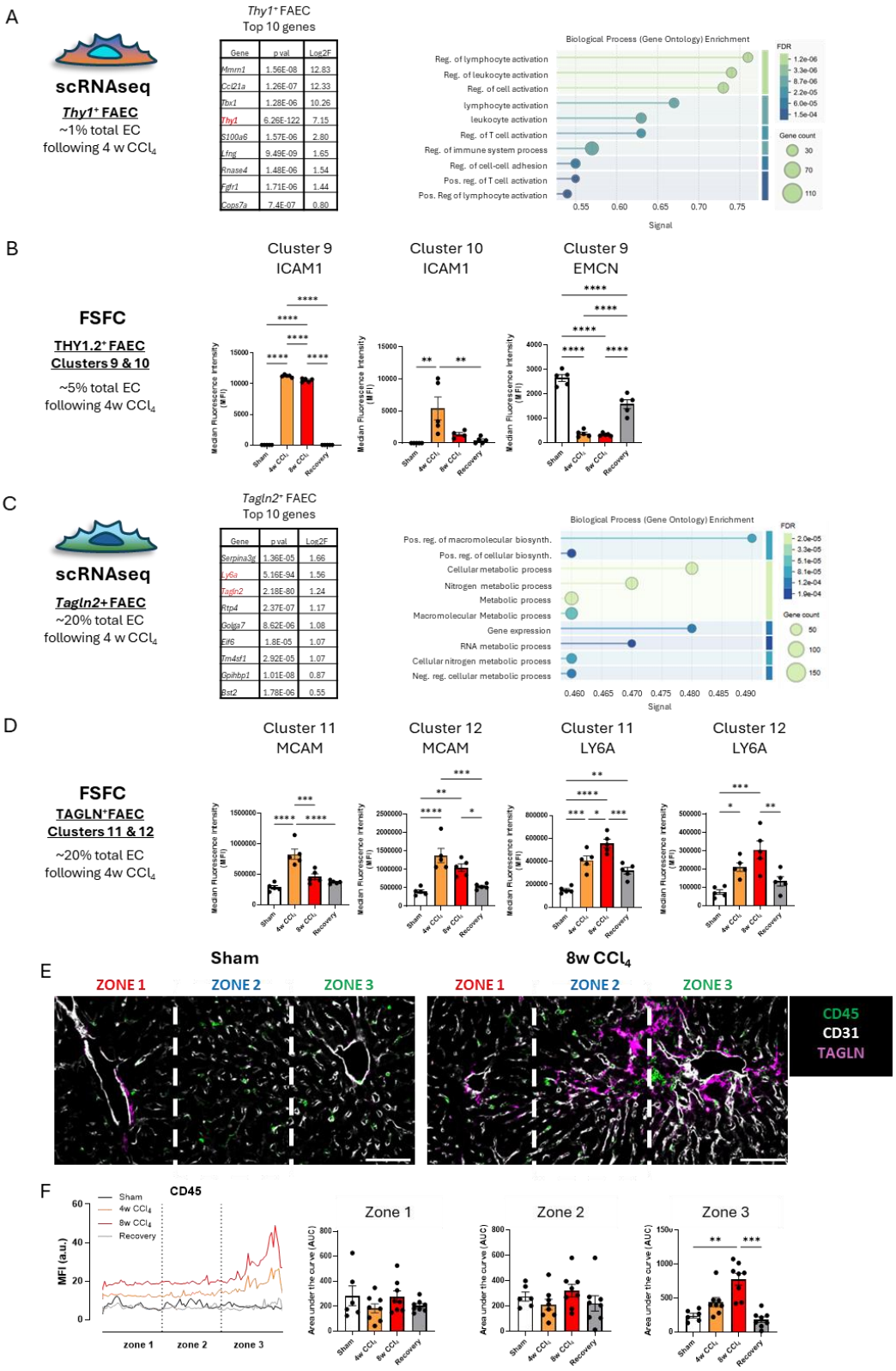

A
scRNAseq
Thy1+ FAEC
~1% total EC
following 4 w CCl4
Thy1+ FAEC
Top 10 genes
Gene | p val | Log2F
Mmrn1 | 1.56E-08 | 12.83
Ccl21a | 1.26E-07 | 12.33
Tbx1 | 1.28E-06 | 10.26
Thy1 | 6.26E-122 | 7.15
S100a6 | 1.57E-06 | 2.80
Lfng | 9.49E-09 | 1.65
Rnase4 | 1.48E-06 | 1.54
Fgfr1 | 1.71E-06 | 1.44
Cops7a | 7.4E-07 | 0.80
Biological Process (Gene Ontology) Enrichment
Reg. of lymphocyte activation
Reg. of leukocyte activation
Reg. of cell activation
lymphocyte activation
leukocyte activation
Reg. of T cell activation
Reg. of immune system process
Reg. of cell-cell adhesion
Pos. reg. of T cell activation
Pos. Reg of lymphocyte activation
Signal
FDR
Gene count
B
FSFC
THY1.2+ FAEC
Clusters 9 & 10
~5% total EC
following 4w CCl4
Cluster 9 ICAM1
Cluster 10 ICAM1
Cluster 9 EMCN
Median Fluorescence Intensity (MFI)
Sham
4w CCl4
8w CCl4
Recovery
C
scRNAseq
Tagln2+ FAEC
~20% total EC
following 4 w CCl4
Tagln2+ FAEC
Top 10 genes
Gene | p val | Log2F
Serpina3g | 1.36E-05 | 1.66
Ly6a | 5.16E-94 | 1.56
Tagln2 | 2.18E-80 | 1.24
Rtp4 | 2.37E-07 | 1.17
Golga7 | 8.62E-06 | 1.08
Ifi6 | 1.8E-05 | 1.07
Tm4sf1 | 2.92E-05 | 1.07
Gpihbp1 | 1.01E-08 | 0.87
Bst2 | 1.78E-06 | 0.55
Biological Process (Gene Ontology) Enrichment
Pos. reg. of macromolecular biosynth.
Pos. reg. of cellular biosynth.
Cellular metabolic process
Nitrogen metabolic process
Metabolic process
Macromolecular Metabolic process
Gene expression
RNA metabolic process
Cellular nitrogen metabolic process
Neg. reg. cellular metabolic process
Signal
FDR
Gene count
D
FSFC
TAGLN+FAEC
Clusters 11 & 12
~20% total EC
following 4w CCl4
Cluster 11 MCAM
Cluster 12 MCAM
Cluster 11 LY6A
Cluster 12 LY6A
Median Fluorescence Intensity (MFI)
Sham
4w CCl4
8w CCl4
Recovery
E
Sham
8w CCl4
ZONE 1
ZONE 2
ZONE 3
CD45
CD31
TAGLN
F
CD45
Sham
4w CCl4
8w CCl4
Recovery
MFI (a.u.)
zone 1
zone 2
zone 3
Zone 1
Zone 2
Zone 3
Area under the curve (AUC)
Sham
4w CCl4
8w CCl4
Recovery

Fig. 6. Transcriptomic and proteomic phenotype of THY1$^+$ and TAGLN$^+$ FAEC. (A) ScRNAseq analysis of *Thy1* expressing liver EC from 4w CCl_4-treated mice compared to all other EC. Table of the top 10 enriched genes and gene ontology analysis of putative biological processes enriched in *Thy1*$^+$ EC based on their differentially expressed genes. (B) FSFC analysis of median fluorescence intensity (MFI) of Thy1.2$^+$ FAEC for ICAM1 and EMCN in cluster 9 and cluster 10. (C) ScRNAseq analysis of *Tagln2* expressing liver EC from 4w CCl_4-treated mice compared to all other EC. Table of the top 10 enriched genes and gene ontology analysis of putative biological processes enriched in *Tagln2*$^+$ EC based on their differentially expressed genes. (D) FSFC analysis of median fluorescence intensity (MFI) of TAGLN$^+$ FAEC for MCAM and LY6A in cluster 11 and 12. (D) Representative confocal images of IF staining for CD45 (green), CD31 (white) and TAGLN (magenta) in Sham and 8w CCl_4-treated livers. Scale bar= 100µm. (E) Profile plot of CD45$^+$ MFI across the portal-central vein axis in Sham, CCl_4-treated and recovery experimental groups. (F) Quantification of the profile plot area under the curve for Zone 1, Zone 2 and Zone 3. (a.u.: arbitrary units). For statistical analysis, one-way ANOVA tests with one-way ANOVA tests with Tukey's multiple comparisons tests were used. N=5 animals/group. *; $p<0.05$, **; $p<0.01$, ***; $p<0.001$, ****; $p<0.0001$.

Fig 7.

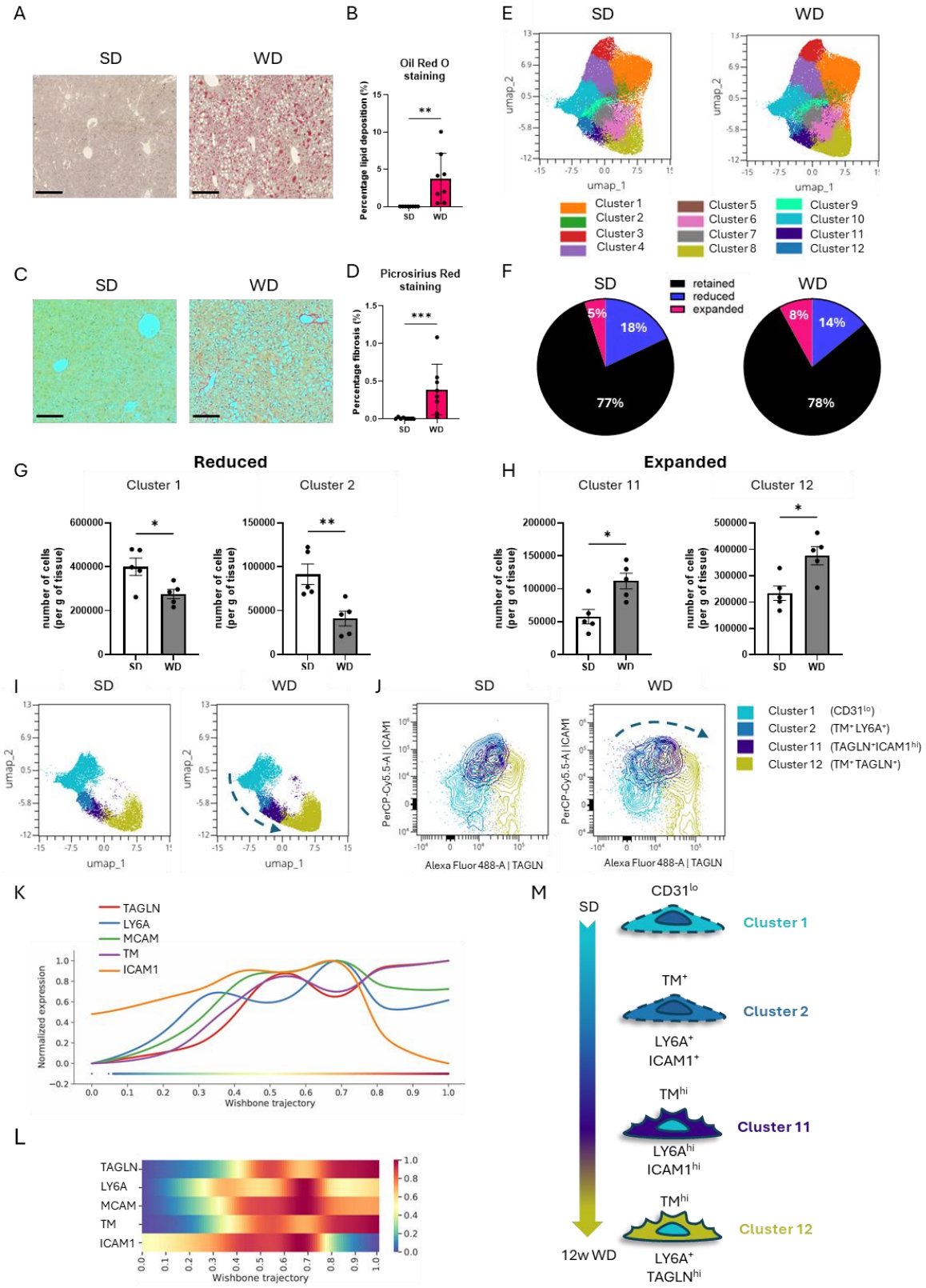

A
SD
WD
B
Oil Red O staining
Percentage lipid deposition (%)
**
SD
WD
E
SD
WD
umap_1
umap_2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Cluster 11
Cluster 12
C
SD
WD
D
Picrosirius Red staining
Percentage fibrosis (%)
***
SD
WD
F
SD
WD
retained
reduced
expanded
5%
18%
77%
8%
14%
78%
G
Reduced
Cluster 1
Cluster 2
number of cells (per g of tissue)
*
**
SD
WD
H
Expanded
Cluster 11
Cluster 12
number of cells (per g of tissue)
*
*
SD
WD
I
SD
WD
umap_1
umap_2
J
SD
WD
PerCP-Cy5.5-A | ICAM1
Alexa Fluor 488-A | TAGLN
Cluster 1 (CD31lo)
Cluster 2 (TM+ LY6A+)
Cluster 11 (TAGLN+ICAM1hi)
Cluster 12 (TM+ TAGLN+)
K
TAGLN
LY6A
MCAM
TM
ICAM1
Normalized expression
Wishbone trajectory
L
TAGLN
LY6A
MCAM
TM
ICAM1
Wishbone trajectory
M
SD
CD31lo
Cluster 1
TM+
Cluster 2
LY6A+
ICAM1+
TMhi
Cluster 11
LY6Ahi
ICAM1hi
TMhi
Cluster 12
LY6A+
TAGLNhi
12w WD

Fig. 7. Expansion of TAGLN$^+$ FAEC in a 12w Western Diet model of MASLD. (A) Representative images of Oil Red O staining of liver samples from 12w standard diet (SD) and Western diet (WD) fed mice. Scale bars = 250 µm. (B) Quantification of Oil Red O staining expressed as a percentage (%) of lipids detected over the total tissue area. (C) Representative images of Picrosirius Red staining of liver samples from 12w SD and WD fed mice. Scale bars = 250 µm. (D) Quantification of Picrosirius Red staining expressed as a percentage (%) of collagen detected over the total tissue area. (E) UMAP plots of CD45$^-$ CD31$^+$ EC clusters from SD or WD fed mice. (F) Pie charts of the percentage composition of total EC characterised as 'Retained' (black), 'Reduced' (blue) and 'Expanded' (pink) for each diet group. (G, H) Quantification of the number of cells per gram of liver tissue for (G) Reduced cluster 1 and cluster 2 and (H) Expanded cluster 11 and cluster 12. (I) UMAP of cluster 1 (turquoise), cluster 2 (blue), cluster 11 (purple) and cluster 12 (yellow) from SD or WD groups. (J) Contour plots of transgelin (TAGLN) vs ICAM1 for clusters 1, 2, 11 and 12 for SD or WD groups. Arrow denotes the shifting phenotypic profile of the clusters from SD to WD. (K) Wishbone trajectory analysis of clusters 1, 2, 11 and 12 for TAGLN (red), LY6A (blue), MCAM (green), TM (purple) and ICAM1 (orange). (L) Heatmap of the expression profile of each marker. (M) Schematic summary of the trajectory and profile of clusters 1, 2, 11 and 12. For statistical analysis, unpaired t-tests were used for figures B, G, H and a Mann-Whitney test was used for figure D. N=5-8 animals/group. *; $p<0.05$, **; $p<0.01$, ***; $p<0.001$.

Fig. 8

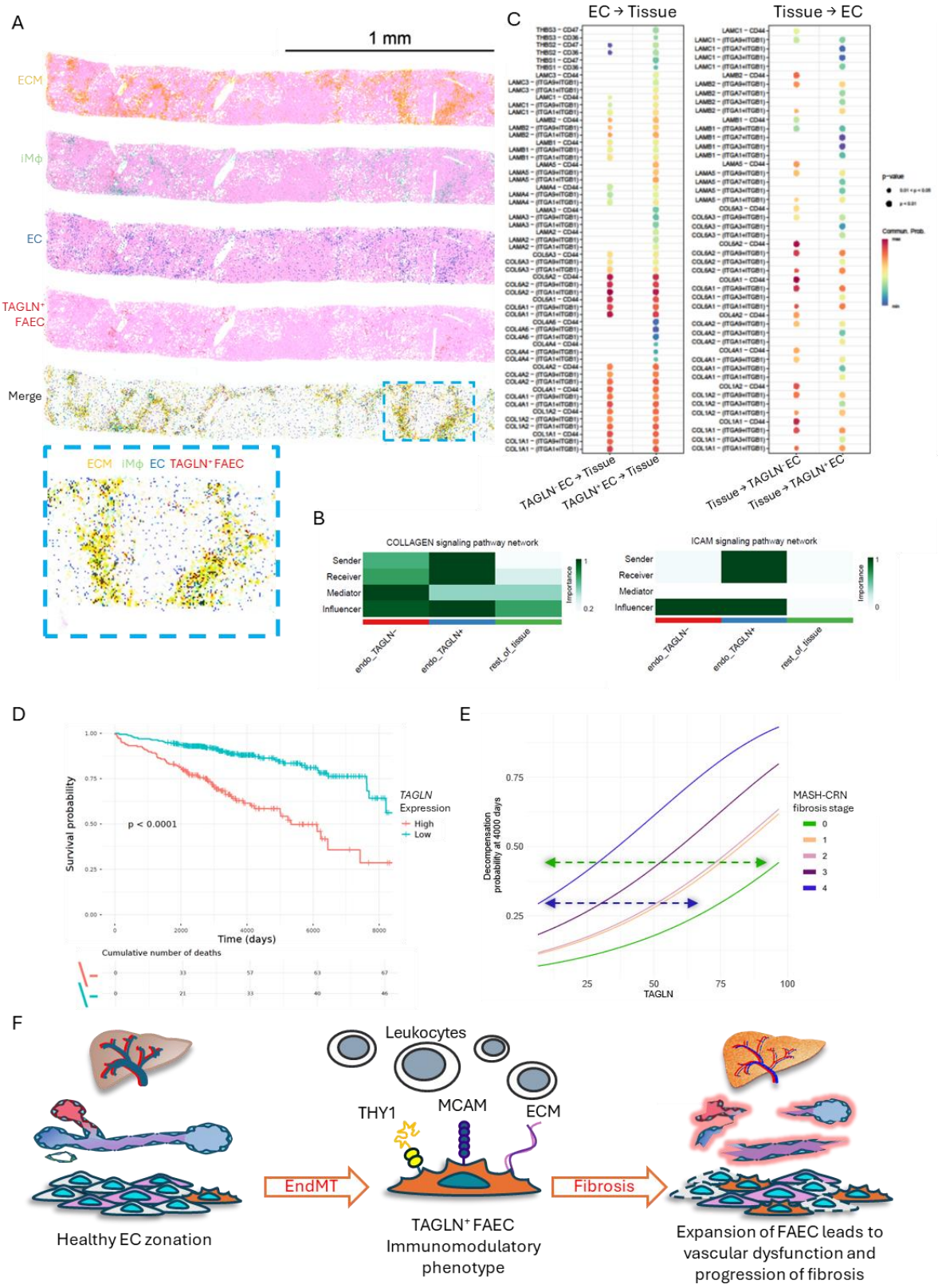
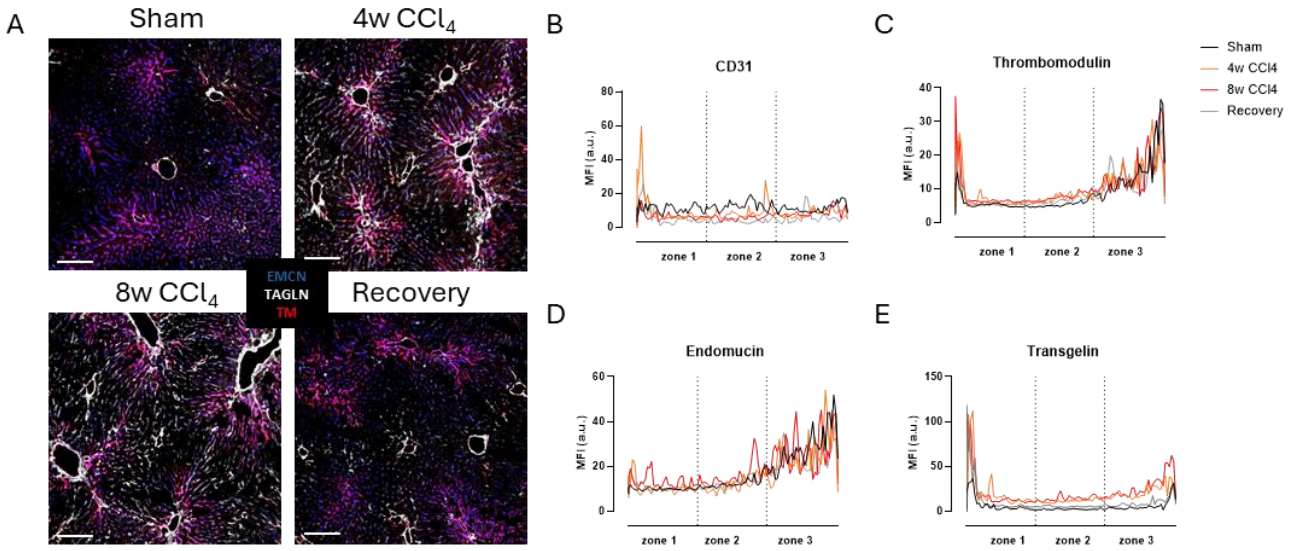

A
1 mm
ECM
iMϕ
EC
TAGLN⁺ FAEC
Merge
ECM iMϕ EC TAGLN⁺FAEC
B
COLLAGEN signaling pathway network
ICAM signaling pathway network
Sender
Receiver
Mediator
Influencer
Importance
endo_TAGLN-
endo_TAGLN+
rest_of_tissue
C
EC → Tissue
Tissue → EC
TAGLN⁻ EC → Tissue
TAGLN⁺ EC → Tissue
Tissue → TAGLN⁻ EC
Tissue → TAGLN⁺ EC
p-value
Commun. Prob.
D
Survival probability
p < 0.0001
TAGLN Expression
High
Low
Time (days)
Cumulative number of deaths
E
Decompensation probability at 4000 days
MASH-CRN fibrosis stage
TAGLN
F
Leukocytes
THY1
MCAM
ECM
EndMT
Fibrosis
Healthy EC zonation
TAGLN⁺ FAEC Immunomodulatory phenotype
Expansion of FAEC leads to vascular dysfunction and progression of fibrosis

Fig. 8. Expression of TAGLN$^+$ FAEC is a conserved feature of MASLD with the potential to be a biomarker of progressive disease. (A) Visualization of spatial transcriptomic clusters for *k*-means shows enrichment data for a patient with Ishak score of 2-3 identifying extracellular matrix (ECM; yellow), inflammatory macrophages (iMϕ), all endothelial cells (EC; blue) and TAGLN$^+$EC (red). Zoomed in region highlighted in blue box (B) Comparison between TAGLN$^-$ EC and TAGLN$^+$EC for collagen and ICAM signalling network displayed as heatmaps. (C) CellChat analysis comparing enriched of putative interaction patterns from EC to tissue and tissue to EC. (D) Kaplan–Meier survival curves comparing patients stratified into high and low TAGLN expression groups (p-value of log-rank test comparing groups) in the biopsy subset of the SteatoSITE cohort (n=505). (E) Cumulative incidence functions for liver-related events at 4000 days post-biopsy as a function of TAGLN expression, stratified by pathologist-scored MASH-CRN fibrosis stage (n=435; F0 (green), F1 (orange), F2 (pink), F3 (purple) and F4 (blue). Scale bar = 50 μm. For statistical analysis a one-way ANOVA with Dunnett's multiple comparisons test was used. **; $p<0.01$. (F) Summary schematic of the role of FAEC in the development and progression of vascular dysfunction, fibrosis and MASLD.

Supplementary Fig. 1



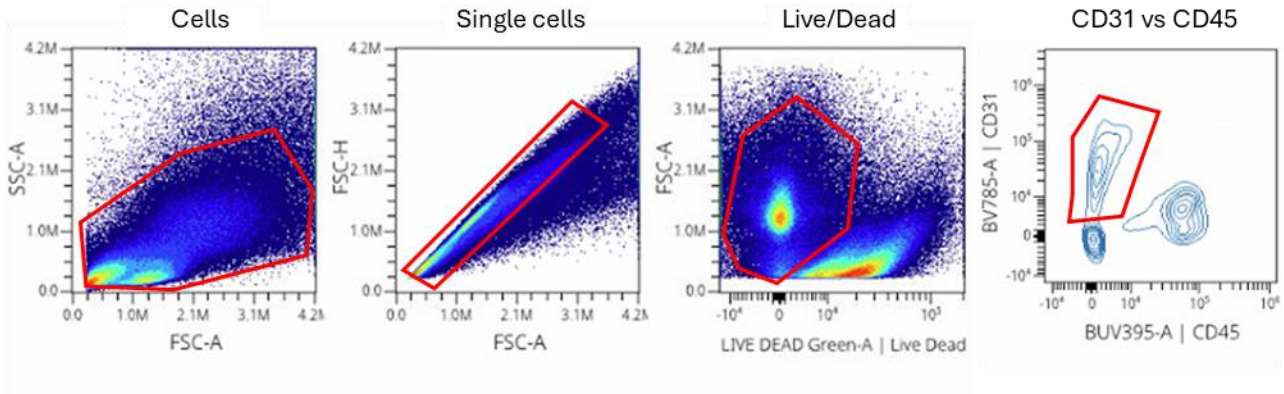
Supplementary Fig. 1. Endothelial plasticity and tissue remodelling in response to CCl_4-induced liver fibrosis. (A) Representative confocal images of thrombomodulin (TM) (red), endomucin (EMCN) (blue) and transgelin (TAGLN) (white) staining in the liver of Sham, 4w and 8w CCl_4-treated and Recovery group. Scale bar= 200µm. Profile plots for (B) CD31, (C) TM, (D) EMCN and (E) TAGLN showing the shift in zonation in Sham (black), 4w CCl_4 (orange), 8w CCl_4 (red) and recovery (grey) groups.

Supplementary Fig. 2

A

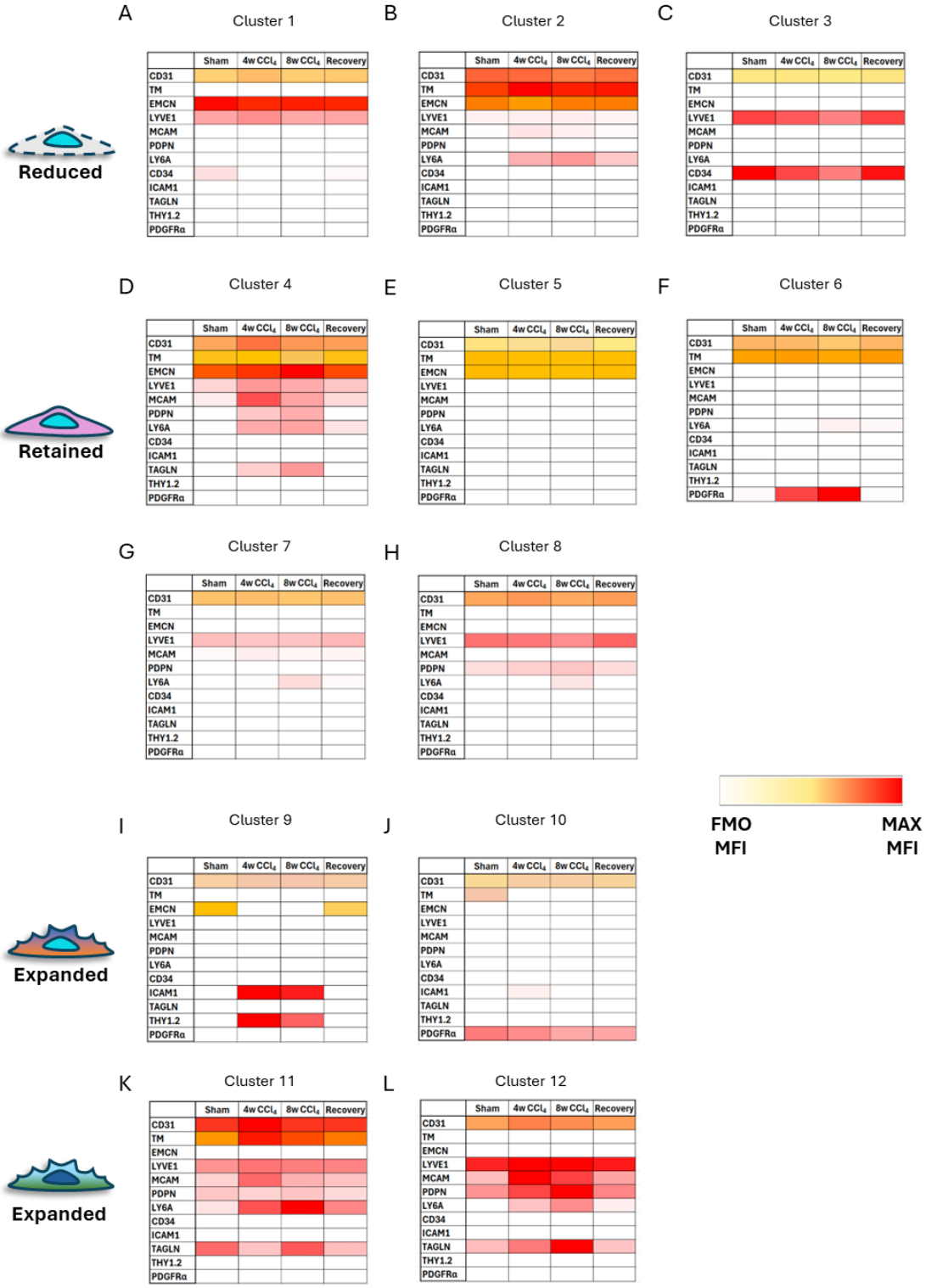
Target	Markers	Target	Markers	Target	Markers
Live Dead	Viability	CD31	EC Identity	LY6A	Activation
CD45	Pan-Leukocyte	TM	EC Identity	ICAM1	Activation
		EMCN	EC Identity	CD34	Activation
		MCAM	EC Identity	TAGLN	Mesenchymal
		LYVE1	EC Identity	PDGFRα	Mesenchymal
		PDPN	EC Identity	THY1.2	Mesenchymal

B



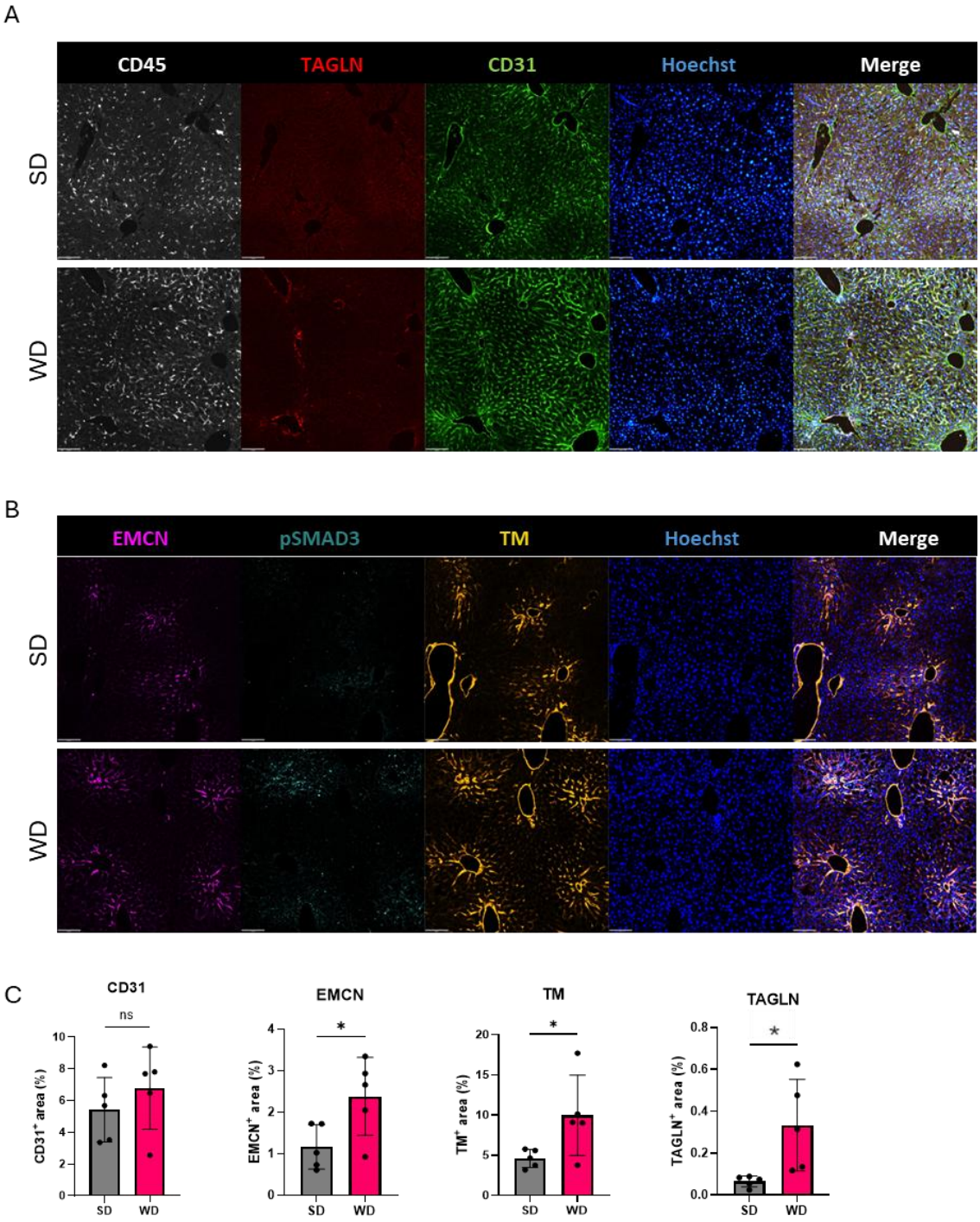
Supplementary Fig. 2. FSFC EC panel and gating. (A) Schematic for endothelial FSFC panel and (B) Gating strategy for the analysis of $CD45^{-}CD31^{+}$ endothelial cells. From left to right: a gate for all cells was firstly drawn using the forward scatter area (FSC-A) vs the side scatter area (SSC-A) plot. Then, single cells were selected using the FSC-A vs forward scatter height (FSC-H) plot. The dead cells were excluded from the analysis using the Live/Dead Fixable stain vs FSC-A plot. In the live cells gate, the $CD45^{-}CD31^{+}$ and $CD45^{+}CD31^{-}$ cells were gated separately using the CD45 vs CD31 plot. The $CD45^{-}CD31^{+}$ gate (red) was used for the analysis of endothelial cells.

Supplementary Fig. 3



Supplementary Fig. 3. Heatmaps of the phenotypic properties of each EC cluster split by treatment group. Heatmaps of median fluorescence intensity (MFI) for (A-C) Reduced clusters 1-3 (D-H) Retained clusters 4-8 and (I-L) Expanded clusters 9-12. The heatmaps were created using full spectrum flow cytometry data. The fluorescence minus one (FMO) MFI for each marker was used as the threshold for positive signal.

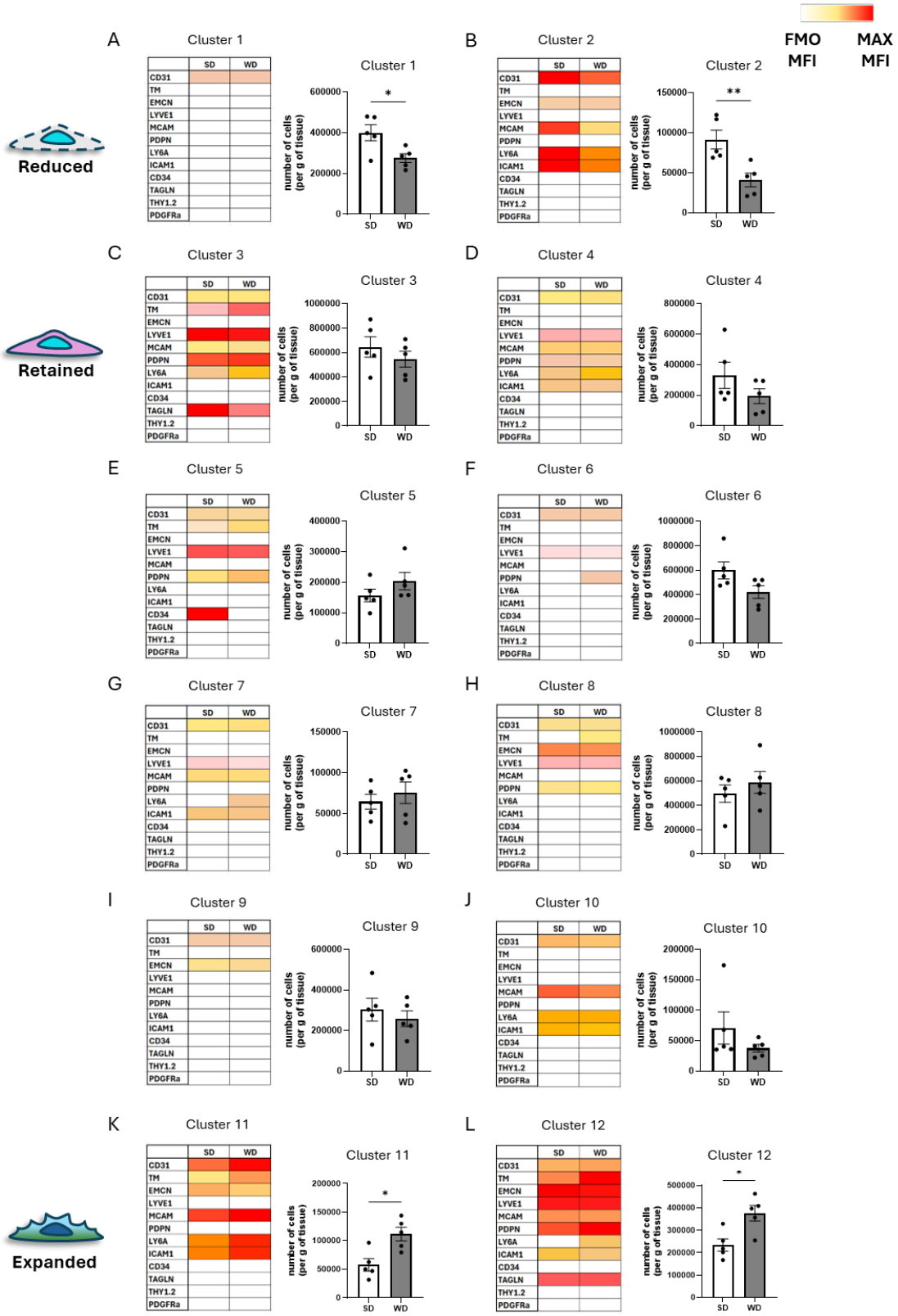
Supplementary Fig. 4



Supplementary Figure 4. Changes in vascular patterning in response to a 12w WD. (A, B) Representative confocal IF images of liver sections from mice on a standard diet (SD) or Western

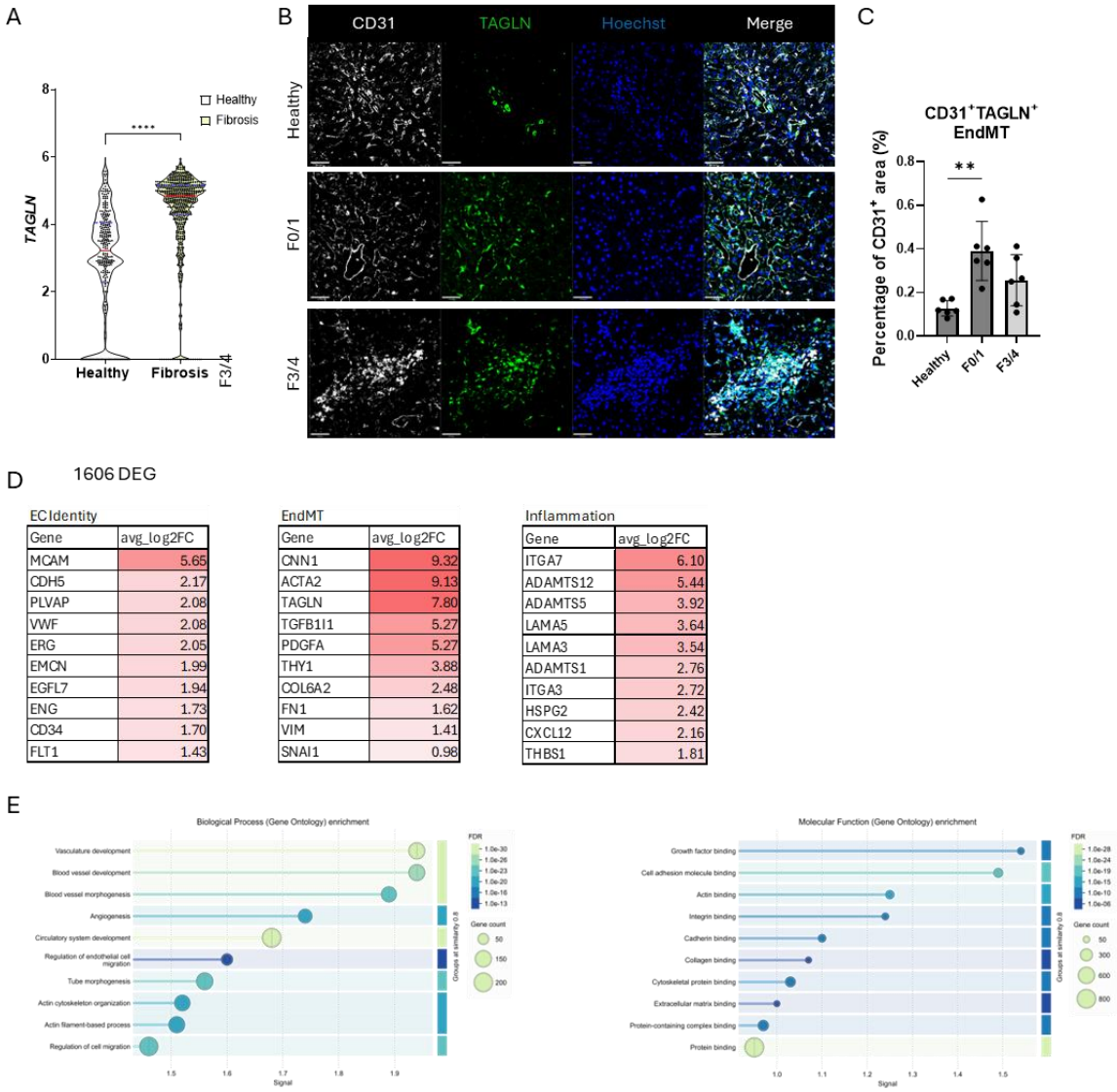
diet (WD) for 12 weeks stained for (A) CD45 (white), transgelin (TAGLN) (red) and CD31 (green) and (B) endomucin (EMCN) (red), pSMAD3 (turquoise) and thrombomodulin (TM) (yellow). Scale bars= 200µm. (C) Quantification of positive fluorescence signal for (C) CD31, EMCN, TM, and TAGLN. For statistical analysis, unpaired t-tests were used. ns; not significant, *; $p<0.05$.

Supplementary Fig. 5



Supplementary Figure 5. Heatmaps of the phenotypic properties of each EC cluster split by diet group. Heatmaps of median fluorescence intensity (MFI) for (A, B) Reduced clusters 1 and 2, (C-J) Retained clusters 3-10 and (K, L) Expanded clusters 11 and 12. The heatmaps were created using full spectrum flow cytometry data. The fluorescence minus one (FMO) MFI for each marker was used as the threshold for positive signal.

Supplementary Fig. 6



ECIdentity	
Gene	avg_log2FC
MCAM	5.65
CDH5	2.17
PLVAP	2.08
VWF	2.08
ERG	2.05
EMCN	1.99
EGFL7	1.94
ENG	1.73
CD34	1.70
FLT1	1.43

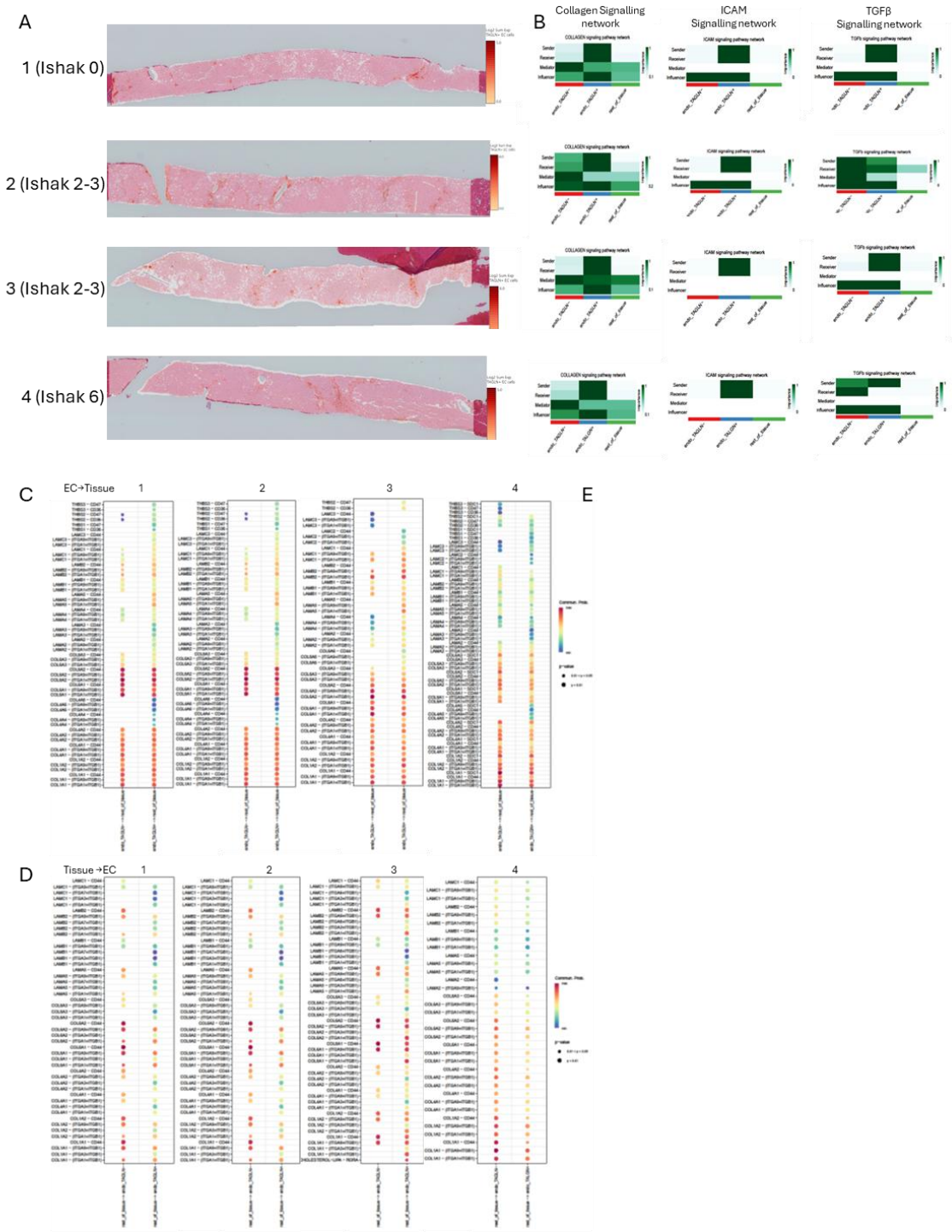
EndMT	
Gene	avg_log2FC
CNN1	9.32
ACTA2	9.13
TAGLN	7.80
TGFB1I1	5.27
PDGFA	5.27
THY1	3.88
COL6A2	2.48
FN1	1.62
VIM	1.41
SNAI1	0.98

Inflammation	
Gene	avg_log2FC
ITGA7	6.10
ADAMTS12	5.44
ADAMTS5	3.92
LAMA5	3.64
LAMA3	3.54
ADAMTS1	2.76
ITGA3	2.72
HSPG2	2.42
CXCL12	2.16
THBS1	1.81

Supplementary Figure 6. Expression of TAGLN⁺ FAEC is a conserved feature of MASLD with the potential to be a biomarker of progressive disease. (A) Violin plot of scRNAseq analysis of cell number (dots) and median expression profile of *TAGLN*⁺ EC from liver biopsies from MASLD patients after bariatric surgery. (B) Representative confocal images of patient biopsies stained for CD31 (white), transgelin (TAGLN) (green) and nuclei (Hoechst; blue). Scale bar= 50 µm. (C) Quantification of CD31⁺TAGLN⁺ EndMT cells expressed as a percentage of total CD31⁺ area (N=6 patients per group). (D) Gene enrichment analysis of *TAGLN*⁺ FAEC presented

as a Table of top 10 enriched EC identity genes, EndMT genes and inflammation identified using scRNAseq. (E) Gene ontology analysis of all significantly enriched genes (Log2FC >0.5, P adj <0.05) for Biological Processes and Molecular Function.

Supplementary Fig. 7



Supplementary Figure 7. Spatial Transcriptomic analysis of four MASLD patients. (A) Localisation of TAGLN^{+}EC in patients with different stratification from F0-F4. (B) Heatmap of

enriched Collagen, ICAM and TGFβ signalling pathways per patient. CellChat analysis for each patient for (C) TAGLN^{-}EC vs TAGLN^{+}EC to tissue and (D) tissue to TAGLN^{-}EC vs TAGLN^{+}EC.

Supplementary Table 1

Biopsy ID	Diagnosis	Age (years)	Ethnicity	Sex (F/M)	Ethanol (ml/ week)	BMI (kg/m^2)	Liver Stiffness Measurement (kPa)	Liver Steatosis Measurement (dB/m)	Histological fibrosis stage (Ishak)	FIB-4	Co-morbidities
1	MASLD	53	White-British	F	170	48.08	12.6	265	0	0.64	Hypertension, Type 2 Diabetes, Atrial Fibrillation, previously resected thyroid cancer
2	MASLD	56	White-British	M	0	44.26	33.0	382	2-3	1.28	Depression, Hypertension, Gout, Obstructive Sleep Apnoea, Osteoarthritis
3	MASLD	21	Mixed – White and Asian	F	0	22.30	16.5	258	2-3	0.58	Depression
4	MASLD	48	White-British	M	80	23.13	15.8	332	6	3.37	Nocturnal epilepsy

Supplementary Table 1. Patient profile for spatial transcriptomic data.

Supplementary Table 2

Target	Host	Clonality	Fluorophore	Dilution	Product number	Supplier
CD31	Goat	polyclonal	AF488	1:100	FAB3628G	R&D Systems
CD31	Sheep	polyclonal	unconjugated	1:100	AF806	R&D Systems
TM	Goat	polyclonal	unconjugated	1:500	AF3894	R&D Systems
EMCN	Rat	Monoclonal (V.7C7.1)	unconjugated	1:200	ab106100	Abcam
LYVE1	Rat	Monoclonal (ALY7)	unconjugated	1:100	48-0443-82	Invitrogen
MCAM	Rat	Monoclonal (ME-9F1)	unconjugated	1:100	130-101-866	Miltenyi
TAGLN	Rabbit	Polyclonal	unconjugated	1:500	ab14106	Abcam
THY1.2	Rat	Monoclonal (G7)	unconjugated	1:100	105201	Biolegend
phospho-SMAD3	Rabbit	Monoclonal (EP823Y)	unconjugated	1:100	ab52903	Abcam
CD45	Rat	F30-11	unconjugated	1:100	103102	Biolegend

Host	Reactivity	Fluorophore	Dilution	Product number	Supplier
donkey	goat	488	1:1000	A32814	Invitrogen
donkey	goat	555	1:1000	A21432	Invitrogen
donkey	rabbit	488	1:1000	A21206	Invitrogen
donkey	rabbit	555	1:1000	A32794	Invitrogen
donkey	rabbit	647	1:1000	A31573	Invitrogen
donkey	rat	488	1:1000	A21208	Invitrogen
donkey	rat	550	1:1000	SA5-10027	Invitrogen
chicken	rat	647	1:1000	A21472	Invitrogen

Supplementary Table 2. List of primary (top) and secondary (bottom) antibodies used for tissue immunofluorescence experiments. (N/A: not applicable, AF: Alexa Fluor)

Supplementary Table 3

Target	Fluorophore	Clone	Dilution for liver samples	Dilution for heart samples	Product number	Supplier
CD31	BV785	390	1:320	1:640	102435	Biolegend
CD34	BV421	MEC14.7	1:160	1:80	119321	Biolegend
CD45	BUV395	30-F11	1:640	1:640	564279	BD Biosciences
EMCN	eFluor660	V7G7	1:640	1:320	50-5851-82	Invitrogen
ICAM1	PerCP-Cy5.5	YN1/1.7.4	1:320	1:160	116124	Biolegend
LY6A	BV480	D7	1:320	1:640	746663	BD Biosciences
LYVE1	PE-Cy5.5	ALY7	1:640	1:640	NBP1-43411	Novus Biologics
MCAM	BV650	ME-9F1	1:320	1:640	740636	BD Biosciences
PDGFRα	APC-eFluor780	APA7	1:160	1:640	47-1401-82	Thermo Scientific
PDPN	PE-Cy7	8.1.1	1:320	1:640	127411	Biolegend
TAGLN	AF488	TAGLN/247	1:320	1:320	NBP2-47688	Novus Bio
THY1.2	AF700	30-H12	1:640	1:640	105320	Biolegend
TM	PE	461714	1:160	1:640	FAB3894P	Biotechne

Supplementary Table 3. Antibody details for the endothelial panel used for full spectrum flow cytometry.