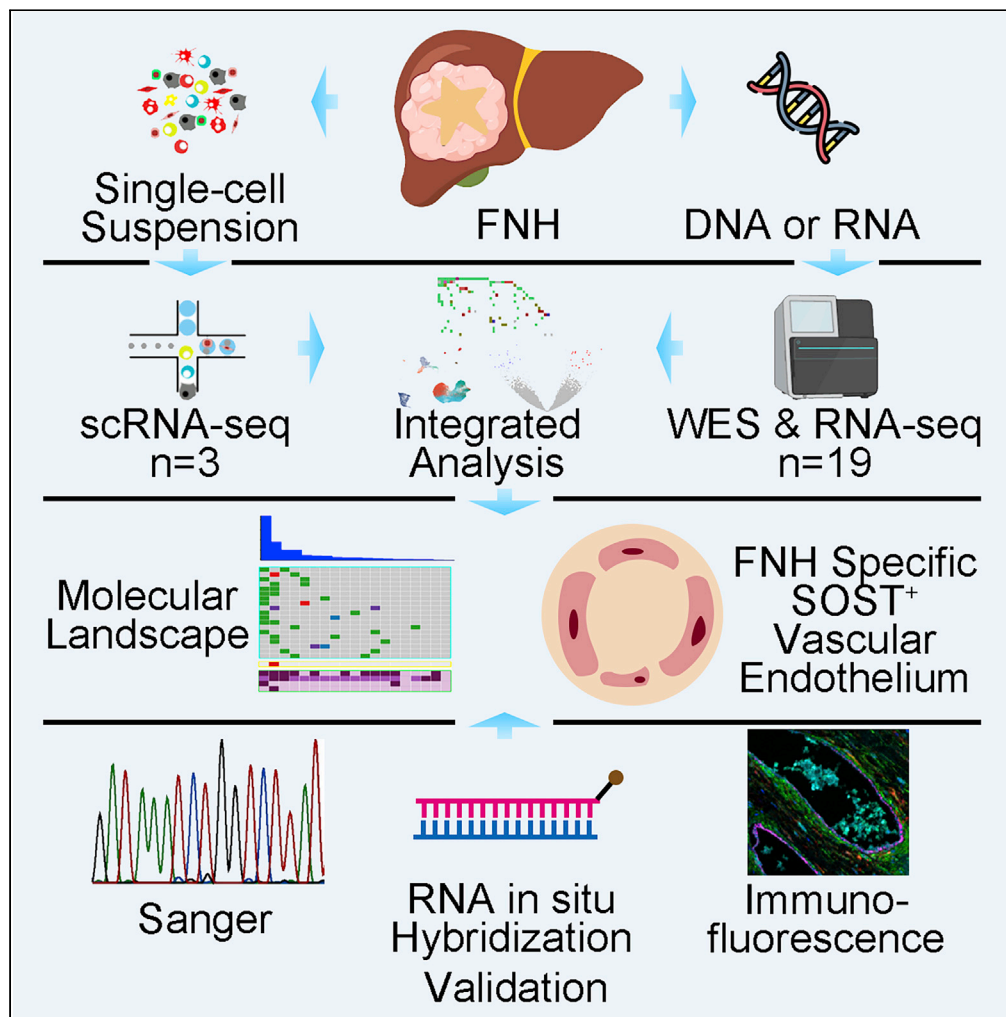


Article

Multi-omics characterization reveals the pathogenesis of liver focal nodular hyperplasia



Yuming Liu, Jinmai Zhang, Zhuo Wang, ..., Jia Fan, Yongyong Shi, Qiang Gao

shiyongyong@gmail.com (Y.S.)
gao.qiang@zs-hospital.sh.cn (Q.G.)

Highlights

FNHs are genetically stable, but high mutation cases exist

FNHs have unique transcriptomic modules, and they alter in atypical FNH

FNH has a unique type of SOST-expressing endothelial cells that may promote fibrosis



Article

Multi-omics characterization reveals the pathogenesis of liver focal nodular hyperplasia

Yuming Liu,^{1,6} Jinmai Zhang,^{2,6} Zhuo Wang,^{2,6} Jiaqiang Ma,¹ Ke Wang,² Dongning Rao,¹ Mao Zhang,¹ Youpei Lin,¹ Yingcheng Wu,¹ Zijian Yang,¹ Liangqing Dong,¹ Zhenbin Ding,¹ Xiaoming Zhang,³ Jia Fan,^{1,4} Yongyong Shi,^{2,*} and Qiang Gao^{1,4,5,7,*}

SUMMARY

The molecular landscape and pathogenesis of focal nodular hyperplasia (FNH) have yet to be elucidated. We performed multi-omics approaches on FNH and paired normal liver tissues from 22 patients, followed by multi-level bioinformatic analyses and experimental validations. Generally, FNH had low mutation burden with low variant allele frequencies, and the mutation frequency significantly correlated with proliferation rate. Although no recurrently deleterious genomic events were found, some putative tumor suppressors or oncogenes were involved. Mutational signatures indicated potential impaired mismatch function and possible poison contact. Integrated analyses unveiled a group of FNH specific endothelial cells that uniquely expressed SOST and probably had strong interaction with fibroblasts through PDGFB/PDGFRB pathway to promote fibrosis. Notably, in one atypical FNH (patient No. 11) with pronounced copy number variations, we observed a unique immune module. Most FNH are benign, but molecularly atypical FNH still exist; endothelial cell derived PDGFB probably promotes the fibrogenic process in FNH.

INTRODUCTION

Liver focal nodular hyperplasia (FNH) ranks the second highest morbidity among benign hepatic tumors after hepatic hemangioma (Nahm et al., 2011). FNH is more prevalent in women than men, especially at childbearing age (Luciani et al., 2002; EASL, 2016). Hitherto, surgical resection is the only standard curative treatment of FNH (Koea and Yeong, 2014), and the surgical procedure will inevitably affect quality of life or even severe postoperative complications. However, if left untreated, the patients may worry about whether FNH could become cancerous, as well as its increasing size over time in some progressive cases (Tajiri et al., 2014; Kudo et al., 2008). Despite the common sense that FNH is a benign tumor with almost no malignant potential (Nahm et al., 2011), its spatial proximity to hepatocellular carcinoma (HCC) within the same patients (Chen et al., 2001; Petsas et al., 2006) and its occurrence linking to chemotherapy-induced liver damage are well documented (Zhu et al., 2021; Torri et al., 2021). Thus, exploring how FNH phenotypic heterogeneity is formed and changed over time may provide clinical insights for FNH and HCC diagnosis and treatment design.

Morphologically, FNH is featured by vascular malformation, ductular overreaction, and fibrous infiltration. Previous studies of FNH mainly focus on single molecules that are cancer or angiogenesis related genes, such as the overexpression of GLUL affecting WNT/ β -catenin pathway (Rebouissou et al., 2008), the map-like distribution of glutamine synthetase (Bioulac Sage et al., 2009), and the upregulation of angiopoietin-1 (Gouw et al., 2010) and CD34 (Maillette DeBuy Wenniger et al., 2010). However, because of hypothesis-based research design and limited molecular profiling methods, clinical concerns mentioned above are not well resolved. Therefore, a data-driven approach that utilizes the state-of-art big data to deduce the molecular features of FNH, their correlation with biological behavior and association with the clinical outcome may complement current understanding of FNH initiation and progression.

Next-generation sequencing has recently been widely utilized to characterize morphologically normal tissues and benign tumors, unveiling their novel pathogenesis and molecular alterations (Kakiuchi and

¹Department of Liver Surgery and Transplantation, Liver Cancer Institute, Zhongshan Hospital, and Key Laboratory of Carcinogenesis and Cancer Invasion (Ministry of Education), Fudan University, Shanghai 200032, China

²Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Collaborative Innovation Center for Brain Science, Shanghai Jiao Tong University, Shanghai 200030, China

³Key Laboratory of Molecular Virology & Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences, Shanghai 200031, China

⁴Key Laboratory of Medical Epigenetics and Metabolism, Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China

⁵State Key Laboratory of Genetic Engineering, Fudan University, Shanghai 200433, China

⁶These authors contributed equally

⁷Lead contact

*Correspondence: shiyongyong@gmail.com (Y.S.), gao.qiang@zs-hospital.sh.cn (Q.G.)

<https://doi.org/10.1016/j.isci.2022.104921>



Ogawa, 2021; Tomasetti, 2019). The mutational landscape of the normal esophageal epithelium (Colom et al., 2020), liver (Brunner et al., 2019; Zhu et al., 2019), and gastrointestinal tract (Li et al., 2021; Lee-Six et al., 2019) have been established to illustrate the dynamic clonal expansion that provides survival advantage. Likewise, it has been revealed that renal angiomyolipoma harbors TSC2 mutations and copy-neutral loss of heterozygosity (Idogawa et al., 2020), and uterine leiomyoma carries chromothripsis that is associated with aggressive cancers (Mehine et al., 2013), indicating that “benign tumors” may not always be genetically “faithful”. As a comprehensive approach, multi-omics studies have uncovered the mystery from premalignant lesions to early stages of cancer, including pre-invasive lung cancer (Teixeira et al., 2019), familial adenomatous polyposis (Li et al., 2020), and hepatocarcinogenesis (Jee et al., 2019), which link genotype with phenotype, offering valuable evidence for precise medication.

Herein, we applied whole exome sequencing (WES), bulk RNA sequencing (RNA-seq), and single-cell RNA sequencing (scRNA-seq) on 22 paired FNH and adjacent normal liver tissues (aNL), followed by multi-level validation. We discovered that most FNH were genetically stable with specific transcriptomic modules. Of importance, FNH specific SOST-expressing vascular endothelial cells (ECs) and PDGFRB⁺ fibroblasts contributing to fibrogenesis in FNH were identified and validated. Surprisingly, some FNH that had atypical molecular characteristics were also identified. Overall, the integrated analyses performed in this study established the genetic and molecular landscape of FNH, and indicated preliminary clues for clinical decision making.

RESULTS

Patient clinical characteristics

Clinical information of all 22 FNH patients was summarized in Table S1. Among them, 19 cases were used for WES and RNA-seq, and 3 cases were applied for scRNA-seq. Eleven male and eight female patients were included, with no one having background liver disease. 68.4% of patients were between 20 and 40 years old, and age was comparable between two genders. Likewise, tumor sizes between the two gender groups were not statistically different (Figure S1A). None of the patients received any chemotherapy, radiotherapy, interventional therapy or drug therapy before surgery.

Individually, P11 simultaneously suffered from HCC in right lobe and FNH in left lobe (Figure S1B). Considering the genetic abnormality of FNH of P07 and P11 (described below), the morphological features of them were confirmed again. For P11, we stained GS, CK19, AFP and ki-67, and typical map-like GS distribution and bile duct hyperplasia were found (Figures S1C–S1F). We next re-checked the hematoxylin-eosin staining image of FNH in P07 with the highest number of somatic mutations, where an FNH-typical central scar and bile duct hyperplasia were seen (Figure S1G). All sequenced FNH samples had morphologically typical features with map-like GS distribution and bile duct hyperplasia.

Not all FNH show genomic integrity

The mean coverage was 99.5 and 99.2% for FNH and aNL, respectively; the depth of WES reached $154 \pm 34X$ (mean \pm SD) and $73 \pm 22X$ (mean \pm SD) for FNH and aNL under mean coverage, respectively. In total, we identified 773 somatic exonic mutations among all 19 patients (Table S2). Of them, 694 mutations were single nucleotide variation (SNV), and the rest 79 mutations were insertion or deletion (Indel). The majority of mutations were missense mutations (473 in total, 61.1%), followed by synonymous mutations (181 in total, 23.4%), and each remaining type was less than 5% (34 nonsense mutations, 29 non-frameshift deletions, 28 frameshift deletions, 17 frameshift insertions, 5 non-frameshift insertions, 5 unknown and 1 nonstop mutation in total) (Figure 1A). The variant allele frequency (VAF) of FNH was 0.132 ± 0.09 (mean \pm SD), indicating that mutations were only prevalent in a small proportion of cells within FNH.

The number of somatic mutations among patients varied significantly (2–220 mutations per patient, median 13 per patient), with patient P07 carrying the largest number of mutations and P04 having the fewest mutations (Figure 1B upper and Table S2). A total of 50 mutations were randomly selected for Sanger validation, reaching a success rate of 88% (Figure S2A and Table S3). We checked genes being repeatedly hit among patients, and the mutations that appeared in more than three patients, homozygous mutations, or multi-hit events among patients were not detected. However, some putative tumor suppressors or oncogenes were affected. Mutations of *ARID1B*, a probable tumor suppressor (Khurshheed et al., 2013), were detected in patients P07 and P14. *ARID1B* mutation (chr6: 157,522,487, G>T, V1574F) in P07 was predicted to be deleterious by different canonical algorithms. Likewise, *ARID1B* mutation (chr6: 157,488,191, G>T,

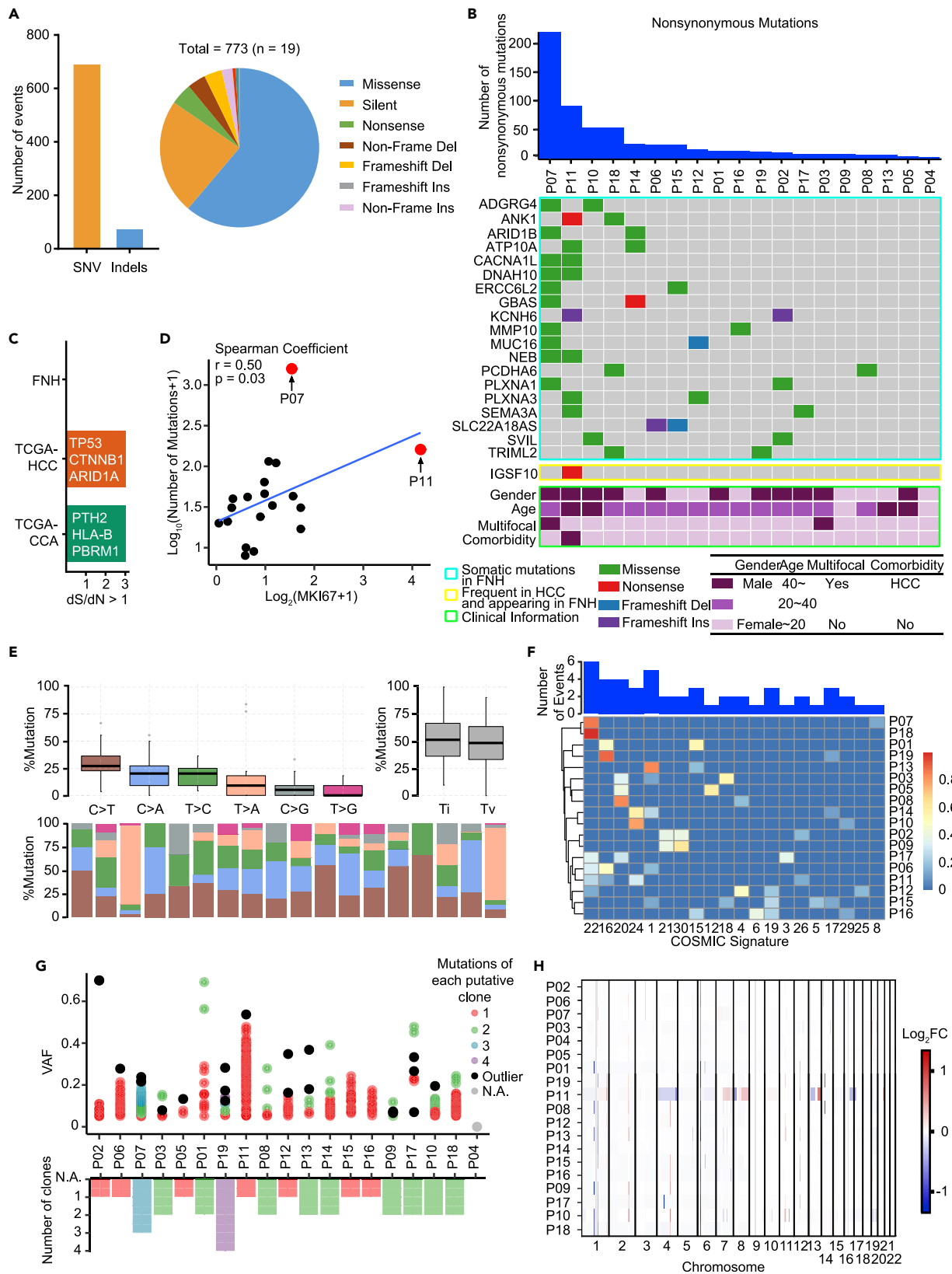


Figure 1. Mutational landscape of FNH

- (A) The constitution of somatic mutations in FNH (n = 19).
 (B) Individual mutational information (upper: the number of nonsynonymous mutations; middle: genes involved in mutation; bottom: basic clinical information).
 (C) dS/dN analysis of FNH, TCGA-HCC and TCGA-CCA.
 (D) The correlation between *MKI67* expression and the number of mutations of 19 FNH samples (Spearman's rank correlation coefficient).
 (E) The ratio of six-type base substitutions. The line and box represent median and upper and lower quartiles, respectively.
 (F) Mutational signatures of each FNH sample.
 (G) Number of clones inferred from mutational data. Each dot represents mutations belonging to each inferred clone.
 (H) CNVs of each FNH sample.
 See also [Figures S1](#) and [S2](#).

G953V) in P14 was also predicted deleterious. The mutations of *MMP10*, which functions as an oncogene ([Deraz et al., 2011](#)), were detected in patients P07 (chr11: 102,651,271, A>T, Y18N) and P16 (chr11: 102,650,265, G>A, P106L), and both mutations were predicted to cause functional damage ([Table S2](#)). Whether FNH shared the same driver mutations with HCC ([Brunner et al., 2019](#); [Fujimoto et al., 2012](#)) was explored, and only *IGSF10* was found mutated in P11 but not functionally harmful. Then, the mutational profiles were compared among FNH, normal liver ([Wijewardhane et al., 2021](#)), cirrhotic liver ([Zhu et al., 2019](#)) and HCC ([Brunner et al., 2019](#); [Fujimoto et al., 2012](#)), and FNH did not share any mutation with normal liver either, whereas FNH and cirrhotic liver only shared *ARID1B* mutation. Such recurrent mutations in cirrhotic liver, like *ARID1A* and *ARID1B*, are viewed to confer adaptive changes that promote fitness and regeneration in response to chronic damage, instead of malignant transformation ([Zhu et al., 2019](#)). Patient P11, who also suffered from HCC, and an elderly patient P10, harbored the second and the third highest number of mutations, indicating that mutations might preferably arise from diseased or aged liver. In addition, dS/dN ([Martincorena et al., 2017](#)) analysis was applied to FNH and TCGA hepatocellular carcinoma/cholangiocarcinoma (TCGA-HCC/CCA) datasets, and no positive mutational selection was found to present in FNH, whereas TCGA-HCC and TCGA-CCA both had positive selection ([Figure 1C](#)), supporting the notion that FNH did not harbor cancer driver mutations. Conclusively, despite that FNH was not driven by somatic gene mutations, high-mutation cases indeed exist, and some potential tumor suppressors and oncogenes were impaired and likely driving clonal expansion.

Mutations often occur during cell proliferation and division ([Cairns, 1975](#)). Thus, the relationship between the cell proliferation rate and the number of mutations was investigated. *MKI67* (gene encoding Ki-67) transcripts per million expression was explored as a cell proliferation marker, and a significant positive correlation between mutation load and *MKI67* expression level ($r = 0.50$, $p = 0.03$, [Figure 1D](#)) was identified. However, the correlation between *MKI67* expression and tumor size was not significant. This phenomenon indicates that mutations of FNH may have an intrinsic connection with cell proliferation, and close follow-up should be applied to FNH with high proliferative capability, regardless of tumor size. These results are consistent with the clinical observation that progressive FNH was not always large lesions ([Tajiri et al., 2014](#); [Kudo et al., 2008](#)).

The mutational spectrum and signatures of FNH were also explored. FNH had a nearly equal number of transition and transversion, where C to T was the predominant type of base substitution (>25% of total mutations), followed by C to A and T to C ([Figure 1E](#)), which is similar to cirrhotic liver ([Zhu et al., 2019](#)). Of them, C to T and T to C were also the main mutational features of HCC and CCA ([Alexandrov et al., 2013](#); [Dong et al., 2018](#)), indicating their commonality in liver tumors. The mutational pattern was deconstructed and compared to COSMIC signature database and signatures with weight ≥ 0.1 were thought significant ([Dong et al., 2018](#)). Liver cancers commonly show signature 16 ([Brunner et al., 2019](#)), and four FNH cases showed signature 16 as well. Poison contact ([Henry et al., 1999](#)) or defective DNA mismatch repair function can contribute to the pathogenesis of liver cancers ([Zekri et al., 2005](#)). Similarly, signatures 1, 22, and 24, which indicate potential exposure to poisons appeared in 8 of 19 FNH samples; four lesions showed signature 20 that implied defective DNA mismatch repair function. Nevertheless, different from cirrhotic liver and HCC ([Brunner et al., 2019](#)), FNH did not show signature 5, indicating that age may not dominantly account for mutational accumulation in FNH ([Figure 1F](#)).

The clonal status of FNH was deduced using VAF ([Figure 1G](#)). Except for P04 having too few mutations, other samples were successfully analyzed for the number of clones. In total, 7 of the 18 FNH lesions were monoclonal, and the remaining 11 were polyclonal, which was consistent with a previous study that

most FNH lesions were polyclonal origin (Paradis et al., 1997). In contrast, the monoclonal origin was well recognized in malignant tumors like HCC and CCA (Cai et al., 2009). Hence, we suspected that the monoclonal FNH of P11, which has the second highest mutational load and the highest proliferative rate, showed partial molecular features akin to HCC. Thus, AFP and Ki-67 were stained for this lesion. Enlarged nucleus, disappeared hepatic plates, and positive AFP or Ki-67 staining were observed in many cells within this case (Figures S1E and S1F), indicating pathologically cancer-like morphology.

Tumor mutational burden (TMB) of FNH was compared to TCGA tumors, and its TMB ranked 32 in all 34 types of tumors (Figure S2B). Among malignant tumors, kidney chromophobe, testicular germ cell tumors, uveal melanoma, acute myeloid leukemia and thyroid carcinoma did not show significantly different TMB with FNH. All TCGA benign tumors including thymoma and pheochromocytoma and paraganglioma did not show significantly different TMB with FNH as well. This result indicates that FNH has relatively low TMB, which is in agreement with its benign attribute. Each mutation was then substituted with its corresponding VAF, entitled VAF-TMB (Figure S2C). Of interest, except for thyroid carcinoma, the other four malignancies mentioned above had significantly higher VAF-TMB than FNH, whereas there was no statistical difference in VAF-TMB between two benign tumors and FNH. Of interest, many of these features are also observed in liver cirrhosis (Zhu et al., 2019). When only considering VAF >5%, FNH had a mean of 10.9% and a median of 9.3% VAF (excluding P11), whereas cirrhotic liver has a mean of 10.5% and a median of 8.7% VAF, indicating that mutations may affect same proportion of cells in two diseases. These results again indicate that FNH, as a benign tumor, has fewer cells affected by mutations than malignant tumors.

Copy number variations of FNH

Neither significantly recurrent arm-level copy number variations (CNVs), nor significantly recurrent focal level CNVs in FNH (Figure 1H) were found. However, significant arm-level CNVs of P11 were detected in chromosomes 4q, 7q, 8p, 8q, 16q, and 17p, which were also commonly detected in HCC (Qin et al., 1999). CNVs were confirmed by inspecting the gene expression level in corresponding regions (Figures S2D). Tumor suppressor genes such as *TP53* and *FAT4* were included in these arm-level regions but no obvious downregulation of these genes was observed, possibly because the functions of these genes were compensated by another intact copy.

The most prominent focal level CNV (copy number = 3) was 113,640,020–114,353,828 at 13q34 including *TFDP1* and *CDC16* with significant upregulation of their expression levels (Figures S2E and S2F). The copy number of *TFDP1* and *CDC16* at 13q34 was frequently amplified to enhance proliferative capability in HCC (Yasui et al., 2002). Altogether, although most of FNH samples did not exhibit common CNVs, P11 had already got features of HCC from both arm-level and focal level CNVs. However, we cannot conclude that FNH has the potential to undergo malignant transformation based on one case, and future integrative analysis enrolling more patients like P11 is needed.

Transcriptomic features of FNH

Principal component analysis (PCA) of the transcriptomic profile revealed pronounced spatial distinction between FNH and aNL (Figure S3A). Next, differentially expressed gene (DEG) analysis between FNH and aNL identified 589 significantly upregulated and 236 significantly downregulated genes. Among significantly upregulated genes, some fibrosis-related genes were found, such as *PDGFB*, *PDGFRB* (Czochra et al., 2006), *CXCL6* (Cai et al., 2018), *NOTCH3* (Chen et al., 2012), and *COL1A1* (Figure 2A). Functional enrichment analysis showed the significant upregulation of extracellular matrix (ECM) production and PI3K/AKT pathway (Figure 2B). Also, the PDGF pathway was significantly enriched, again implying its contribution to the fibroblastic process. Meanwhile, Ingenuity Pathway Analysis (IPA) unearthed possible cell types that were responsible for transcriptomic alterations. The most significant canonical pathways were hepatic stellate cell activation (Figure S3B) and PDGFB/PDGFRB axis was the only detected input signal for stellate cell activation in FNH (Figure S3C). Thus, PDGFB/PDGFRB-PI3K/AKT was the most possible pathway that may cause fibrosis in FNH.

FNH transcriptome was compared with TCGA-HCC, TCGA-CCA and GTEx normal liver using PCA. As expected, FNH, aNL and GTEx normal liver were spatially adjacent, whereas TCGA-HCC and TCGA-CCA were located oppositely (Figure 2C). This result preliminarily indicates the benign attribute of FNH. Surprisingly, we found P11 was away from other FNH counterparts and spatially close to TCGA-HCC. Further, consensus clustering was utilized and five classes were confirmed referring to consensus cumulative

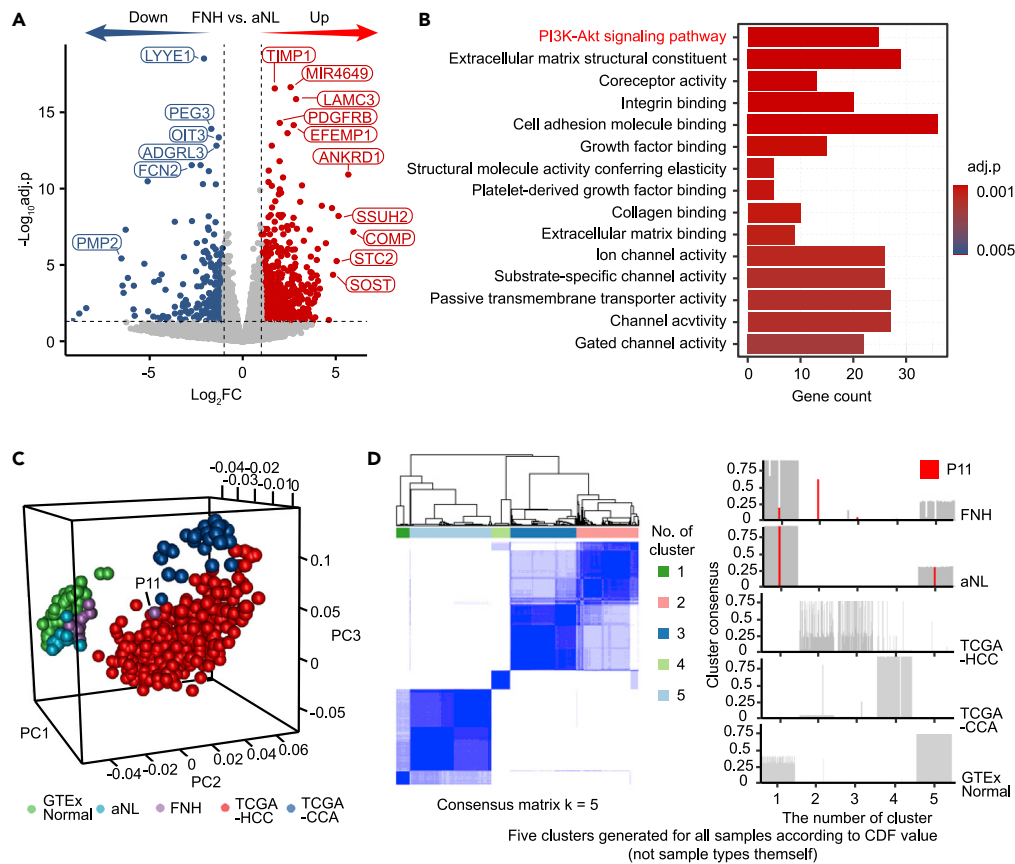


Figure 2. Transcriptomic features of FNH and co-analyzing with canonical database

(A) Volcano plot shows DEGs between FNH and aNL.

(B) Functional enrichment analysis of significantly upregulated genes in FNH.

(C) PCA of FNH, aNL, TCGA-HCC, TCGA-CCA and GTEx normal liver transcriptome.

(D) Consensus clustering of FNH, aNL, TCGA-HCC, TCGA-CCA and GTEx normal liver transcriptome ($k = 5$). Red highlights the position of FNH P11.

See also [Figure S3](#).

distribution function (CDF) value ([Figure 2D](#) left and [S3D](#)). Similarly, FNH, aNL and GTEx normal liver were clustered together mostly in clusters 1 and 5, whereas TCGA-HCC was in clusters 2 or 3 and TCGA-CCA was in cluster 4. Once again, P11 was clustered in clusters 1, 2 and 3, which suggests its similarity to TCGA-HCC. As a comparison, aNL of P11 was still clustered in clusters 1 and 5 as other NLs ([Figure 2D](#) right).

FNH has specific transcriptomic modules

The transcriptomic data of five types of samples (FNH, aNL, TCGA-HCC, TCGA-CCA and GTEx normal) were decomposed using independent component analysis (ICA). In total, 35 independent components (ICs) reached the stable status of the model followed by unsupervised clustering, and four modules were generated ([Figure 3A](#)). The point biserial correlation coefficients between meta-samples and binary vectors (e.g., all FNH samples were defined as 1 and other types of samples were defined as 0) were calculated as previously described ([Aynaoud et al., 2020](#)). It is worth noting that IC13 and IC7 in modules 3 and 4, respectively, highly correlated with FNH and aNL, whereas ICs such as IC2, IC11, IC17 and IC29 in module 2 correlated with malignancies ([Figures 3A](#) and [3B](#)).

Each IC was compared with BIODICA ([Captier et al., 2022](#)) ([Figure S3E](#)) or MSigDB ([Liberzon et al., 2011](#)) database or manually annotated. Importantly, cellular senescence was most significantly enriched in IC13, which suggest an intact “brake” function in response to proliferation ([Campisi and D’Adda, 2007](#)). Other features such as CTNNB1 signature ([Rebouissou et al., 2008](#)), estrogen-related gene expression ([Aldinger, 1977](#)) and NOTCH signaling in fibrogenesis ([Chen et al., 2012](#)) have been reported in FNH or liver

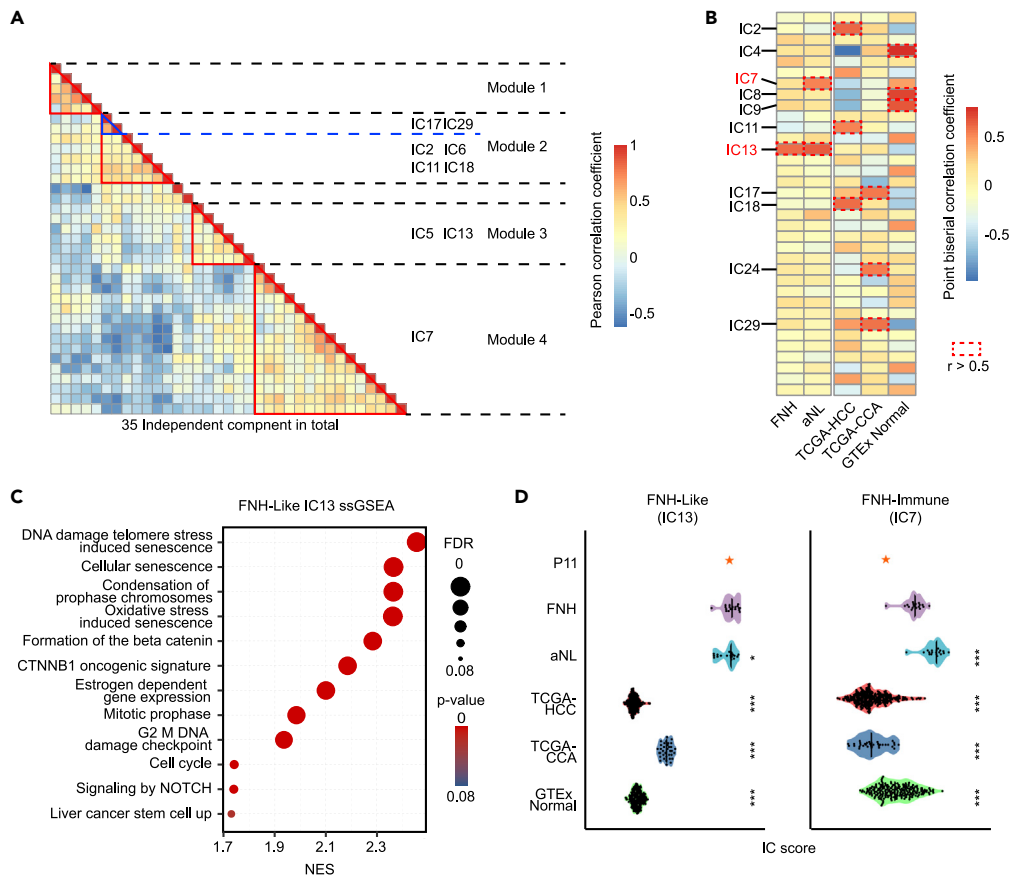


Figure 3. FNH owing its specific transcriptomic components

(A) ICA decomposes 35 independent components from FNH, aNL, TCGA-HCC, TCGA-CCA and GTEx normal liver transcriptome. Unsupervised clustering generates 4 basic transcriptomic modules (Pearson correlation coefficient). (B) IC7 and IC13 highly correlated with FNH or aNL (The point-biserial correlation coefficient). (C) ssGSEA on IC13 depicting many features of FNH. NES, normalized enrichment score. (D) Scoring each type of sample using ICs with biological significance (Mann–Whitney U test). The line represents median. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. See also [Figure S3](#).

fibrosis. Thus, IC13 was entitled FNH-like IC ([Figure 3C](#)). Likewise, IC7 was abundant in macrophages and endothelial signatures, thus named FNH-Immune ([Table S4](#)).

Next, each sample was scored using the leading genes in ICs that highly correlated with samples and had biological significance. As expected, FNH-Like score was significantly higher in FNH than in other types of samples, whereas the FNH-Immune score in FNH was significantly higher than other types of samples lower than aNL ([Figure 3D](#)). Other ICs also delineated other features of FNH, such as fibrosis, ECM generation, moderately elevated proliferative capacity and ductular proliferation ([Figure S3F](#)). Meanwhile, the FNH-Immune score of P11 FNH was significantly lower than P11 aNL, with significantly elevated proliferation score but FNH-Like score did not decrease. This result molecularly proved that lesion of P11 was FNH and further explained what alterations happened in P11 that made its transcriptome analogous to HCC mentioned above.

Single-cell landscape of FNH

To further explore FNH molecular characteristics and pathogenesis, three freshly dissected paired FNH and aNL were dissociated into single cell suspension, followed by flow cytometry sorting living cells and scRNA-seq. In total, 54,221 cells (including 27,474 cells from FNH, 26,747 cells from aNL) passed quality control, with 27 clusters of cells generated and visualized by uniform manifold approximation and

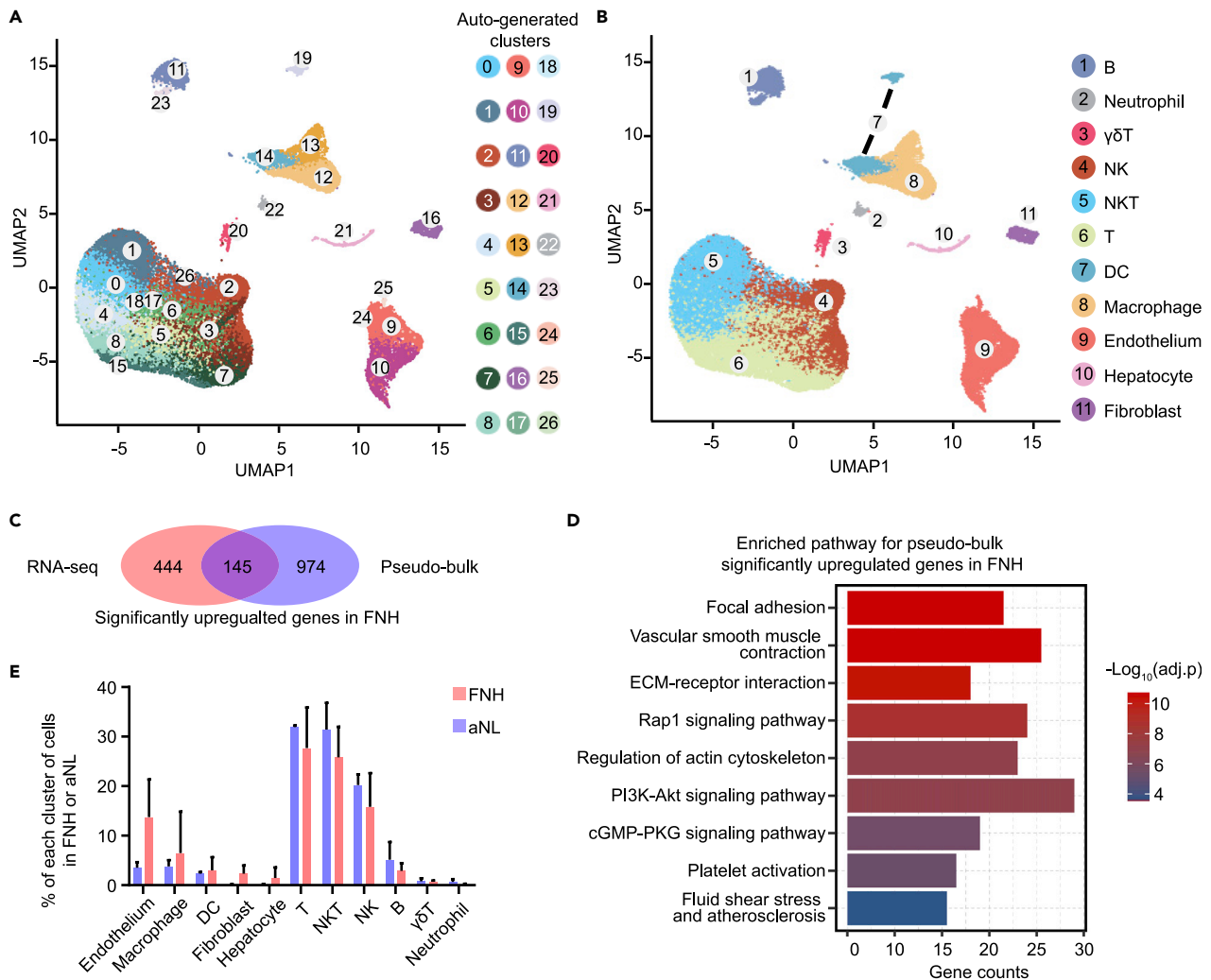


Figure 4. The single cell landscape of FNH

(A) The single-cell landscape of FNH. 26 clusters were automatically generated under the resolution of 1.2.

(B) Annotations of cells in each cluster.

(C) Venn graph showing 145 intersected significantly upregulated genes in FNH between RNA-seq and pseudo-bulk DEG analysis.

(D) Pathway enrichment analysis for significantly upregulated genes in pseudo-bulk FNH samples.

(E) The proportion of different type of cells in FNH and aNL (paired FNH and aNL, n = 3). Data are represented as mean \pm SD.

See also [Figure S4](#).

projection (UMAP) ([Figure 4A](#)). Each cluster was then annotated automatically and verified manually, resulting in 11 major cell types ([Figures 4B and S4A](#)).

To evaluate whether the main features captured by scRNA-seq, three paired pseudo-bulk samples were generated using the scRNA-seq data, followed by DEG analysis. Among significantly upregulated genes in FNH detected by bulk RNA-seq and pseudo-bulk DEG analyses, 145 intersected genes were discovered ([Figure 4C](#)). Many essential genes that related to FNH intrinsic features, such as *CXCL6*, *PDGFB*, *PDGFRB*, and *NOTCH3*, were all included. At single cell level, *CXCL6* was found mainly in hepatocytes, *PDGFB* in endothelial cells, *PDGFRB* and *NOTCH3* in fibroblasts ([Figure S4B](#)), supporting the IPA results ([Figure S3C](#)). Also, the pathways enriched in pseudo-bulk samples showed a high degree of consistency with those in bulk analysis, again indicating the enhanced ECM interaction, the possible involvement of PI3K-AKT signaling and underlying vascular changes in FNH ([Figure 4D](#)).

The proportions of each type of cells were next analyzed to explore FNH specific components (Figure 4E). Because of the limited number of fibroblasts captured by scRNA-seq (577 cells from FNH and 20 cells from aNL), we were only able to analyze as a whole. Even so, ECM deposition related genes, such as *COL4A1* and *COL4A2*, were observed significantly upregulated in the fibroblasts of FNH compared to aNL. Many ECM formation or interaction pathways were enriched in the fibroblasts of FNH as well. FNH had significantly more endothelial cells compared to aNL, which was consistent with the pathological feature of vascular proliferation. Among immune cells, macrophages dominated FNH with an abundance of over 5%, which was proved by the deconvolution of bulk RNA-seq data. Importantly, FNH-Immune consisted of genes mostly representing macrophage, which conspicuously distinguished FNH from other types of samples. Therefore, macrophages and endothelial cells that were highly enriched in FNH were selected for further analysis.

Kupffer cells are abundant in FNH

In total, 2,817 macrophages (1,987 from FNH and 830 from aNL) with three clusters were identified (Figure 5A). Macrophage cluster 1 (M ϕ _1) expressed a high level of *CD206*, *CD163* and *MARCO*, showing the identity of the non-inflammatory Kupffer cell (MacParland et al., 2018). M ϕ _2 highly expressed *S100A8/9* and *EREG* (Figure 5B), indicative of the role of pro-inflammation (MacParland et al., 2018; Cao et al., 2019). The proportion of M ϕ _1 and M ϕ _2 significantly differed between FNH and aNL, with FNH having more Kupffer cells whereas aNL owning more pro-inflammatory macrophages (Figure 5C). Co-staining of CD68 and CD206 in additional six samples confirmed the results as well (Figure 5D).

To further investigate the distribution of M ϕ _1 MARCO⁺ macrophages, the FNH scRNA-seq data from this study were combined with scRNA-seq from an in-house HCC dataset and a normal liver dataset in a previous study (MacParland et al., 2018). MARCO was hardly detectable in macrophages from HCC, consistent with the previous study (Sun et al., 2017), whereas it showed a relatively high expression in normal liver datasets, FNH and aNL (Figures 5E and 5F). Of note, in P11, FNH rather than corresponding aNL, showed a nearly negative signal of MARCO, similar to that in HCC, whereas the remaining 18 FNH and paired aNL showed obvious staining of MARCO (Figure 5G). Clinically, superparamagnetic iron oxide contrast is able to be taken by Kupffer cells in liver, thus Kupffer-cell-poor lesions can be differentiated (Tanaka et al., 1996). Therefore, superparamagnetic iron oxide MR contrast may be used to monitor potential molecular transformation based on the abundance of Kupffer cells in certain FNH non-invasively.

FNH had a unique type of SOST⁺ endothelial cells

Previous morphological studies assumed that abnormal blood flow might fuel hyperplasia (Rebouissou et al., 2008), driving us to further ask whether FNH had unique endothelial features. A total of 4,194 endothelial cells (3,561 from FNH and 633 from aNL) were identified with 10 clusters generated (Figure 6A). All ECs cells of FNH and aNL detected were positive for vascular markers *PECAM1* (CD31) and *FLT1* (Hirakawa et al., 2003) but negative for lymphatic EC marker *PDPN* (Figure S5A) (Amatschek et al., 2007).

We next explored FNH-enriched endothelial clusters and their specific genes. Compared to aNL, the most prominent differences concentrated on cluster 4 and cluster 6 (Figure 6B) that expressed both venous marker *COUP-TFII* and arterial marker *DLL4* but negative for sinus marker *MRC1* (Figure S5B) (deHaan et al., 2020; Diez et al., 2007), indicating an ambiguous status between artery and vein. DEG analysis was performed between all ECs of FNH and aNL. The most upregulated genes (Figure 6C), such as *COL4A1*, *SOST*, *EDNRB*, were mostly expressed in clusters 4 and 6 (Figures S5C–S5E). The expression of *COL4A1* indicates an increased activity in ECM production and high *EDNRB* expression on ECs indicates an enhanced ability of proliferation and migration (Ziche et al., 1995). *SELE*, found in activated or proliferating endothelial cells (Nishiwaki et al., 2007), was the gene with the highest fold change and its distribution was relatively equal among clusters (Figure S5F), indicating the diffusive molecular alterations in ECs of FNH. Then, the relationship between endothelial clusters was inferred using pseudo-trajectory analysis. Surprisingly, cluster 6 with *SOST* expression was mostly located at the middle and generated a new branch, whereas all ECs from aNL were found at the ends of the trajectory (Figure S5G). The result indicates that FNH had a unique type of ECs expressing *SOST*, independent of any ECs in aNL. To corroborate that *SOST*⁺ ECs were FNH specific, the intersection of FNH upregulated genes in RNA-seq, scRNA-seq and marker genes of cluster 6 were acquired, and *SOST* was within it (Figure S6A). Moreover, other types of liver samples, including fetal liver, normal liver, cirrhotic liver, HCC and CCA, were all negative for *SOST*

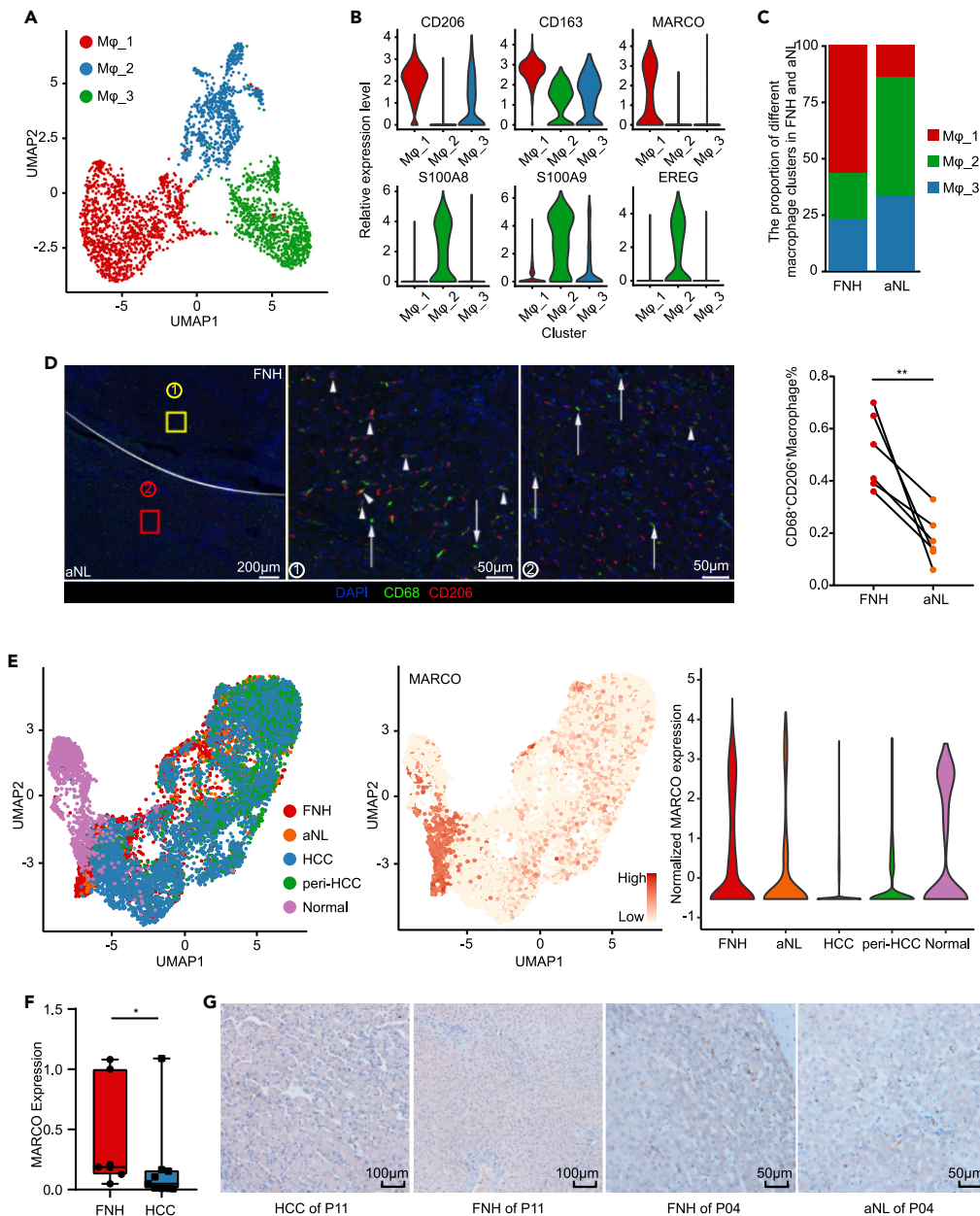


Figure 5. The features of macrophages in FNH

(A) UMAP showing three main types of macrophages in FNH.
 (B) The marker genes of each clusters.
 (C) The relative ratio of each type of macrophages in FNH and aNL.
 (D) IHC confirming the constitutional differences of CD68⁺CD206⁺ macrophages between FNH and aNL (paired FNH and aNL, n = 6; Paired Student's t test).
 (E) Single-cell data exhibiting a decrease of MARCO⁺ macrophages in HCC.
 (F) qRT-PCR data indicating the decreased level of MARCO in HCC (FNH, n = 7; HCC, n = 9; Mann-Whitney U test). The line and box represent median and upper and lower quartiles, respectively.
 (G) IHC confirming significant fewer MARCO⁺ macrophages in HCC of P11 and FNH of P11 compared to aNL of P11.
 *, p < 0.05; **, p < 0.01; ***, p < 0.001.

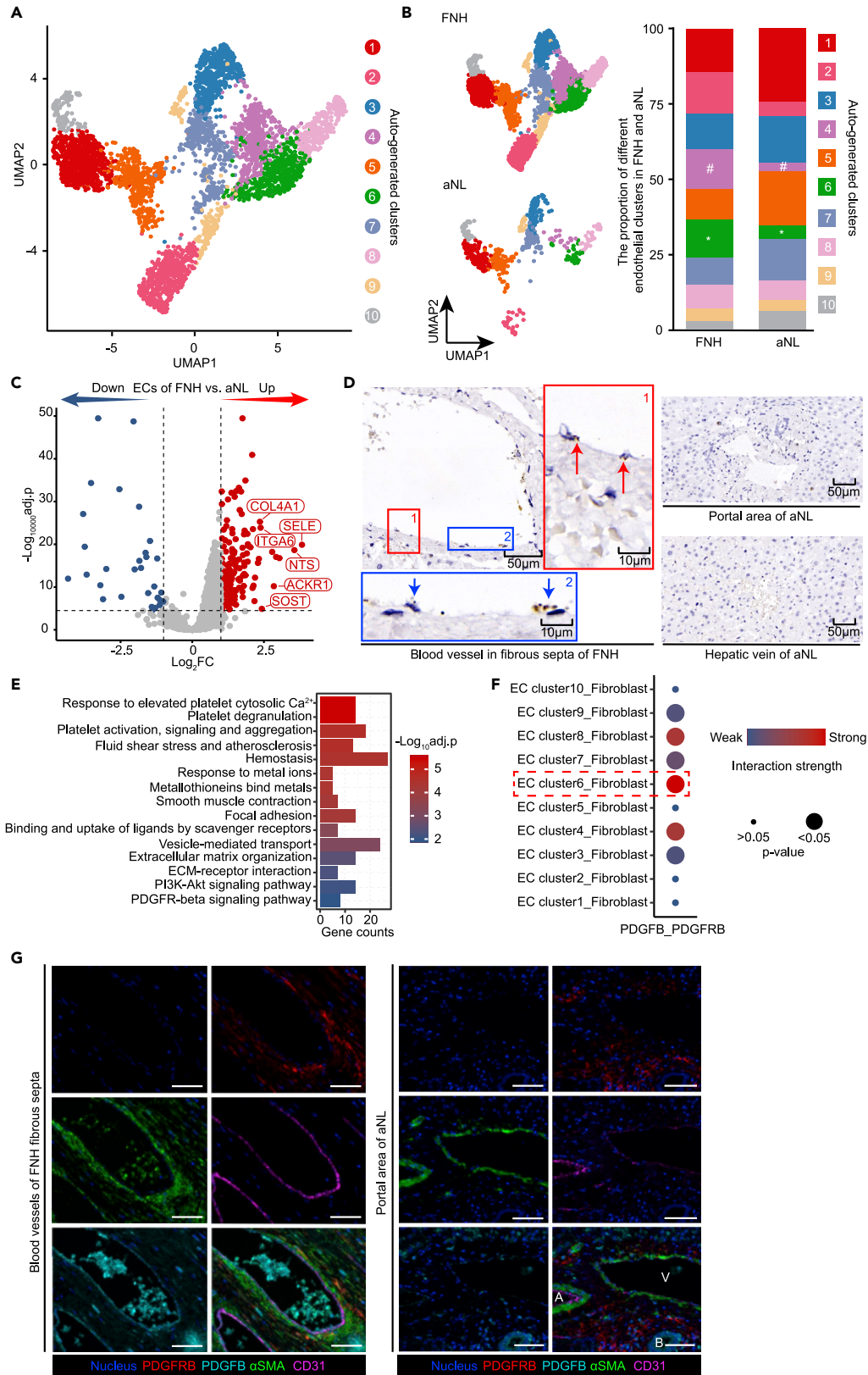


Figure 6. ECs of FNH probably promoting fibrogenesis and showing an SOST⁺ subtype

- (A) The landscape of ECs in FNH.
 (B) Ratios of all EC clusters in FNH and aNL. # and * denotes cluster 4 and 6, respectively.
 (C) DEGs between ECs of FNH and aNL at single cell level.
 (D) RNA ISH proving the expression of SOST in ECs of blood vessels in fibrous septa in FNH but not in any structure of aNL.
 (E) Functional enrichment analysis of significantly upregulated genes of ECs in FNH.
 (F) Cell-cell interaction analysis indicating ECs in cluster 6 having the strongest interaction with fibroblasts.
 (G) Representative mIF showing PDGF-B^{high} vascular ECs are spatially adjacent to PDGFRB⁺ fibroblasts in fibrous septa of FNH; whereas no PDGF-B^{high} ECs in aNL nor PDGFRB⁺ fibroblasts closely around vascular ECs in aNL. Scale bar, 50µm. A, artery; B, bile duct; V, vein.
 See also [Figures S5–S7](#).

expression ([Figures S6B–S6F](#)). The expression of SOST in vascular ECs in fibrous septa of FNH but not in aNL was further confirmed by RNA *in situ* hybridization ([Figure 6D](#)).

The functional enrichment analysis was performed at the whole EC level between FNH and aNL. The ECs of FNH showed significant enrichment of ECM associated pathways, PDGFBR signaling pathway and fluid shear stress pathway. Then, we re-analyzed the data excluding clusters 4 and 6, finding that ECM associated and PDGFBR signaling pathways were not enriched any more whereas fluid shear pathway was still enriched ([Figure 6E](#)). This result indicates that ECs of FNH may be broadly influenced by abnormal blood flow, and FNH specific ECs of clusters 4 and 6 are probably the main contributors of ECM deposition. In addition, cluster 4 exhibited a stronger function of vasculogenesis, whereas cluster 6 was more active in PDGFRB signaling, but the crosstalk between ECs and immune cells decreased ([Figure S6G](#)). Moreover, metabolic signature analysis figured out that cluster 6 had stronger oxidative phosphorylation, citric acid cycle and glutathione metabolism but lower arginine metabolism, which suggests a more active functional and metabolic status but impaired endothelial protection function ([Figure S6H](#)) ([Zhang et al., 2006](#); [Prasad et al., 1999](#); [Jongkind et al., 1989](#)). Together, FNH had a specific type of SOST⁺ vascular ECs with higher ECM related activity and higher metabolic activity.

Considering the involvement of PDGFRB signaling in FNH specific ECs, cell-cell interaction analysis was then performed. There were six clusters of ECs showing interaction with PDGFRB⁺ fibroblasts, among which cluster 6 had the strongest interaction ([Figure 6F](#)). Next, multiplex immunostaining of PDGFB, PDGFRB, αSMA, and CD31 in four paired FNH and aNL showed strong PDGFB intensity in vascular ECs from FNH instead of aNL. Correspondingly, PDGFRB⁺ fibroblasts were found spatially proximal to PDGFB⁺ ECs in FNH but not in aNL as well ([Figures 6G and S7A–S7C](#)), highlighting that PDGFB⁺ ECs could activate fibroblast and promote fibrosis through spatially proximal PDGFB/PDGFRB pair ([Czochra et al., 2006](#)). In addition, the two cells within 4 µm were thought actively interacted ([Sheng et al., 2021](#)). Therefore, ECs within FNH may promote fibrosis through upregulating ECM pathways and activating fibroblasts by producing PDGFB.

DISCUSSION

FNH is the second most common benign hepatic tumor, the prevalence of which is reported between 0.4 and 3% ([Maillette DeBuy Wenniger et al., 2010](#)). Transcriptomic signature related to FNH pathogenesis has been reported ([Rebouissou et al., 2008](#)), but the existence and relevance of the genomic and micro-environmental alterations remain less clear. Herein, we delineated the multi-omics characteristics of FNH in the scope of genome, transcriptome and single cell transcriptome and revealed that, despite mutations in tumor suppressors or oncogenes, FNH possesses a relatively stable genome with specific transcriptomic features. Combined with single cell data, we found FNH harbored abundant MARCO⁺ Kupffer cells and a unique type of SOST⁺ ECs that may contribute to the fibrotic process through PDGFB/PDGFRB axis. Interestingly, an atypical FNH was identified that gained an extra copy of 13q34 and excessive proliferative ability. Our work advanced the comprehensive understanding of FNH, which may further improve the clinical management of FNH.

We observed no cancer driver events in FNH at genomic level, which is consistent with the previously reported results that the well-known driver genes in HCC do not appear in FNH ([Cai et al., 2009](#)). However, lesions with a high number of mutations and high proliferative rate were discovered. Despite that only passenger mutations were found, the mutational signatures still indicated impaired DNA repair functions in FNH. Though the mutational profiles between FNH and liver cirrhosis are generally different, FNH indeed

shares some genomic features with liver cirrhosis (Zhu et al., 2019), indicating that genomic events are confined to a fraction of hepatocytes restricted by hyperplastic nodes. Thus, we recommended that FNH especially atypical FNH with a high proliferative rate should undergo surgical resection in time, and proliferation markers such as ki-67 should be stained routinely.

Our RNA-seq and scRNA-seq data ascertain the fibrogenic nature of FNH, in agreement with its morphological feature (Rebouissou et al., 2008). We also revealed that EC-fibroblast interaction may be driven by PDGFB/PDGFRB interaction through PI3K-AKT pathway. One mechanism for fibrosis in FNH can be summarized as ECs that under abnormal blood flow secrete PDGFB, by which ECs activate and recruit hepatic stellate cells to form fibroblasts, and further *in vivo* study is needed for validation. Notably, PDGFB/PDGFRB-PI3K/AKT was the only enriched pathway in FNH instead of famous TFGFB1/TGFBR1 signaling, which differs from corresponding pathways reported in hepatitis-B-related liver fibrosis or alcohol-associated liver cirrhosis (Kisseleva and Brenner, 2021). As for the source of fibroblasts, other than stellate cell formed fibroblasts, portal fibroblasts are another possible source of ECM producing myofibroblast. However, portal fibroblasts cannot be driven by PDGF (Kisseleva and Brenner, 2021; Wells et al., 2004). These differences suggest the driving factor of FNH is different from those hepatocyte-injured diseases.

Interestingly, FNH specific ECs of cluster 6, showing the strongest interaction with fibroblasts, also uniquely express *SOST*. *SOST* was initially discovered in bone and elevated while encountering mechanical stimulation (Robling et al., 2008). Studies about *SOST* and ECs are scarce. So far as we know, only one study reported that sclerostin increased the proliferation of human umbilical vein ECs *in vitro* (Oranger et al., 2016). Nevertheless, which factors drive the expression of *SOST* in FNH and what is the exact biological significance are still elusive.

P11 is the most special among all samples sequenced. FNH of P11 has the second largest number of mutations that were monoclonal origin, but there are no noxious mutations in famous oncogenes or tumor suppressors. However, CNVs in FNH of P11 suggests the similarity to early HCC. FNH of P11 preserves the main transcriptomic feature, but acquires extra proliferative capability. Essentially, MARCO⁺ macrophages decreasing is the main immune alteration in P11. Thus, Kupffer cell sensitive contrast agent may help non-invasively trace FNH or alarm atypical FNH. However, further in-depth studies are needed to confirm this conclusion.

In conclusion, by integrating WES, RNA-seq and scRNA-seq data, we revealed that most FNH are genetically stable but some molecularly atypical FNH still exist, discovered a new type of FNH specific ECs, and offered a probable mechanism of fibrogenesis in FNH. Our work advances the understanding of molecular pathogenesis of FNH and provides clues for clinical management.

Limitations of the study

Limitations of this study included the lack of cellular or animal models, which handicaps the study of mechanisms. Although WES yielded an average of 13% VAF mutations in FNH, it was possible that considerable somatic mutations with less VAF were left undetected. Ultra-high sequencing depth or single cell genome sequencing may better help reveal those hidden but important secrets.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Clinical sample acquisition
- METHOD DETAILS
 - Specimens processing
 - DNA extraction, library construction, WES and sanger validation
 - RNA extraction, library construction and RNA-seq
 - WES data processing, somatic mutation detection
 - RNA-seq data processing

- dN/dS analysis
- Mutational signature analysis
- Identification of CNVs
- Clone number estimation
- TMB calculation and comparison with TCGA datasets
- DEG detection and pathway analysis
- PCA of FNH, TCGA and GTEx transcriptome
- Consensus clustering of FNH, TCGA and GTEx samples
- ICA and correlation with five types of samples
- Annotation of FNH specific IC and samples scoring
- Single-cell suspension preparation
- Library construction and scRNA-sequencing
- scRNA sequencing data processing
- Unsupervised clustering and determination of cell type
- Deconvolution of RNA-seq data
- Pseudo-bulk sample generation
- Gene set enrichment of single-cell data
- Pseudo-trajectory inference
- Metabolic gene signature scoring
- Intercellular ligand-receptor analysis
- Immunofluorescence and immunohistochemistry
- RNA extraction and qRT-PCR
- RNA *in situ* hybridization
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.104921>.

ACKNOWLEDGMENTS

We thank Dr. Yumeng Sun and Dr. Lu Meng, from Institute Pasteur of Shanghai, Chinese Academy of Sciences for technical assistance. We thank Ms. Yifei Cheng from Department of Liver Surgery and Transplantation, Liver Cancer Institute, Zhongshan Hospital for the important support of written English. We thank all members of the laboratory for their insightful discussions. We would also like to acknowledge the funding of the National Natural Science Foundation of China (Nos. 81961128025, 82121002, 32070679, U1804284, 81871055), Program of Shanghai Academic Research Leader (No. 19XD1420700), the National Key R&D Program of China (2021YFC2702100, 2019YFA0905400, 2017YFC0908105), Shanghai Municipal Science and Technology Major Project (2017SHZDZX01), Taishan Scholar Program of Shandong Province (tsqn201812153), Natural Science Foundation of Shandong Province (ZR2019YQ14) and Zhupei Foundation of Zhongshan Hospital (Zhupei 2021-001).

AUTHOR CONTRIBUTIONS

Q.G. and Y.-Y.S. conceived and directed the project. Y.-M.L., L.-Q.D., and D.-N.R. collected surgical samples and clinical data. J.-M.Z., K.W., and Y.-M.L. wrote the code for analysis and Y.-C.W. and Q.G. were responsible for verification. J.-Q.M. designed immunohistochemistry panel. Y.-M.L. performed experiments and M.Z., Y.-P.L., and Z.-J.Y. provided experimental support. J.F. and Z.-B.D. provided important intellectual input. Y.-M.L. and J.-M.Z. wrote the manuscript with the assistance of Z.W. and X.-M.Z. All authors critically revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 27, 2022

Revised: July 17, 2022

Accepted: August 8, 2022

Published: September 16, 2022

REFERENCES

- Aizarani, N., Saviano, A., Sagar, M., Maily, L., Durand, S., Herman, J.S., Pessaux, P., Baumert, T.F., and Grün, D. (2019). A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* 572, 199–204. <https://doi.org/10.1038/s41586-019-1373-2>.
- Aldinger, K. (1977). Focal nodular hyperplasia of the liver associated with high-dosage estrogens. *Arch. Intern. Med.* 137, 357–359. <https://doi.org/10.1001/archinte.1977.03630150057016>.
- Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Borresen-Dale, A.L., et al. (2013). Signatures of mutational processes in human cancer. *Nature* 500, 415–421. <https://doi.org/10.1038/nature12477>.
- Amatschek, S., Kriehuber, E., Bauer, W., Reininger, B., Meraner, P., Wolpi, A., Schweifer, N., Haslinger, C., Stingl, G., and Maurer, D. (2007). Blood and lymphatic endothelial cell-specific differentiation programs are stringently controlled by the tissue environment. *Blood* 109, 4777–4785. <https://doi.org/10.1182/blood-2006-10-053280>.
- Aran, D., Hu, Z., and Butte, A.J. (2017). Xcell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18, 220. <https://doi.org/10.1186/s13059-017-1349-1>.
- Aran, D., Looney, A.P., Liu, L., Wu, E., Fong, V., Hsu, A., Chak, S., Naikawadi, R.P., Wolters, P.J., Abate, A.R., et al. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* 20, 163–172. <https://doi.org/10.1038/s41590-018-0276-y>.
- Aynaud, M.M., Mirabeau, O., Gruel, N., Grossetête, S., Boeva, V., Durand, S., Surdez, D., Saulnier, O., Zaïdi, S., Gribova, S., et al. (2020). Transcriptional programs define intratumoral heterogeneity of Ewing sarcoma at single-cell resolution. *Cell Rep.* 30, 1767–1779.e6. <https://doi.org/10.1016/j.celrep.2020.01.049>.
- Bioulac Sage, P., Laumonier, H., Rullier, A., Cubel, G., Laurent, C., Zucman Rossi, J., and Balabaud, C. (2009). Over-expression of glutamine synthetase in focal nodular hyperplasia: a novel easy diagnostic tool in surgical pathology. *Liver Int.* 29, 459–465. <https://doi.org/10.1111/j.1478-3231.2008.01849.x>.
- Bonizzoni, P., Della Vedova, G., Dondi, R., and Jiang, T. (2008). On the approximation of correlation clustering and consensus clustering. *J. Comput. Syst. Sci.* 74, 671–696. <https://doi.org/10.1016/j.jcss.2007.06.024>.
- Brunner, S.F., Roberts, N.D., Wylie, L.A., Moore, L., Aitken, S.J., Davies, S.E., Sanders, M.A., Ellis, P., Alder, C., Hooks, Y., et al. (2019). Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. *Nature* 574, 538–542. <https://doi.org/10.1038/s41586-019-1670-9>.
- Captier, N., Merlevede, J., Molkenov, A., Seisenova, A., Zhubanchaliyev, A., Nazarov, P.V., Barillot, E., Kairo, U., and Zinovyev, A. (2022). BIODICA: a computational environment for independent component analysis. *Bioinformatics* 38, 2963–2964. <https://doi.org/10.1093/bioinformatics/btac204>.
- Chen, Y.X., Weng, Z.H., and Zhang, S.L. (2012). Notch3 regulates the activation of hepatic stellate cells. *World J. Gastroenterol.* 18, 1397–1403. <https://doi.org/10.3748/wjg.v18.i12.1397>.
- Campisi, J., and D’Adda, D.F.F. (2007). Cellular senescence: when bad things happen to good cells. *Nat. Rev. Mol. Cell Biol.* 8, 729–740. <https://doi.org/10.1038/nrm2233>.
- Cai, Y.R., Gong, L., Teng, X.Y., Zhang, H.T., Wang, C.F., Wei, G.L., Guo, L., Ding, F., Liu, Z.H., Pan, Q.J., and Su, Q. (2009). Clonality and allelotypic analyses of focal nodular hyperplasia compared with hepatocellular adenoma and carcinoma. *World J. Gastroenterol.* 15, 4695–4708. <https://doi.org/10.3748/wjg.15.4695>.
- Cairns, J. (1975). Mutation selection and the natural history of cancer. *Nature* 255, 197–200. <https://doi.org/10.1038/255197a0>.
- Chen, T.C., Chou, T.B., Ng, K.F., Hsieh, L.L., and Chou, Y.H. (2001). Hepatocellular carcinoma associated with focal nodular hyperplasia. *Virchows Arch.* 438, 408–411. <https://doi.org/10.1007/s004280000348>.
- Colom, B., Alcolea, M.P., Piedrafita, G., Hall, M.W.J., Wabik, A., Dentre, S.C., Fowler, J.C., Herms, A., King, C., Ong, S.H., et al. (2020). Spatial competition shapes the dynamic mutational landscape of normal esophageal epithelium. *Nat. Genet.* 52, 604–614. <https://doi.org/10.1038/s41588-020-0624-3>.
- Cao, W., Luo, L.L., Chen, W.W., Liang, L., Zhang, R.R., Zhao, Y.L., Chen, J., and Yue, J. (2019). Polymorphism in the EREG gene confers susceptibility to tuberculosis. *BMC Med. Genet.* 20, 7. <https://doi.org/10.1186/s12881-018-0729-z>.
- Cai, X., Li, Z., Zhang, Q., Qu, Y., Xu, M., Wan, X., and Lu, L. (2018). CXCL6-EGFR-induced Kupffer cells secrete TGF- β 1 promoting hepatic stellate cell activation via the SMAD2/BRD4/C-MYC/EZH2 pathway in liver fibrosis. *J. Cell Mol. Med.* 22, 5050–5061. <https://doi.org/10.1111/jcmm.13787>.
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Czochra, P., Klopčič, B., Meyer, E., Herkel, J., Garcia-Lazaro, J.F., Thieringer, F., Schirmacher, P., Biesterfeld, S., Galle, P.R., Lohse, A.W., and Kanzler, S. (2006). Liver fibrosis induced by hepatic overexpression of PDGF-B in transgenic mice. *J. Hepatol.* 45, 419–428. <https://doi.org/10.1016/j.jhep.2006.04.010>.
- Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* 10, giab008. <https://doi.org/10.1093/gigascience/giab008>.
- Deraz, E.M., Kudo, Y., Yoshida, M., Obayashi, M., Tsunematsu, T., Tani, H., Siriwardena, S.B.S.M., Keikhaee, M.R., Kiekhraee, M.R., Qi, G., et al. (2011). MMP-10/stromelysin-2 promotes invasion of head and neck cancer. *PLoS One* 6, e25438. <https://doi.org/10.1371/journal.pone.0025438>.
- Dong, L.Q., Shi, Y., Ma, L.J., Yang, L.X., Wang, X.Y., Zhang, S., Wang, Z.C., Duan, M., Zhang, Z., Liu, L.Z., et al. (2018). Spatial and temporal clonal evolution of intrahepatic cholangiocarcinoma. *J. Hepatol.* 69, 89–98. <https://doi.org/10.1016/j.jhep.2018.02.029>.
- de Haan, W., Øie, C., Benkheil, M., Dheedene, W., Vinckier, S., Coppello, G., Aranguren, X.L., Beerens, M., Jaekers, J., Topal, B., et al. (2020). Unraveling the transcriptional determinants of liver sinusoidal endothelial cell specialization. *Am. J. Physiol. Gastrointest. Liver Physiol.* 318, G803–G815. <https://doi.org/10.1152/ajpgi.00215.2019>.
- Diez, H., Fischer, A., Winkler, A., Hu, C.J., Hatzopoulos, A.K., Breier, G., and Gessler, M. (2007). Hypoxia-mediated activation of DLL4-NOTCH-HEY2 signaling in endothelial progenitor cells and adoption of arterial cell fate. *Exp. Cell Res.* 313, 1–9. <https://doi.org/10.1016/j.yexcr.2006.09.009>.
- EASL (2016). EASL clinical practice guidelines on the management of benign liver tumours. *J. Hepatol.* 65, 386–398.
- Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* 15, 1484–1506. <https://doi.org/10.1038/s41596-020-0292-x>.
- Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarizing analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. <https://doi.org/10.1016/j.bone.2017.03.001>.
- Fujimoto, A., Totoki, Y., Abe, T., Boroevich, K.A., Hosoda, F., Nguyen, H.H., Aoki, M., Hosono, N., Kubo, M., Miya, F., et al. (2012). Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat. Genet.* 44, 760–764. <https://doi.org/10.1038/ng.2291>.
- Gouw, A.S.H., Zeng, W., Buischool, M., Platteeuw, I., van den Heuvel, M.C., Poppema, S., de Jong, K.P., and Molema, G. (2010). Molecular characterization of the vascular features of focal nodular hyperplasia and hepatocellular adenoma: a role for angiotensin-1. *Hepatology* 52, 540–549. <https://doi.org/10.1002/hep.23700>.
- Hirakawa, S., Hong, Y.K., Harvey, N., Schacht, V., Matsuda, K., Libermann, T., and Detmar, M. (2003). Identification of vascular lineage-specific genes by transcriptional profiling of isolated blood vascular and lymphatic endothelial cells. *Am. J. Pathol.* 162, 575–586. [https://doi.org/10.1016/S0002-9440\(10\)63851-5](https://doi.org/10.1016/S0002-9440(10)63851-5).
- Henry, S.H., Bosch, F.X., Troxell, T.C., and Bolger, P.M. (1999). Public health: reducing liver cancer-global control of aflatoxin. *Science* 286, 2453–2454. <https://doi.org/10.1126/science.286.5449.2453>.
- Hyvärinen, A., and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Netw.* 13, 411–430. [https://doi.org/10.1016/S0893-6080\(00\)00026-5](https://doi.org/10.1016/S0893-6080(00)00026-5).

- Himberg, J., Hyvärinen, A., and Esposito, F. (2004). Validating the independent components of neuroimaging time series via clustering and visualization. *Neuroimage* 22, 1214–1222. <https://doi.org/10.1016/j.neuroimage.2004.03.027>.
- Idegawa, M., Hida, T., Tanaka, T., Ohira, N., Tange, S., Sasaki, Y., Uhara, H., Masumori, N., Tokino, T., and Natori, H. (2020). Renal angiomyolipoma harboring a missense mutation of TSC2 with copy-neutral loss of heterozygosity. *Cancer Biol. Ther.* 21, 315–319. <https://doi.org/10.1080/15384047.2019.1702406>.
- Jongkind, J.F., Verkerk, A., and Baggen, R.G. (1989). Glutathione metabolism of human vascular endothelial cells under peroxidative stress. *Free Radic. Biol. Med.* 7, 507–512. [https://doi.org/10.1016/S0735-1097\(99\)00216-8](https://doi.org/10.1016/S0735-1097(99)00216-8).
- Jee, B.A., Choi, J.H., Rhee, H., Yoon, S., Kwon, S.M., Nahm, J.H., Yoo, J.E., Jeon, Y., Choi, G.H., Woo, H.G., and Park, Y.N. (2019). Dynamics of genomic, epigenomic, and transcriptomic aberrations during stepwise hepatocarcinogenesis. *Cancer Res.* 79, 5500–5512. <https://doi.org/10.1158/0008-5472.CAN-19-0991>.
- Kakiuchi, N., and Ogawa, S. (2021). Clonal expansion in non-cancer tissues. *Nat. Rev. Cancer* 21, 239–256. <https://doi.org/10.1038/s41568-021-00335-3>.
- Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* 16, 1289–1296. <https://doi.org/10.1038/s41592-019-0619-0>.
- Koea, J.B., and Yeong, M.L. (2014). Focal nodular hyperplasia and hepatocellular carcinoma: uncommon companions? *Pathology* 46, 348–350. <https://doi.org/10.1097/PAT.000000000000102>.
- Kudo, M., Zheng, R.Q., Chung, H., Yonekura, T., Yagi, M., Nakashima, O., and Kojiro, M. (2008). Long-term follow-up of atypical progressive focal nodular hyperplasia increasing in size and number implicates its pathogenesis. *Am. J. Gastroenterol.* 103, 2153–2155. https://doi.org/10.1111/j.1572-0241.2008.01982_14.x.
- Kisseleva, T., and Brenner, D. (2021). Molecular and cellular mechanisms of liver fibrosis and its regression. *Nat. Rev. Gastroenterol. Hepatol.* 18, 151–166. <https://doi.org/10.1038/s41575-020-00372-7>.
- Khursheed, M., Kolla, J.N., Kotapalli, V., Gupta, N., Gowrishankar, S., Uppin, S.G., Sastry, R.A., Koganti, S., Sundaram, C., Pollack, J.R., and Bashyam, M.D. (2013). ARID1B, a member of the human SWI/SNF chromatin remodeling complex, exhibits tumour-suppressor activities in pancreatic cancer cell lines. *Br. J. Cancer* 108, 2056–2062. <https://doi.org/10.1038/bjc.2013.200>.
- Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. <https://doi.org/10.1038/s41587-019-0201-4>.
- Kosinski, M., and Biecek, P. (2021). RTCGA: the cancer genome atlas data integration. R Package Version 1.24.0. <https://rtcga.github.io/RTCGA>.
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. <https://doi.org/10.1093/bioinformatics/btr260>.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. <https://doi.org/10.1093/bioinformatics/btt656>.
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Lee-Six, H., Olafsson, S., Ellis, P., Osborne, R.J., Sanders, M.A., Moore, L., Georgakopoulos, N., Torrente, F., Noorani, A., Goddard, M., et al. (2019). The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* 574, 532–537. <https://doi.org/10.1038/s41586-019-1672-7>.
- Li, J., Wang, R., Zhou, X., Wang, W., Gao, S., Mao, Y., Wu, X., Guo, L., Liu, H., Wen, L., et al. (2020). Genomic and transcriptomic profiling of carcinogenesis in patients with familial adenomatous polyposis. *Gut* 69, 1283–1293. <https://doi.org/10.1136/gutjnl-2019-319438>.
- Li, R., Di, L., Li, J., Fan, W., Liu, Y., Guo, W., Liu, W., Liu, L., Li, Q., Chen, L., et al. (2021). A body map of somatic mutagenesis in morphologically normal human tissues. *Nature* 597, 398–403. <https://doi.org/10.1038/s41586-021-03836-1>.
- Luciani, A., Kobeiter, H., Maison, P., Cherqui, D., Zafrani, E.S., Dhumeaux, D., and Mathieu, D. (2002). Focal nodular hyperplasia of the liver in men: is presentation the same in men and women? *Gut* 50, 877–880. <https://doi.org/10.1136/gut.50.6.877>.
- Maillette De Buy Wenniger, L., Terpstra, V., and Beuers, U. (2010). Focal nodular hyperplasia and hepatic adenoma: epidemiology and pathology. *Dig. Surg.* 27, 24–31. <https://doi.org/10.1159/000268404>.
- MacParland, S.A., Liu, J.C., Ma, X.Z., Innes, B.T., Bartzack, A.M., Gage, B.K., Manuel, J., Khuu, N., Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat. Commun.* 9, 4383. <https://doi.org/10.1038/s41467-018-06318-7>.
- Mehine, M., Kaasinen, E., Mäkinen, N., Katainen, R., Kämpjärvi, K., Pitkänen, E., Heinonen, H.R., Bützow, R., Kilpivaara, O., Kuosmanen, A., et al. (2013). Characterization of uterine leiomyomas by whole-genome sequencing. *N. Engl. J. Med.* 369, 43–53. <https://doi.org/10.1056/NEJMoa1302736>.
- Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Van Loo, P., McLaren, S., Wedge, D.C., Fullam, A., Alexandrov, L.B., Tubio, J.M., et al. (2015). High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* 348, 880–886. <https://doi.org/10.1126/science.1226806>.
- Martincorena, I., Raine, K.M., Gerstung, M., Dawson, K.J., Haase, K., Van Loo, P., Davies, H., Stratton, M.R., and Campbell, P.J. (2017). Universal patterns of selection in cancer and somatic tissues. *Cell* 171, 1029–1041.e21. <https://doi.org/10.1016/j.cell.2017.09.042>.
- Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., and Koeffler, H.P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28, 1747–1756. <https://doi.org/10.1101/gr.239244.118>.
- Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34, 267–273. <https://doi.org/10.1038/ng1180>.
- Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37, 773–782. <https://doi.org/10.1038/s41587-019-0114-2>.
- Nahm, C.B., Ng, K., Lockie, P., Samra, J.S., and Hugh, T.J. (2011). Focal nodular hyperplasia—a review of myths and truths. *J. Gastrointest. Surg.* 15, 2275–2283. <https://doi.org/10.1007/s11605-011-1680-x>.
- Nishiwaki, Y., Yoshida, M., Iwaguro, H., Masuda, H., Nitta, N., Asahara, T., and Isobe, M. (2007). Endothelial E-selectin potentiates neovascularization via endothelial progenitor cell-dependent and -independent mechanisms. *Arterioscler. Thromb. Vasc. Biol.* 27, 512–518. <https://doi.org/10.1161/01.ATV.0000254812.23238.2b>.
- Oranger, A., Brunetti, G., Colaianni, G., Tamma, R., Carbone, C., Lippo, L., Mori, G., Pignataro, P., Cirulli, N., Zerlotin, R., et al. (2016). Sclerostin stimulates angiogenesis in human endothelial cells. *Bone* 101, 26–36.
- Prasad, A., Andrews, N.P., Padder, F.A., Husain, M., and Quyyumi, A.A. (1999). Glutathione reverses endothelial dysfunction and improves nitric oxide bioavailability. *J. Am. Coll. Cardiol.* 34, 507–514.
- Paradis, V., Laurent, A., Flejou, J.F., Vidaud, M., and Bedossa, P. (1997). Evidence for the polyclonal nature of focal nodular hyperplasia of the liver by the study of X-chromosome inactivation. *Hepatology* 26, 891–895. <https://doi.org/10.1002/hep.510260414>.
- Petsas, T., Tsamandas, A., Tsota, I., Karavias, D., Karatza, C., Vassiliou, V., and Kardamakis, D. (2006). A case of hepatocellular carcinoma arising within large focal nodular hyperplasia with review of the literature. *World J. Gastroenterol.* 12, 6567–6571. <https://doi.org/10.3748/wjg.v12.i40.6567>.
- Qin, L.X., Tang, Z.Y., Sham, J.S., Ma, Z.C., Ye, S.L., Zhou, X.D., Wu, Z.Q., Trent, J.M., and Guan, X.Y. (1999). The association of chromosome 8p deletion and tumor metastasis in human hepatocellular carcinoma. *Cancer Res.* 59, 5662–5665.
- Robling, A.G., Niziolek, P.J., Baldrige, L.A., Condon, K.W., Allen, M.R., Alam, I., Mantila, S.M.,

- Gluhak-Heinrich, J., Bellido, T.M., Harris, S.E., and Turner, C.H. (2008). Mechanical stimulation of bone in vivo reduces osteocyte expression of SOST/sclerostin. *J. Biol. Chem.* 283, 5866–5875. <https://doi.org/10.1074/jbc.M705092200>.
- Ramachandran, P., Dobie, R., Wilson-Kanamori, J.R., Dora, E.F., Henderson, B.E.P., Luu, N.T., Portman, J.R., Matchett, K.P., Brice, M., Marwick, J.A., et al. (2019). Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* 575, 512–518. <https://doi.org/10.1038/s41586-019-1631-3>.
- Rebouissou, S., Couchy, G., Libbrecht, L., Balabaud, C., Imbeaud, S., Auffray, C., Roskams, T., Bioulac-Sage, P., and Zucman-Rossi, J. (2008). The β -catenin pathway is activated in focal nodular hyperplasia but not in cirrhotic FNH-like nodules. *J. Hepatol.* 49, 61–71. <https://doi.org/10.1016/j.jhep.2008.03.013>.
- Rajput, A., Bocklage, T., Greenbaum, A., Lee, J.H., and Ness, S.A. (2017). Mutant-allele tumor heterogeneity scores correlate with risk of metastases in colon cancer. *Clin. Colorectal Cancer* 16, e165–e170. <https://doi.org/10.1016/j.clcc.2016.11.004>.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 1572–1573. <https://doi.org/10.1093/bioinformatics/btq170>.
- Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B.S., and Swanton, C. (2016). Deconstructsig: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* 17, 31. <https://doi.org/10.1186/s13059-016-0893-4>.
- Sharma, A., Seow, J.J.W., Dutertre, C., Pai, R., Blériot, C., Mishra, A., Wong, R.M.M., Singh, G.S.N., Sudhagar, S., Khalilnezhad, S., et al. (2020). Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell* 183, 377–394. <https://doi.org/10.1016/j.cell.2020.08.040>.
- Sheng, J., Zhang, J., Wang, L., Tano, V., Tang, J., Wang, X., Wu, J., Song, J., Zhao, Y., Rong, J., et al. (2021). Topological analysis of hepatocellular carcinoma tumour microenvironment based on imaging mass cytometry reveals cellular neighbourhood regulated reversely by macrophages with different ontogeny. *Gut* 71, 1176–1191. <https://doi.org/10.1136/gutjnl-2021-324339>.
- Sun, H., Song, J., Weng, C., Xu, J., Huang, M., Huang, Q., Sun, R., Xiao, W., and Sun, C. (2017). Association of decreased expression of the macrophage scavenger receptor MARCO with tumor progression and poor prognosis in human hepatocellular carcinoma. *J. Gastroenterol. Hepatol.* 32, 1107–1114. <https://doi.org/10.1111/jgh.13633>.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of Single-Cell data. *Cell* 177, 1888–1902.e21. <https://doi.org/10.1016/j.cell.2019.05.031>.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102, 15545–15550. <https://doi.org/10.1073/pnas.0506580102>.
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32, 381–386. <https://doi.org/10.1038/nbt.2859>.
- Tanaka, M., Nakashima, O., Wada, Y., Kage, M., and Kojiro, M. (1996). Pathomorphological study of Kupffer cells in hepatocellular carcinoma and hyperplastic nodular lesions in the liver. *Hepatology* 24, 807–812. <https://doi.org/10.1053/jhep.1996.v24.pm0008855180>.
- Talevich, E., Shain, A.H., Botton, T., and Bastian, B.C. (2016). CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol.* 12, e1004873. <https://doi.org/10.1371/journal.pcbi.1004873>.
- Tajiri, K., Tsuneyama, K., Kawai, K., Atarashi, Y., Minemura, M., Sawada, S., Tsukada, K., Imura, J., and Sugiyama, T. (2014). A case of progressing focal nodular hyperplasia and its molecular expression pattern. *Clin. J. Gastroenterol.* 7, 271–277. <https://doi.org/10.1007/s12328-014-0483-5>.
- Teixeira, V.H., Pipinikas, C.P., Pennycuik, A., Lee-Six, H., Chandrasekharan, D., Beane, J., Morris, T.J., Karpathakis, A., Feber, A., Breeze, C.E., et al. (2019). Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nat. Med.* 25, 517–525. <https://doi.org/10.1038/s41591-018-0323-0>.
- Torri, G.B., Soldatelli, M.D., Luersen, G.F., and Almeida Ghezzi, C.L. (2021). Imaging of chemotherapy-induced liver toxicity: an illustrated overview. *Hepat. Oncol.* 8, HEP32. <https://doi.org/10.2217/hep-2020-0026>.
- Tomasetti, C. (2019). Mutated clones are the new normal. *Science* 364, 938–939. <https://doi.org/10.1126/science.aax5525>.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. <https://doi.org/10.1038/nprot.2012.016>.
- Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38, e164. <https://doi.org/10.1093/nar/gkq603>.
- Wilkerson, M.D., and Ayes, D.N.H. (2010). Consensusclusterplus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26, 1572–1573.
- Wijewardhane, N., Dressler, L., and Ciccarelli, F.D. (2021). Normal somatic mutations in cancer transformation. *Cancer Cell* 39, 125–129. <https://doi.org/10.1016/j.ccell.2020.11.002>.
- Wickham, H. (2011). ggplot2. *WIREs. Comp. Stat.* 3, 180–185. <https://doi.org/10.1002/wics.147>.
- Wells, R.G., Kruglov, E., and Dranoff, J.A. (2004). Autocrine release of TGF- β by portal fibroblasts regulates cell growth. *FEBS Lett.* 559, 107–110. [https://doi.org/10.1016/S0014-5793\(04\)00037-7](https://doi.org/10.1016/S0014-5793(04)00037-7).
- Yasui, K., Arai, S., Zhao, C., Imoto, I., Ueda, M., Nagai, H., Emi, M., and Inazawa, J. (2002). TFDP1, CUL4A, and CDC16 identified as targets for amplification at 13q34 in hepatocellular carcinomas. *Hepatology* 35, 1476–1484. <https://doi.org/10.1053/jhep.2002.33683>.
- Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. <https://doi.org/10.1089/omi.2011.0118>.
- Zhu, K., Wang, W., Luo, R., Song, D., Wang, X., Gao, Q., Fan, J., Zhou, J., Rao, S., and Wang, X. (2021). Newly detected liver nodules with a history of colorectal cancer: are they metastatic? Review of 2, 632 cases in a single center. *Ann. Transl. Med.* 9, 1079. <https://doi.org/10.21037/atm-20-8153>.
- Zhu, M., Lu, T., Jia, Y., Luo, X., Gopal, P., Li, L., Odewole, M., Renteria, V., Singal, A.G., Jang, Y., et al. (2019). Somatic mutations increase hepatic clonal fitness and regeneration in chronic liver disease. *Cell* 177, 608–621.e12. <https://doi.org/10.1016/j.cell.2019.03.026>.
- Zekri, A.R.N., Sabry, G.M., Bahnassy, A.A., Shalaby, K.A., Abdel-Wahab, S.A., and Zakaria, S. (2005). Mismatch repair genes (HMLH1, HPMS1, HPMS2, GTBP/HMSH6, HMSH2) in the pathogenesis of hepatocellular carcinoma. *World J. Gastroenterol.* 11, 3020–3026. <https://doi.org/10.3748/wjg.v11.i20.3020>.
- Ziche, M., Morbidelli, L., Donnini, S., and Ledda, F. (1995). ETB receptors promote proliferation and migration of endothelial cells. *J. Cardiovasc. Pharmacol.* 26, S284–S286. <https://doi.org/10.1097/00005344-199506263-00085>.
- Zhang, M., Yang, H., Wan, L., Wang, Z., Wang, H., Ge, C., Liu, Y., Hao, Y., Zhang, D., and Shi, G. (2020). Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. *J. Hepatol.* 73, 1118–1130. <https://doi.org/10.1016/j.jhep.2020.05.039>.
- Zhang, W.Z., Venardos, K., Chin-Dusting, J., and Kaye, D.M. (2006). Adverse effects of cigarette smoke on no bioavailability. *Hypertension* 48, 278–285. <https://doi.org/10.1161/01.HYP.0000231509.27406.42>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CD68	abcam	Cat# ab213363
CD206	abnova	Cat# 5C11
MARCO	ATLAS ANTIBODIES	Cat# HPA063793
PDGFRB	abcam	Cat# ab32570
PDGFB	abcam	Cat# ab23914
α SMA	abcam	Cat# ab5694
CD31	abcam	Cat# ab76533
GS	abcam	Cat# ab176562
CK19	abcam	Cat# ab52625
Ki-67	abcam	Cat# ab15580
Chemicals, peptides, and recombinant proteins		
DMEM, without glucose, L-glutamine, phenol red, sodium pyruvate and sodium bicarbonate	Sigma-Aldrich	Cat#D5030
Fetal Bovine Serum	Gibco	Cat# 10099141
Opal™ 7-Color Manual IHC Kit 50 slides	PerkinElmer	NEL811001KT
Collagenase IV	Gibco	Cat# 17104019
Dispase II	Gibco	Cat# 17105041
Critical commercial assays		
RNAscope® 2.5 HD Assay - Brown	Advanced Cell Diagnostics	Cat#322300
RNAscope® SOST Probes	Advanced Cell Diagnostics	Cat# 452941
RNAprep Pure Tissue Kit	TIANGENE	Cat#DP431
NEBNext® Ultra™ II Directional RNA Library Prep Kit	New England BioLabs	Cat#E7765
TIANamp Genomic DNA kit	TIANGEN	Cat#DP130227
Agilent SureSelect XT Human All Exon V5 kit	Agilent	Cat#5190-6213
Chromium Single Cell 3' Reagent Kits	10× Genomics	Cat# PN-1000057
Deposited data		
Normal liver single cell data	(Aizarani et al., 2019)	GSE124395
Normal liver single cell data	MacParland et al., 2018	GSE115469
Liver cirrhosis single cell data	(Ramachandran et al., 2019)	GSE136103
HCC and fetal liver single cell data	(Sharma et al., 2020)	GSE156625
Intrahepatic CCA single cell data	(Zhang et al., 2020)	GSE138709
Oligonucleotides		
Primers for Sanger sequencing, see Table S3	This paper	N/A
Primer for SOST	This paper	N/A
Forward: CGTCAGGATTGTCGGCAGTA		
Reverse: CCTTTGGAGTAACCCAGCAT		
Software and algorithms		
R V3.6.2	R Core Team	https://www.r-project.org/
GraphPad Prism V8.0	GraphPad Software Inc	https://www.graphpad.com/scientific-software/prism
Burrows- Wheeler Alignment V0.7.12	Li and Durbin, 2010	http://bio-bwa.sourceforge.net/

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FastQC V0.11.8	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Samtools V0.1.19	Danecek et al., 2021	https://github.com/samtools/samtools
Mutect2	Broad Institute	https://gatk.broadinstitute.org/hc/en-us/articles/360037593851-Mutect2
CNVkit V0.9.5	Talevich et al., 2016	https://cnvkit.readthedocs.io/en/stable/
dNdScv	Martincorena et al., 2017	https://github.com/im3sanger/dndscv
Maftools V3.1	Mayakonda et al., 2018	https://www.bioconductor.org/packages/release/bioc/html/maftools.html
TopHat2 V2.1.1	(Kim et al., 2019)	http://ccb.jhu.edu/software/tophat/index.shtml
Cufflinks V2.1.1	(Trapnell et al., 2012)	http://cole-trapnell-lab.github.io/cufflinks/
DESeq2 V1.14.1	Love et al., 2014	http://bioconductor.org/packages/release/bioc/html/DESeq2.html
edgeR V3.12	(Robinson et al., 2009)	https://bioconductor.org/packages/release/bioc/html/edgeR.html
ggplot2 V3.3.0	(Wickham, 2011)	https://cran.r-project.org/web/packages/ggplot2/index.html
Biodica	Captier N et al.	https://github.com/LabBandSB/BIODICA
Consensusclusterplus V1.5	(Wilkinson and Ayes, 2010)	http://bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html
Cellranger V3.1	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
Seurat V3.1.4	Stuart et al., 2019	https://cran.r-project.org/web/packages/Seurat/index.html
Monocle V2.1.4	Trapnell lab	https://www.bioconductor.org/packages/release/bioc/html/monocle.html
Cibersortx	Newman et al., 2019	https://cibersortx.stanford.edu/
ggsci V2.9	Xiao and Li	https://CRAN.R-project.org/package=ggsci

RESOURCE AVAILABILITY**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Prof. Qiang Gao (gao.qiang@zs-hospital.sh.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All sequencing raw data, including WES, RNA-seq and scRNA-seq, are deposited in NODE (<https://www.biosino.org/node>) under the accession number OEP002718. This paper does not report original code. Any additional information required to reanalyze the data reported in this work paper is available from the **lead contact** upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS**Clinical sample acquisition**

19 Paired FNH and aNL samples for WES and RNA-seq were all acquired from patients underwent curative resection from January 2012 to December 2013 at Zhongshan Hospital. Three paired FNH and aNL samples for scRNA-seq were from December 2019 to January 2020 at Zhongshan Hospital. Samples of Hematoxylin-Eosin staining, IHC, fluorescent immunostaining and qRT-PCR were included in the 19 sample cohorts. Patients for ISH were from February 2021. The study was approved by the Research Ethics Committee of Zhongshan Hospital, and written informed consent was obtained from each patient.

METHOD DETAILS

Specimens processing

After resected, paired FNH and aNL for sequencing were transferred in 10% FBS (Gibco, USA) RPMI-1640 medium (Sigma-Aldrich, USA), followed by washing for three times by PBS solution and visibly necrotic and hemorrhagic parts were carefully removed. Then, the clean samples were put into liquid nitrogen for snap frozen and stored in -80°C for further processing. The time from sample collection to storage at -80°C was strictly controlled within 30 minutes. aNL was used for germline analysis. All surgically resected samples were macroscopically and microscopically diagnosed by two experienced pathologists to assure FNH diagnosis.

DNA extraction, library construction, WES and sanger validation

TIANamp Genomic DNA kit (TIANGEN, Beijing, China) standard procedure was applied to extract tissue DNA. Quality control was measured by Nanodrop 2000, and DNA content per sample >200 ng and concentration >20 ng/ μL was accepted for following steps. Covaris™ S2 Ultrasonicator System (Covaris, Woburn, MA, USA) was used for DNA segmentation of 150 bp. 200ng DNA for each sample was used for library construction. Library was built in accordance to Agilent SureSelect XT Human All Exon V5 kit (Agilent, CA, USA) guideline. 150 bp paired-end sequencing was performed on Illumina X-ten. Primers for Sanger sequencing were summarized in [Table S3](#).

RNA extraction, library construction and RNA-seq

Total RNA was extracted from freshly frozen FNH and aNL samples using RNeasy Pure Tissue Kit (QIAGEN, Beijing, China) followed by Nanodrop 2000 for quality control with the threshold of total RNA >4 μg . 200 ng RNA for each sample was sent for library construction. We prepared RNA-seq library with NEBNext® Ultra™ II Directional RNA Library Prep Kit (NEB, MA, USA) for Illumina in accordance with the manufacturer's instructions. 150 bp paired-end sequencing was performed on Illumina X-ten (Illumina, CA, USA).

WES data processing, somatic mutation detection

FastQC V1.9 (<https://github.com/s-andrews/FastQC>) and MultiQC V1.9 (Ewels et al., 2016) were used for quality control of primary WES data, fastp tool V0.20.1 (Chen et al., 2018) was used to filter data and sequence was mapped to hg19 using BWA V0.7.15 (Li and Durbin, 2010). We used GATK BaseRecalibrator V4.0.6.0 (<https://gatk.broadinstitute.org/hc/en-us/articles/360036898312-BaseRecalibrator>) to adjust base quality. Filtered data was sent to GATK Mutect2 (<https://gatk.broadinstitute.org/hc/en-us/articles/360037593851-Mutect2>) to call somatic mutation followed by Annovar V2017 (Wang et al., 2010) annotation. Mutations with alteration depth $<5\text{X}$ or total depth of FNH $<20\text{X}$ were removed and we used IGV (<https://software.broadinstitute.org/software/igv/igvtools>) to verify data. Important databases related to Annovar were COSMIC (<http://cancer.sanger.ac.uk/cosmic>), dbSNP146 (<http://www.ncbi.nlm.nih.gov/snp>), 1000genomes (<http://www.1000genomes.org/>) and EXAC (<http://exac.broadinstitute.org/>).

RNA-seq data processing

Raw data was imported to FastQC V1.9 (<https://github.com/s-andrews/FastQC>) for quality control and was subsequently sent to fastp V0.20.1 (Chen et al., 2018) for filtering low quality sequence. Reads were mapped to the human genome hg19 with HISAT2 V2.0.5 (Kim et al., 2019) and Cufflinks (Trapnell et al., 2012). SAM to BAM transformation was finished by Samtools V1.7 (Daneczek et al., 2021). The final count files annotated by gencode.v33.annotation.gtf were produced from FeatureCounts V1.22.2 (Liao et al., 2014) processed BAM files.

dN/dS analysis

To estimate whether FNH has positive selection on mutations, we performed dN/dS analysis referred to the previously described method (Martincorena et al., 2015). VCF files including silent and non-silent mutations exported from the WES upstream analysis were imported to dNdScv R package (Martincorena et al., 2017), and genes with $w_{\text{mis_cv}}$, $w_{\text{non_cv}}$ or $w_{\text{ind_cv}} > 1$, and the corresponding q value <0.05 were considered as positive selection. RTCGA and RTCGA.mutations.20160128 packages (Kosinski and Biecek, 2021) were used to download TCGA liver hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA) mutational information. Same analytical process was also applied to TCGA data.

Mutational signature analysis

Due to the relatively low mutation number in each sample compared with malignancies, we manually merge SNVs from all 19 samples as one integrated VCF file. Then, the VCF file was imported to maftools R package (Mayakonda et al., 2018) for illustration and further analysis. The mutational signature was acquired using deconstructSigs (Rosenthal et al., 2016), and the cosine similarity was calculated by comparing the signature with COSMIC predefined Mutational Signatures V2. The cosine similarity ranges from 0 to 1, and the closer the number is to 1, the higher the similarity between each pair.

Identification of CNVs

Both FNH and aNL BAM files were imported into CNVkit (Talevich et al., 2016), and the aNL file was pooled and set as reference. The final results were generated using default parameter. The copy number was thought to be non-diploid if $\log_2(\text{copy number fold change}) > 0.585$ or ≤ -1 . And the arm-level changes were defined as ≥ 0.1 fold change detected on $\geq 50\%$ length of that chromosome arm.

Clone number estimation

We estimated clone number with maftools (Mayakonda et al., 2018) and the variant allele frequency (VAF) was used to calculate Mutant-Allele Tumor Heterogeneity (MATH) score. MATH score is a method to measure intra-tumor genetic heterogeneity. When evaluating malignant tumors, high MATH score is correlated to lower survival rate (Rajput et al., 2017). Here, we put this method into benign tumor to estimate whether mutations derived from monoclonal cells. Samples with >5 somatic mutations were calculated.

TMB calculation and comparison with TCGA datasets

33 TCGA tumor mutational profiles were acquired with RTCGA and RTCGA.mutations.20160128 packages (Kosinski and Biecek, 2021). Setting the filtering depth threshold of our dataset as the reference, we ruled out low quality mutations using the depth-adjusted threshold for TCGA samples. To make our dataset and TCGA datasets comparable, we then adjusted sequenced exome length and sorted each tumor by median mutational number. Mann-Whitney U test was performed between FNH and each tumor, respectively. When considering VAF which means the scale a mutation could affect, we substituted each single point in previous graph with VAF of each mutation. Mann-Whitney U test was performed as described above.

DEG detection and pathway analysis

Different expressed genes were identified by DESeq2 (Love et al., 2014) R package V3.12 embedding shrinkage estimation for dispersion and fold change. Read counts were corrected for their patient origin, and genes expressed more than three patients were included for further analysis. We used $\log_2(\text{fold change}) \geq 1$ or ≤ -1 and FDR q value < 0.05 as the threshold for significant gene selection. For function and pathway enrichment, we used clusterProfiler (Yu et al., 2012) R package V3.14.0 and ToppGene tool (<https://toppgene.cchmc.org/enrichment.jsp>) implementing canonical database such as KEGG, GO, REACTOME. Results with FDR q value < 0.05 was thought significant.

PCA of FNH, TCGA and GTEx transcriptome

HCC and CCA read counts were downloaded from GDC portal (<https://portal.gdc.cancer.gov/>), and normal liver read counts was downloaded from GTEx portal (<https://www.gtexportal.org/home>). Only protein coding genes were extracted from the original matrixes, and these protein coding matrixes including FNH were merged together. We used DESeq2 R package V3.12 (Love et al., 2014) embedded variance stabilizing transformation that normalizes the count data by dividing the normalization factors. Genes expressed less than 10% of the corresponding sample were excluded from further analysis. PCA was performed using DESeq2 R package described above and the 3D plot was graphed by plot3D R package V1.3 (<https://CRAN.R-project.org/package=plot3D>) using top 3 principle components.

Consensus clustering of FNH, TCGA and GTEx samples

Consensus clustering is a method of aggregation of clustering. It refers to that the given clusters of a dataset desire to fit for a better existing clustering. Thus, this algorithm can reconcile data from different samples or results from different runs of an algorithm (Bonizzoni et al., 2008). Therefore, we used consensus clustering embedded ConsensusClusterPlus R package V3.12 (Wilkerson and Ayes, 2010) which can also determine the ideal cluster number to comprehensively analyze data from our dataset, TCGA and GTEx. We normalized read count from different dataset via weighted trimmed mean of M-values (TMM) algorithm

built in edgeR R package V3.12 (Robinson et al., 2009). We then used median absolute deviation (MAD) to select how many genes were included for the next step. The similarity between each sample and each unsupervised generated cluster was established by 1000 iterations.

ICA and correlation with five types of samples

ICA is a decomposition method that extracts a group of independent signal from a multivariate signal. The process can be described as $X = AS$, and X represents the raw signal, A is the loading matrix and S is the weight matrix of each independent component (IC). FastICA (Hyvärinen and Oja, 2000), an efficient algorithm implementing ICA, orchestrate lasso algorithm (Himberg et al., 2004) to synergistically explore the stable IC number and subsequently acquire a robust IC result. We used MATLAB developed BIODICA software windows GUI version (<https://github.com/LabBandSB/BIODICA>) which relies on FastIC and lasso, running for 100 times respectively for 10 to 70 ICs to inspect the compactness, and finally decided the optimal IC result. Then, the correlation coefficient between each IC and every BIODICA built-in biologically significant gene set. Biologically prominent ICs with correlation coefficient >0.4 were directly entitled as the corresponding name, such as “proliferation” and “fibrosis”. We set the binary vector for each sample within specific type as 1 and outside as 0. To establish the relation between ICs and sample types, we calculated the biserial correlation coefficient between the binary vector and the metasample (sample versus component). Lastly, the simple Pearson correlations were calculated between each IC pairs (Aynaud et al., 2020). In our study, the correlation coefficient >0.5 was considered significant.

Annotation of FNH specific IC and samples scoring

The FNH specific metagenes (IC versus genes) were ranked by the contribution factor by decreasing. The ranked list was then imported into GSEA_4.0.3 (Subramanian et al., 2005; Mootha et al., 2003), and GSEA Preranked tool with 1000 permutation was performed. MSigDB V7.2 (Gouw et al., 2010) H, C2, C5 and cell signatures obtained from xCell (Aran et al., 2017) database were set as the standard signature. Terms whose normalized enrichment score (NES) > 1.5 and FDR q value <0.05 were significant. Of each gene for each IC in metagene matrix, contribution factor >3 were ideal representative for the IC. Therefore, we used the sum of transcripts per million (TPM) of each representative gene as the IC score. Hypothesis testing was carried out using Mann-Whitney U test between FNH and sample x ($x = \text{aNL, TCGA-HCC, TCGA-CCA and GTEx Normal}$).

Single-cell suspension preparation

Three paired freshly resected FNH and aNL samples were immediately transferred into a 4°C DMEM (Gibco, CA, USA) filled 50 mL centrifugal tube with 10% fetal bovine serum (Gibco, CA, USA) and samples were transported to our lab on ice. Clean tissue was trimmed to 5 × 5 × 5 mm in size and was subsequently emerged into 10 mL complex digestive enzyme system including 1 mg/mL collagenase IV (Gibco, CA, USA), and 1 U/mL dispase II (Gibco, CA, USA). The 10 mL system was constantly stirred at 37°C for 40 min. The dissociated solution was filtered through a 40- μm cell-strainer nylon mesh (BD, NJ, USA) followed by 700 g centrifugation for 10 min. Supernatant was removed and the cell sediment was washed twice with MACS solution (PBS containing 1% FBS, 0.5% EDTA, and 0.05% gentamycin). Cells were finally dissolved in sorting buffer (PBS with 1% FBS), stained with DRAQ5 1:200 for 15 min and DAPI 1:200 for 15 min. Living cells were then assessed and sorted by MoFlo Astros EQ (Beckman Coulter, CA, USA) Cell Sorter into 10% FBS DMEM solution. Living cells took up $> 90\%$ of total cells.

Library construction and scRNA-sequencing

We used Chromium Single Cell 3' Reagent Kits V3 (10× Genomics, CA, USA) to construct libraries according to manufacturer's instructions. We aimed to harvest 6000 cells per cells from each channel. Libraries were constructed using 10× Genomics Single-Cell Instrument and drops were then generated. Samples were sequenced on Illumina NovaSeq.

scRNA sequencing data processing

Raw data was processed by Cell Ranger Single-Cell Software TABLE suite (10× Genomics, CA, USA) for doublet removing, barcode processing, alignment and quality control. Reads were mapped to GRCh38 annotated with corresponding annotation files in ENSEMBL database by Cellranger pipeline V3.0.1 (<https://github.com/10XGenomics/cellranger>). To optimally balance cell number and quality, we kept cells with both $600 < \text{genes} < 4000$, $1200 < \text{UMIs} < 20,000$ and percentage of mitochondrial genes $< 20\%$. Moreover,

we manually checked the marker genes of each cluster and deleted those clusters marked with contradictory canonical genes. A total of 27,474 cells from FNH and 26,747 cells from aNL finally passed the quality control, with $1,686 \pm 793$ (mean \pm SD) genes for cells of FNH and $1,387 \pm 527$ (mean \pm SD) genes for cells of aNL.

Unsupervised clustering and determination of cell type

We used log-transformed method to normalize counts and the top 2000 highly variable genes were subsequently scaled for principle component analysis. Seurat V3.2.3 (Stuart et al., 2019) embedding Harmony algorithm (Korsunsky et al., 2019), a strong method to minimize technical or biological confounders to integrate different datasets and preserve biological characteristics was used to combine cells derived from three patients. Top 20 principle components were used for clustering and Uniform Manifold Approximation and Projection (UMAP) was applied for visualization. Marker genes were detected using FindAllMarker function. Genes expressed in >25% cells in each cluster and $\log_2(\text{fold change}) \geq 1$ were recognized as marker genes. Next, we employed singleR V1.0.6 R package (Aran et al., 2019) to preliminarily annotate clusters followed by manual verification using canonical marker genes published in papers or documented in databases.

Deconvolution of RNA-seq data

Our bulk RNA-seq data was deconvoluted using Cibersortx (Newman et al., 2019), LM22 signature matrix was used as background reference and 500 permutations were applied. Absolute mode was finally set to acquire results.

Pseudo-bulk sample generation

The pseudo-bulk samples were generated by adding all single-cell read counts individually. Genes expressed by less than 10 cells were excluded from analysis and DEG and pathway enrichment analysis were performed the same way described in "Differentially Expressed Genes Detection and Pathway Analysis". DEGs with adjusted p value <0.05 and $|\log_2(\text{fold change})| \geq 1$ were considered significant.

Gene set enrichment of single-cell data

Gene set variation analysis (GSVA), depending on a non-parametric unsupervised method, transform an expression matrix into a relative enrichment score matrix. The count files were scaled to conform to Gaussian distribution and imported into GSVA function. For visualization, we used z-score to plot heatmap.

Pseudo-trajectory inference

We used Monocle V2.12.0 R package (Trapnell et al., 2014) to construct the pseudo-trajectory of endothelial cells. Differentially expressed genes (expressed >10% cells) were chosen using differentialGeneTest function and the genes with FDR q value <0.01 were selected for DDRTree which is responsible for trajectory construction.

Metabolic gene signature scoring

We obtained metabolic gene signatures from the published article (Trapnell et al., 2014). Both Seurat object and signature genes were imported to AddModuleScore function of Seurat, which calculates the difference of the average expression of target genes and control genes randomly sampled from each bin. We set 100 control genes and 24 bins for analysis. Each cell would get a score of the signature.

Intercellular ligand-receptor analysis

We used CellPhoneDB V2.0 to infer the ligand-receptor crosstalk between single cells (Efremova et al., 2020). Significant secreting ligand and receptor was selected and visualized by ggplot2 (Wickham, 2011) and ggsci (<https://CRAN.R-project.org/package=ggsci>).

Immunofluorescence and immunohistochemistry

The paraffin section was baked for 1 h and dewaxed by 10 min xylene for three times followed by 100%, 95%, 85% and 75% gradient dehydration. Then, the section was washed for twice and placed in boiling antigen retrieval buffer (citrate PH = 6 or EDTA PH = 9) for 10 min. Next, the section was incubated with 3% H₂O₂ for 10 min for blocking endogenous peroxidase. Non-specific binding was blocked by goat serum

(Vector, CA, USA) for 30 min. The section was incubated with primary antibody over night at 4°C or for 1 h at room temperature. After washing for three times, the section was then incubated with corresponding second antibody for 25 min. Following three-time washing, Opal® (PerkinElmer, MA, USA) fluorescence dye kit was used to stain the target. The whole process from antigen retrieval to fluorescence staining was repeated till the last marker was finished. Finally, nuclei were stained using DAPI (Sigma-Aldrich, USA) and slides were mounted with fluorescence mounting media (DAKO, CA, USA). Slides were scanned and analyzed using PerkinElmer Vectra3® platform.

Immunohistochemistry followed the same steps as immunofluorescence, however, after second body incubation, the section was processed stepwise by hematoxylin, acidic differentiation solution and bluing buffer. Then, the section was dehydrated by 100% ethanol and xylene followed by resin mounting.

RNA extraction and qRT-PCR

Frozen tissue was minced thoroughly in liquid nitrogen precooled mortar into powder. After liquid nitrogen completely evaporated, 1 mL Trizol (Invitrogen, CA, USA) was added followed by 200 µL chloroform. Centrifugation was performed and the supernatant was moved to 500 µL isopropanol for RNA precipitation. Centrifugation was performed again and 75% ethanol (prepared with DEPC water) was added for washing.

RNA with OD260/OD280 > 1.8 was considered qualified. Reverse transcription and qRT-PCR were performed according to the manufacturer's instruction (Yeasen, Shanghai, China) using QuantStudio 3 (ABI, MA, USA).

Primers for MARCO:

F: CGTCAGGATTGTCGGCAGTA.

R: CCTTTGGAGTAACCCAGCAT.

RNA *in situ* hybridization

RNAscope® 2.5 HD Assay - Brown and RNAscope® SOST Probes were purchased from Advanced Cell Diagnostics (ACD, CA, USA). Freshly resected FNH tissue was immediately transferred to cold PBS to remove as much blood as possible. Then, samples were immersed in 10% neutral buffered formalin for 24 h for paraffin embedding. After deparaffinized, RNAscope® Hydrogen Peroxide and RNAscope® Target Retrieval Reagents were used to quench endogenous peroxidase and retrieve antigens respectively, followed by protein digestion using RNAscope® Protease Plus. Finally, targeted RNA was hybridized by SOST probe and RNAscope® 2.5 HD Detection Reagents – BROWN was used to amplified the signal.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

Standard statistical methods were utilized, including Student's t test, Paired Student's t test and Mann-Whitney U test. For clinical data comparison, single-omic and multi-omic analyses, Student's t test was used to compare the age and tumor size between two gender groups. Mann-Whitney U test was applied to the comparison of TMB or VAF-TMB between FNH and each TCGA tumor samples, IC scores between FNH and each type of samples. For experiment data, Student's t test was used to compare the MARCO expression level between FNH and HCC. Paired Student's t test was utilized to measure the content difference of macrophages between FNH and paired aNL. All statistical tests were two-sided, and statistical significance was considered when P value < 0.05. For continuous variables versus continuous variables, Spearman's rank or Pearson correlation was used. For category versus continuous variables, the point-biserial correlation was used. Statistical analysis was performed using GraphPad Prism 8. Correlation analyses were performed in R v3.6.3. Data were represented using the mean ± standard deviation (SD), unless indicated otherwise.