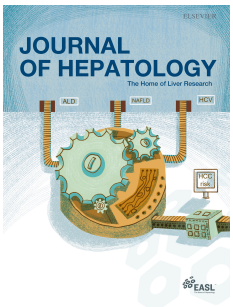
# Journal Pre-proof




Hepatocytes undergo punctuated expansion dynamics from a periportal stem cell niche in normal human liver

A.M. Passman, M.J. Haughey, E. Carlotti, M.J. Williams, B. Cereser, M.L. Lin, S. Devkumar, J.P. Gabriel, E. Gringeri, U. Cillo, F.P. Russo, M. Hoare, J. ChinAleong, M. Jansen, N.A. Wright, H.M. Kocher, W. Huang, M.R. Alison, S.A.C. McDonald

PII: S0168-8278(23)00223-4

DOI: https://doi.org/10.1016/j.jhep.2023.03.044

Reference: JHEPAT 9124

To appear in: *Journal of Hepatology*

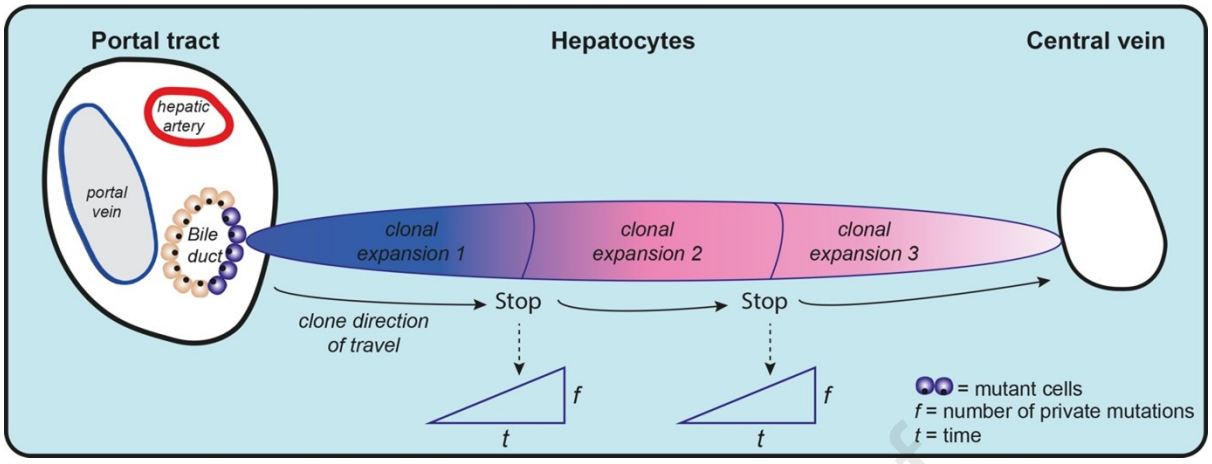
Received Date: 8 June 2022

Revised Date: 29 March 2023

Accepted Date: 30 March 2023

Please cite this article as: Passman A, Haughey M, Carlotti E, Williams M, Cereser B, Lin M, Devkumar S, Gabriel J, Gringeri E, Cillo U, Russo F, Hoare M, ChinAleong J, Jansen M, Wright N, Kocher H, Huang W, Alison M, McDonald S, Hepatocytes undergo punctuated expansion dynamics from a periportal stem cell niche in normal human liver, *Journal of Hepatology* (2023), doi: https://doi.org/10.1016/j.jhep.2023.03.044.

*Punctuated hepatocyte expansion from a common cell of origin with biliary epithelium*



# Hepatocytes undergo punctuated expansion dynamics from a periportal stem cell niche in normal human liver

Passman AM[1*], Haughey MJ[2*], Carlotti E[1], Williams MJ[1], Cereser B[3], Lin ML[1], Devkumar[1] S, Gabriel JP[3], Gringeri E[4], Cillo U[4], Russo FP[4], Hoare M[5], ChinAleong J[6], Jansen M[7,8], Wright NA[3], Kocher HM[3,9,10], Huang W[2,11§], Alison MR[3§] & McDonald SAC[1§]

[1]Centre for Cancer Genomics and Computational Biology, [3]Centre for Tumour Biology, [9]Cancer Tissue Bank, Barts Cancer Institute, Queen Mary University of London, London, UK.

[2]School of Mathematical Sciences, Queen Mary University of London, London, UK.

[4]Department of Surgery, Oncology and Gastroenterology, University of Padova, Padova, Italy

[5]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK.

[6]Histopathology, Barts Health NHS Trust, London, UK.

[7]Department of Cellular Pathology, [8]UCL Cancer Centre, University College London, London, UK.

[10]Barts and the London HPB Centre, The Royal London Hospital, Barts Health NHS Trust, London, UK.

[11]Group of Theoretical Biology, The State Key Laboratory of Biocontrol, School of Life Science, Sun Yat-sen University, Guangzhou, China

*Contributed equally to this work

§ these authors jointly supervised this work

**Corresponding Author**: Dr Stuart McDonald, Centre for Cancer Genomics and Computational Biology, 3rd Floor, John Vane Science Centre, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London, UK, EC1M 3BQ.

Email: s.a.mcdonald@qmul.ac.uk

Tel: +44 207 882 3800

## Author contributions

W.H., M.R.A. and S.A.C.M. were responsible for concept development and design. A.M.P., M.J.H., E.C., M.J.W., B.C., M-L.L., S.D., E.A.S. and J.P.G. were responsible for data acquisition. A.M.P., M.J.H., M.J, N.A.W., W.H., M.R.A. and S.A.C.M. were responsible for data analysis and interpretation. H.M.K., M.M.H, J.C-A., E.G., U.C., F.P.R., provided materials and clinical or pathological support. A.M.P., M.J.H., W.H and S.A.C.M wrote the manuscript. All authors identified above have critically reviewed the paper and approved the final version.

## Competing interests

The authors declare no competing interests, financial or otherwise.

## Keywords

Stem cell, clonal expansion, mitochondria

## Abstract

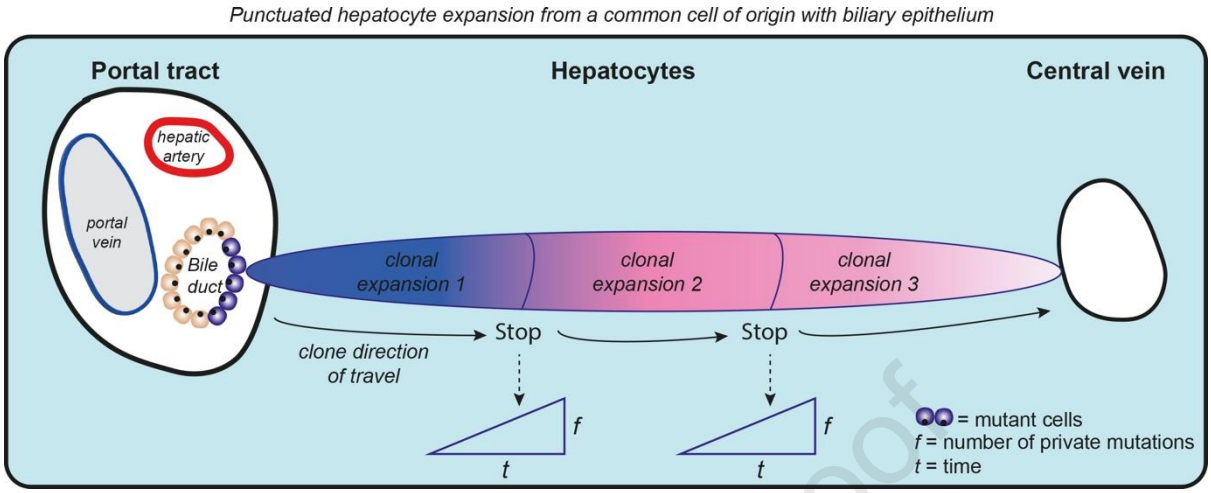
**Background:** Normal human liver is thought to be generally quiescent, however clonal hepatocyte expansions have been observed but neither their cellular source nor their expansion dynamics have been determined. Knowing the hepatocyte cell of origin, and their subsequent dynamics and trajectory within the human liver will provide an important basis to understand disease-associated dysregulation.

**Methods:** Here we use *in vivo* lineage tracing and a combination of methylation sequence analysis to demonstrate normal human hepatocyte ancestry. We exploit next generation mitochondrial sequencing to determine hepatocyte clonal expansion dynamics across spatially-distinct areas of laser-captured, microdissected clones, in tandem with computational modelling in morphologically-normal human liver.

**Results:** Hepatocyte clones and rare $SOX9^{+}$ hepatocyte progenitors commonly associate with portal tracts and we present evidence that clones can lineage-trace with cholangiocytes, indicating the presence of a bipotential common ancestor at this niche. Within clones, we demonstrate methylation CpG sequence diversity patterns indicative of periportal, but not pericentral ancestral origins, indicating a portal to central vein expansion trajectory. Using spatial analysis of mtDNA variants by next-generation sequencing coupled with mathematical modelling and Bayesian inference across the portal-central axis, we demonstrate that patterns of mtDNA variants reveal large numbers of spatially-restricted mutations in conjunction with limited numbers of clonal mutations.

**Conclusions:** These datasets support the existence of a periportal progenitor niche and indicate that clonal patches exhibit punctuated but slow growth, then quiesce, likely due to acute environmental stimuli. These findings crucially contribute to our understanding of hepatocyte dynamics in normal human liver.

## Graphical abstract




## Lay abstract

The liver is mainly composed of hepatocytes but we know little regarding the source of these cells or how they multiply over time within disease-free human liver. In this study, we determine a source of new hepatocytes by combining many different lab-based methods and computational predictions to show that hepatocytes share a common cell of origin with bile ducts. Both our experimental and computational data also demonstrate hepatocyte clones are likely to expand in slow waves across the liver in a specific trajectory, but often lie dormant for many years. These data show for the first time the expansion dynamics of hepatocytes in normal liver.

## Background

The liver possesses an extensive playbook for cell renewal and recovery from injury. Decades of partial hepatectomy studies have established that mature hepatocytes, regardless of their anatomic location, are capable of restoring the vast majority of the liver mass (1, 2). Hepatocytes are seemingly inexhaustible in their replicative capacity, without signs of malignant transformation (3, 4) and are capable of dedifferentiation, expansion and redifferentiation, both *in vitro* (5) and in rodent damage models (6, 7), and thus, may themselves be capable of regenerating the biliary epithelium (8, 9). Self-renewal and multipotentiality, properties of stem cells, are unusual for a differentiated cell. As such, the necessity for a dedicated hepatic stem cell niche, as is present in many other organs has been questioned.

Regardless of hepatocyte plasticity, cells variously known as liver progenitor cells, oval cells or hepatic progenitor cells have in fact demonstrated bipotentiality and are thought to reside close to, or at the portal tract (PT) (10, 11, 12). The contribution of these cells to liver regeneration *in vivo* has been widely reported using rodent injury models (13, 14, 15). This has been the source of some controversy, with several lineage-tracing studies failing to detect significant proportions of progenitor-derived hepatocytes upon regeneration from injury (16, 17). Evidence is now mounting that widespread impairment of hepatocyte replication is required for activation of the progenitor cell regenerative compartment (13, 14, 18). This finding is particularly relevant for human disease in which widespread hepatocyte replicative impairment is commonplace.

Whilst there is a reasonable degree of clarity regarding the cellular responses upon liver injury, there is still conflicting opinion regarding homeostatic dynamics. Early studies utilised tritiated thymidine to track hepatocytes in normal rat livers over time (19, 20). In these studies, hepatocyte migration along the portal tract-central vein axis was observed; a phenomenon coined ‘the streaming liver’. Subsequent studies have proposed homeostatic neo-hepatocyte generation from all three zones in the hepatic lobule; periportal (21), mid-zonal (17, 22) and pericentral (23). Further still, there is evidence that homeostatic turnover results from zone-agnostic, broadly distributed hepatocytes (24, 25, 26, 27). Beyond rodent models, there is evidence supporting the existence of a periportal progenitor niche in human liver. Huch *et al.* were able to generate bipotential organoids from $EPCAM^{+}$ ductal cells

isolated from normal human livers (28). Similarly, a periportal, bipotent, organoid-forming cell population was identified from a single-cell RNAseq analysis of normal human liver cells (29).

Our approach to studying homeostatic dynamics in normal human liver centres on demarcation of clonal boundaries within these tissues. In contrast to diseased liver, clones within the typically quiescent normal human liver are often modest in size (30) however, intriguingly, clonal expansions are indeed detectable (31, 32). Lineage tracing studies using a visually recognisable deficiency of the mitochondrial enzyme cytochrome *c* oxidase (CCO) caused by acquisition of somatic mutations, have demonstrated the presence of clonally-related cells presumably derived from long-lived stem cells (32, 33). Studies utilising this technique in liver suggested a spatial association between the clonal patches and the PT, however, further evidence that the periportal region is a stem/progenitor niche in normal human liver, is lacking. Whether potential progenitor cells exist as $SOX9^+$ hybrid hepatocytes (34) or as a subpopulation of biliary epithelial cells, is currently unknown.

Mitochondrial DNA mutations clonally expand within cells by random intracellular drift and mitochondria do not require cell division for the generation of somatic mutations (35). Cell division facilitates the spread of mutant mitochondria through tissues and the rate of cell division determines the speed and the physical direction this takes (36). Should cell division be constant then large clones comprised of clonally-related cells will be detectable over time, however if cells within a clonal patch are quiescent, each will generate its own independent mitochondrial mutation repertoire not shared with its neighbours. In this study, we exploit this unique feature of mutation spread, using it to determine the behaviour of clonal expansions, as a proxy for cellular dynamics within the normal human liver.

Here, we demonstrate for the first time, the spatial dynamics and the origins of human hepatocytes *in vivo*. We utilise a range of techniques including CpG methylation sequencing, mitochondrial DNA next generation sequencing and mathematical modelling to establish their origins and the nature of their expansion. Collectively, our data support the existence of a periportal stem cell niche that feeds the clonal expansion of hepatocyte patches in normal human liver. These patches most likely have a common cell of origin with biliary epithelium and expand in a slow and punctuated fashion, with a portal to central trajectory, and are largely quiescent, potentially for decades thereafter.

## Results

### Characterisation of CCO-deficient clonal expansions

We previously established the use of cytochrome *c* oxidase (CCO) deficiency to investigate dynamics of clonal cell populations in human tissue including the liver (31, 32, 36). This method allows histochemical identification of clonal patches of cells that have inherited a lack of CCO activity from a long-lived cell of origin. Here, we have utilised this technique permitting the visualisation of blue, CCO-deficient clonal hepatocyte expansions, henceforth referred to as 'patches'. Observations from large-area samples reveal clones extending from the portal tree **(Fig. 1a)**. To characterise the patches in our cohort, we investigated their abundance with age. Detection of CCO-deficiency by enzyme histochemistry requires mitochondrial homoplasmy, or at least a high degree of heteroplasmy that takes decades to develop stochastically (35). Indeed, in our cohort, the number of patches observed per $mm^2$ of tissue increased significantly with patient age **(Fig. 1b;** *P=0.0017***)**. Despite CCO-deficiency, blue patches were previously determined to be metabolically indistinguishable from their adjacent, brown, CCO-proficient counterparts (31). We digitally counted the number of nuclei in blue, and equally sized neighbouring, brown areas and found the former had significantly more nuclei/area **(Fig. 1c;** *P=0.0067***)**. These nuclei were of similar size indicating a similar ploidy profile **(Fig. 1d).** Hepatocytes collected by laser-capture microdissection (LCM) from throughout the lobule had a 2.2-fold greater mitochondrial DNA (mtDNA) copy number in blue vs brown regions **(Fig. 1e;** *P=0.0009***)**. This is a previously reported compensatory mechanism that attempts to restore residual wild-type mtDNA to optimum levels (37). No mtDNA copy number differences were observed when comparing periportal vs pericentral hepatocytes within brown **(Supplementary Fig. S1 a&b)** or blue **(Supplementary Fig. S1 c&d)** regions. We also evaluated proliferation by Ki67 immunostaining, calculated as a percentage of total cell nuclei rather than by area due to the increased cell density in blue patches. As expected in normal liver, overall Ki67 positivity was low, with no significant difference in the proportion of proliferative cells in CCO-deficient patches vs adjacent CCO-proficient regions **(Fig. 1f)**.

To further determine if CCO-deficiency impacts hepatocyte biology, we performed Smart-3SEQ (3' RNAseq) on laser-capture microdissected areas taken from CCO-deficient and

proficient hepatocytes in periportal, centrilobular and pericentral regions. (**Supplementary Fig. S2a&b**). After correcting for patient-patient variability, we investigated the impact of patient therapy, the presence of patient metastases (i.e. not within our samples) and the effects of CCO-deficiency on gene expression (**Supplementary Fig. S3a-d**). It is important to note that our liver samples were obtained distal to any cause of resection (metastases or otherwise). Using principal component and unsupervised heatmap analyses, we found no informative clustering of samples based on therapy or the presence of patient metastases. We did observe a degree of clustering based on CCO-deficiency (**Supplementary Fig. S3c&d)**.

Hepatocyte zonation remained intact (**Supplementary Fig. S4a&b**) as we observed a host of previously reported genes with periportal zonation such as *HAL*, *C7*, *SDS*, *RAPGEF5*, *NAV2* and *LEPR* and pericentral zonation such as *SLCO1B3*, *CYP1A2*, *CYP3A4*, *GLUL* and *SLC1A2* (29, 38). We compared CCO-deficient with CCO-proficient microdissections using 'fast gene set enrichment analysis' and observed significant negative enrichment of the oxidative phosphorylation pathway in pericentral CCO-deficient samples as expected. A negative periportal enrichment was observed, however, this was not significant. The apoptosis pathway was significantly negatively-enriched within CCO-deficient samples of both comparisons, suggesting CCO-deficiency is not pro-apoptotic. The reactive oxygen species pathway, a crude indicator for localised tissue damage, had no significant alteration, irrespective of location within the lobule. We additionally compared all (periportal, centrilobular and pericentral) CCO-deficient and proficient microdissections and found both the oxidative phosphorylation and apoptosis pathways had significant negative enrichment. The reactive oxygen species pathway was not significantly enriched **(Supplementary Fig. S4c-e)**. In conclusion, our data broadly reflects previous observations that CCO-deficiency does not have major deleterious effects on hepatocyte biology.

We next sought to localise the source of CCO-deficient clonal expansions, particularly as stem cells or committed progenitors are thought to be sufficiently long-lived to acquire homoplasmy or the high level of heteroplasmy required for visualisation as CCO-deficient and thus be the cells of origin of CCO-deficient patches (31). Observationally, CCO-deficient patches appear to emanate from terminal portal tracts **(Fig. 1a)**. This propensity for periportal patch localisation was proposed in our previous investigations of histologically normal human

liver (31, 32), but never quantified. Here, we have determined the closest structure, either portal tract (PT) or central vein (CV), for 377 CCO-deficient patches of ~10 cells or greater in 17 normal human livers **(Fig. 1g)**. In these livers, an average of 76.9% of patches were located closer to PTs than CVs, whilst only 21.5% were closer to CVs. A much smaller proportion (1.6%) were sufficiently large to associate with both structures **(Fig. 1g)**. Patch localisation thus favours the presence of a periportal stem/progenitor niche in phenotypically normal human liver.

**Hepatocytes and cholangiocytes share a common somatic ancestor**

If CCO-deficient patches originate periportally, the biliary epithelium or nearby progenitor-like, hybrid hepatocytes (34) may well be a source. To investigate this possibility, we studied 54 patient samples and observed only three with CCO-deficient biliary epithelium and nearby deficient hepatocyte patches. In one such case, several CCO-deficient patches were found adjacent to terminal portal tracts that branched from a conducting portal tract containing the CCO-deficient biliary epithelia **(Fig. 2a&b)**. CCO-proficient **(Fig. 2b; Patch 1)** and deficient hepatocytes patches **(Fig. 2b; Patch 2&3)** as well as the nearby CCO-deficient biliary epithelium **(Fig. 2b)** were laser-microdissected and Sanger-sequenced. Two CCO-deficient biliary areas **(Fig. 2c-e)** shared a *m.8251G>A* mutation (in *MT-CO2* coding for subunit COII of cytochrome *c* oxidase) that was not present within in the stroma of other portal tracts from the same tissue **(Fig. 2 i&j)**. This mutation was also undetected within CCO-proficient areas of the biliary epithelium and hepatocytes within Patch 1 **(Supplementary Fig. S5)**. Two CCO-deficient areas tested within Patch 2 shared a *m.1214A>G* mutation (in *MT-RNR1*), but did not have the *m.8251G>A* mutation present within the CCO-deficient biliary epithelium (data not shown). However, the *m.8251G>A* mutation was present within CCO-deficient hepatocytes within Patch 3 and absent from CCO-proficient hepatocytes nearby it **(Fig. 2f-j)**. The chances of a common mutation independently arising in these two cell types is extremely low; $2.48\times10^{9}$:1 (32, 36). We were unable to determine clonal relationship within other samples with CCO-deficient biliary epithelium **(Supplementary Fig. S6a-d)** due to insufficient CCO-deficient cells in serial sections cut for LCM. Regardless, these data demonstrate a strong likelihood for the existence of a somatic periportal bipotential common ancestor of biliary

257 epithelial cells and hepatocytes. To our knowledge, this is the first such clonal demonstration
258 in normal human liver.

259

260 **SOX9$^+$ and Ki67$^+$ hepatocytes are present in normal human liver but SOX9$^+$ hepatocytes are**
261 **predominantly periportal**

262 Normal liver is largely quiescent, however should hepatocyte turnover in this tissue be
263 actively fed from periportal cells, this region may be more proliferative than the parenchyma
264 at the CV. We annotated all PTs and CVs within normal livers from 13 patients using the
265 pathology software QuPath, and we detected Ki67$^+$ cells within 50 µm expansions from the
266 boundaries of these structures **(Fig. 3a&b)**. We observed no significant difference in the
267 proportion of Ki67$^+$ cells at the PT vs CV **(Fig. 3c)**. We also investigated the possibility of a
268 SOX9$^+$ progenitor as a source of neo-hepatocyte generation in normal liver, irrespective of
269 CCO-status (34). As SOX9 is also expressed in KRT19$^+$ ductal cells, we performed dual IHC for
270 SOX9 (Brown; DAB) and KRT19 (Blue; Vector Blue AP) to enable their exclusion **(Fig. 3d&e)**.
271 QuPath was trained to differentiate between SOX9$^+$KRT19$^+$ ductal cells (combined blue &
272 brown immunostaining) and SOX9$^+$KRT19$^-$ hepatocytes (stained brown only) **(Fig. 3e&ei).** Blue
273 and brown digital detection masks overlay and highlight these detections within 50 µm
274 expansions surrounding each CV and PT **(Fig. 3eii)**. SOX9$^+$ hepatocytes were detected in
275 significantly greater abundance within 50 µm of PTs relative to CVs; a mean of 14.85 vs 5.89
276 cells/mm$^2$ **(Fig. 3f;** ***P<0.0001*****)**. Collectively, these results are in line with the observations of
277 Font-Burgada (34).

278

279 **Methylation diversity supports the proposal of a periportal stem cell niche**

280 We next sought to investigate the expansion dynamics of clonal patches in normal human
281 liver. For this, we used methylation diversity as a dynamic clonal marker – one that changes
282 with time and size of clonal expansion (39, 40). Methylation patterns record cell ancestry as
283 they are inherited with high fidelity by daughter cells. The greater the methylation diversity
284 between cells, the more distant their ancestral relationship (39). We used two markers of
285 diversity: intra-patch distance (calculated as the minimum number of changes required for
286 each sample to have the same sequence) and the number of unique methylation sequences

(tags). First, we measured methylation diversity within large periportal and pericentral regions without regard to clonality and found no difference in the epigenetic distance between groups **(Supplementary Fig. S7a-c)**. Next, CCO-deficient periportal, centrilobular and pericentral cuts were analysed from each patch and no differences were observed **(Supplementary Fig. S7d-f)**. If the streaming liver hypothesis (19) is correct, and hepatocytes migrate along the portal-central axis, the ancestral and epigenetic distance should increase with anatomical distance from the stem cell niche. Cuts were made within the patches, at varying distances from the PT along the portal-central axis **(Supplementary Fig. S7g-i)**, but no correlation was observed between epigenetic and anatomical distance. These data indicate the epigenetic diversity measurable using this strategy was maximised.

Smaller extractions, remaining close to the putative stem cell niche are more likely to remain within the limits of this diversity measure. To this end, we investigated methylation diversity in patches of increasing size that abutted the portal tract. Measuring diversity is heavily influenced by the number of cells analysed thus micro-dissections were size-controlled. Extractions were made within small, medium and large CCO-deficient patches directly abutting PTs **(Fig. 4a&b)** and CVs. Larger clones had multiple dissections, each the same size but analysed independently. Methylation tag diversity increased with patch-abuttal size at the PT **(Fig. 4c)**. A significant, positive correlation was observed between epigenetic distance and patch PT abuttal length **(Fig. 4d)**. No such correlation was observed for patches abutting the CV **(Fig. 4e)**. These data demonstrate an ancestral relationship between hepatocytes at the PT but not the CV, where it dissipates due to the number of cell divisions taken for clones to expand to the CV.

We hypothesise there are no differences in expansion dynamics between CCO-deficient and proficient patches. Indeed, we found no significant difference in the % methylation of our CpG sites between CCO-deficient and proficient cells **(Supplementary Fig. S7j)** indicating patch dynamics are similar. Given the relationship determined within known clonal boundaries, we also measured methylation diversity in size and spatial location-matched CCO-proficient PT and CV abutting areas. Intra-patch distance vs increasing abuttal size were non-significant and more weakly correlated than their CCO-deficient counterparts **(Supplementary Fig. S7m&n)**. This observation demonstrates the benefit of examination within known clonal boundaries, which are otherwise difficult to determine in normal liver.

Similar overall results were achieved using another measure of diversity; the simple number of unique methylation patterns or 'tags' observed within a given extraction **(Supplementary Fig. S7c, f, i, k, l, o & p)**.

## Clonal hepatocytes display large numbers of unique mitochondrial DNA variants across PT-CV axes

Our data thus far indicated the presence of a periportal progenitor niche being the origin of CCO-deficient patches, however the spatio-temporal dynamics of these expansions could not be determined with the limited resolution of our methylation assay. To address this, we performed deep LCM NGS of mtDNA genomes, a technique with far greater resolution, enabling analysis across the whole patch. In a similar fashion to methylation diversity, mtDNA mutations can be exploited to determine the relatedness of cells across a patch. As mitochondria will continue to replicate in the absence of cell division, mutations can accumulate even in largely quiescent tissues such as the normal liver.

If hepatocytes continuously stream from their cell of origin, we hypothesise that we would detect a distinct stepwise pattern of inheritance of mtDNA variants across each patch. Older cells, hypothetically at the CV, would have had more time to develop independent mitochondrial variants compared to cells at their PT origin. Faster streaming would see an increase in public (clonal) variants detected across the whole patch as there is less time for large proportions of independent variants to develop. Alternatively, should patches grow in a discontinuous or punctuated fashion, with intervening quiescence, we would observe a large degree of independent variants across the whole patch.

Four to five 'cuts' of size-matched hepatocytes were laser microdissected along the PT-CV axis in CCO-deficient patches that bridged PTs and CVs **(Fig. 5a).** After DNA extraction, mtDNA genomes were sequenced by NGS and the variants within each cut were called. In total, 68 cuts across 14 PT-CV axes were extracted from 10 patches across five patient samples. The loci of all PT and CV variants were also plotted for each patch and no discernible differences were observed **(Fig. 5b).** Individual patch data are shown in **Supplementary Fig. S8-10**. The mutation types for cuts adjacent to PTs and CVs were compared to investigate whether differences between these locations could be observed **(Fig. 5c)**. Overall, the base

348 substitutions were very similar, with the T>C substitution predominant in cuts from both
349 locations. This substitution has been associated with aging (41), a phenomenon coinciding
350 with the development of the large patches used for this analysis **(Fig. 1a)**.

351 Most PT-CV axes contained a large percentage of unique private variants; (those not
352 shared by neighbouring pairs of cuts) **(Fig. 5a)**. By virtue of their shared CCO-deficiency, all
353 cells within a given patch share a common ancestor, yet these data indicate cells in each cut
354 are largely genetically distant from their physical neighbours. We propose that this pattern is
355 suggestive of punctuated expansion of clonal patches, with extensive periods of quiescence.
356 This would allow sufficient time for the cells in each cut to develop variants that are unique
357 when compared to their neighbours **(Fig. 5d).**

358 Additionally, we sequenced CCO-proficient areas, some directly adjacent to the CCO-
359 deficient areas previously examined, to determine whether CCO-deficiency impacted on the
360 variant composition of CCO-deficient clones. We observed no notable differences in the
361 overall distribution of variants **(Supplementary Fig. S11a)**. We did, however, observe a
362 change in the variant signature; while T>C variants remained dominant, T>A variants showed
363 an increase in CCO-proficient areas relative to their deficient counterparts **(Supplementary**
364 **Fig. S11b)**. Furthermore, we also observed an increase in the overall number of unique
365 variants in CCO-proficient areas **(Supplementary Fig. S11c)**. This can be explained by
366 multiclonal sampling due to an inability to discern clonal boundaries. Crossing such
367 boundaries also drastically reduced the number of clonal variants detected **(Supplementary**
368 **Fig. S11d)**.

369

370 **Spatial modelling supports neo-hepatocyte generation from a periportal niche.**

371 We built a spatial simulation model of liver tissue dynamics to explore scenarios that further
372 test our hypothesis of a periportal niche. We compared all possible mtDNA spatial patterns
373 predicted under different model assumptions with the patterns observed in NGS data **(Fig.**
374 **5a)**. We explored two modes of clonal expansion: homeostatic cell turnover dominated by
375 proliferation of randomly distributed 'non-PT' parenchymal cells or, alternatively, faster
376 replication by a small number of stationary 'PT cell' progenitors at the centre of the model
377 **(Fig. 6a)**. These two regimes are described by 'β', with the value of $\beta<1$ and $\beta>1$ respectively.

Rounds of homeostasis take place characterised by the death of non-PT cells with subsequent replacement and mitochondrial mutation **(Fig. 6b)**. Our model explored randomly distributed non-PT death as well as a CV-centric cell killing model. In both, cell death leaves open space, allowing for occupation during cell replacement. Diving cells displace neighbours into nearby empty spaces, made available by prior cell death, to create space to divide **(Fig. 6c).** Cell movement in our model is thus passive, generated via our displacement mechanics, with no active migration. A traditional straight-line cell-pushing algorithm ultimately produced very similar dynamics to this adapted algorithm **(Supplementary Fig. S12a)**. After each round of homeostasis, non-PT cells in discrete zones along a random radial axis are sampled to record their mutational burden **(Fig. 6d)**. As mutational burden depends on the number of cells sampled ($N_s$) per zone, we used $N_s = 60$ in accordance with the average cells numbers in each cut of our NGS experiment. Our model also captures the phenomenon of mitochondrial mutations without cell division. We used a literature based mitochondrial mutation rate of 0.1 per cell division (42, 43). Further details are in **Methods** and **Supplementary Methods S1**.

Each simulation was run over 100 rounds of homeostasis under various combinations of β and number of PT cells ($N_{PT}$). To compare our simulated zonal mtDNA mutation patterns with those along each PT-CV axis obtained in our experimental NGS data **(Fig. 5a&b)**, Approximate Bayesian Computation (ABC) rejection sampling was implemented. This was applied to estimate the joint posterior distribution of the β and $N_{PT}$ parameters that best simulate our experimental NGS data. Approximately 2,000 β and $N_{PT}$ parameter pairs were sampled from the posterior distribution for each experimental sample, and geometric medians were used as a point estimator for the best-suited combination of parameters **(Supplementary Fig. S9-10)**.

With randomly distributed non-PT death, our experimental samples were best fit by $\beta>1$, signifying the presence of small numbers of static PT progenitors that replicate faster than the randomly distributed parenchymal cells **(Fig. 6e)**. A low sensitivity to the $N_{PT}$ parameter in our model, meant we did not infer the number of PT cells. With $N_{PT} = 10$, we compared these results to our CV-centric killing model. At all β values investigated, resulted in a greater total variant burden compared with the prior, homogeneous non-PT death model. Importantly, both models converge at below $\beta \approx 5$. Given we predict values of $1 < \beta < 5$ for the majority of our experimental NGS samples, this implies a 'PT'-dominant expansion niche

is present irrespective of the distribution of cell death **(Supplementary Fig. S12b)**. Collectively, this modelling supports our hypothesis for the generation of CCO-deficient patches from a confined niche and not solely from randomly distributed hepatocytes.

**Clonal patches expand slowly and quiesce.**

We further tested our expansion dynamics hypotheses by simulating the distribution of mtDNA variants across a simpler, 1-dimensional PT-CV axis **(Supplementary Methods S2)**. The simulation takes place over two phases; first, a clonal expansion occurs either slowly or rapidly to populate the axis. As was determined in our prior model, generation of new hepatocytes by the PT-cell positioned at the simulated portal tract occurs more rapidly than by randomly distributed hepatocytes along the axis. At the conclusion of clonal expansion, a second phase consisting of either hepatocyte quiescence, or streaming occurs **(Fig. 7a&b)**. As the duration between clonal expansion and observation of the patch within the biological samples is unknown, the second phase is simulated across various time periods (1 month, 5 years, 10 years and 20 years). During simulated quiescence all cells stop dividing, however, mtDNA mutation persists as they are capable of replication in the absence of cell division. Under streaming dynamics, PT-cell replication persists, and as new hepatocytes are generated, those at the simulated central vein are pushed out and lost from the system, akin to an active conveyor belt of hepatocytes. Extensive methodological details and parameters are located in **Supplementary Methods S2**.

At the conclusion of both phases, the axis is divided into 5 sections of equal cell number, mimicking the patch cuts in the experimental data **(Fig. 5a)**. Unique private variants are detected within each section and compared in neighbouring pairs as in our biological data. In our simulations, regardless of phase 2 dynamics, the proportion of unique variants in each pair of sections approaches 1 (100%) as the time since phase 1 (clonal expansion) increases **(Fig. 7a&b)**. Interestingly, quiescent dynamics **(Fig. 7a)** serve to decrease the proportion of public variants (those shared across all sections) over time whilst a conveyor-like streaming dynamic **(Fig. 7b)** achieves the opposite in this respect. In all instances, an initial rapid expansion appears to accelerate the trend formed under a slow expansion system. Whilst there is some variability in the unique variant plots obtained from our NGS data **(Fig. 7c)**, most

approach 100% unique variants across the patch. This is most closely represented by longer phase 2 (5-20 years) simulations. Our biological data has public variant proportions largely <10%, which is most consistent with quiescent phase 2 dynamics. To determine the simulation that most closely resembled our average NGS data (**Fig. 7c**; black line & boxplot), we found the simulated scenario which minimised the Euclidean distance to the average NGS data with respect to the unique private variants curve and public variant proportion. A slow expansion with 20 years quiescence was determined to be the best fit for both the unique private variants and public variants data. As such, for the average patch, this modelling lends support to a slow/punctuated clonal expansion with long periods of quiescence.

## Discussion

There is a broad understanding that injury-related expansion dynamics are different to those within normal homeostatic liver, however, these dynamics are poorly understood. The current evidence is contentious and contradictory, with few studies investigating human livers. Under disease conditions, large clones are known to dominate, but these are smaller in normal liver (30). As such, investigation of normal liver dynamics requires a means of accurately revealing clonal boundaries. We achieve this by visualising clones that share an inherited CCO deficiency. Here we uncover an informative spatial bias of these clonal expansions and within their confines, we measure gene expression, methylation diversity and mtDNA variants, gaining insights into their ancestry. We pair these findings with mathematical models and Bayesian inference to identify the most likely hepatocyte origins and expansion dynamics within our ostensibly normal human liver samples.

Some commentators have been dismissive of the possibility of facultative stem cells, particularly within the biliary epithelial population of the liver (16)**.** Models of hereditary tyrosinaemia type1 have recorded the massive capacity for expansion of transplanted hepatocytes in such severely damaged livers (4), while similar large scale hepatocytic repopulation abilities have been noted in livers severely poisoned by a hepatotoxic transgene targeting the liver in albumin-urokinase-type plasminogen activator (Alb-uPA) mice (44, 45). More recently, studies employing dietary models of liver damage combined with cell lineage tracing, found no evidence of a significant input of biliary-derived cells to the resultant re-

population. These investigations led to the conclusion that hepatocyte self-duplication was the sole mechanism responsible for liver regeneration after injury (16). However, all these models failed to recapitulate the seemingly key role of hepatocyte senescence in determining the nature of the regenerative response (46). At least in liver regeneration after injury, a non-competitive environment is required for biliary cells to contribute to hepatocyte replenishment as seen when Mdm2 (negative regulator of p53) is deleted in hepatocytes (15), or when p21 is over-expressed in hepatocytes (13) or when β-catenin is deleted specifically in hepatocytes (47). Such findings suggest that under certain circumstances the liver can be considered as a maturational lineage system emanating from the periportal area, similar to the lineage systems prevailing in other epithelia, notably the intestine (48) and epidermis.

These observations beg the question as to whether a similar compartmentalization operates under more homeostatic conditions, where the requirement for hepatocyte replacement is perhaps less urgent? Indeed, such an organisation was first proposed 30 years ago based upon gradients in cell size, ploidy, growth potential, gene expression and matrix deposition (49, 50). Further studies revealed that small cells with stem cell attributes such as clonogenicity resided in the proximal biliary tree (canal of Hering), and might be the origin of biphenotypic transit amplifying hepatoblasts and subsequently hepatocytes (51, 52).

There is evidence for homeostatic hepatocyte turnover from specific liver zones (17, 21, 23) as well as from non-zonal, randomly distributed hepatocytes (24, 25). In our study, CCO-deficient patch location, $SOX9^{+}$ hepatocyte proximity to the PT, methylation diversity analysis and spatial modelling all support the existence of a periportal progenitor niche for the origin of clonal hepatocyte expansion. We demonstrate clonal patches have a periportal spatial bias, suggesting this as their origin. As CCO-deficiency likely initially develops in long-lived stem/progenitors (32), this localisation likely designates a periportal progenitor niche. In support of this niche, we demonstrate a clonal, common somatic ancestor of the biliary epithelium and hepatocytes. This finding raises the likelihood of a periportal bipotential progenitor as a source of clonal hepatocyte expansions. We also identify the presence of putative $SOX9^{+}$ $KRT19^{-}$ progenitor-like hepatocytes (34, 53) in greater abundance immediately surrounding portal tracts. Using methylation as a tool to determine ancestry, we were able to establish a relatedness of hepatocytes at the PT that was absent at the CV, again indicative of the PT as a site of origin for clonal expansions. Furthermore, the pattern of mtDNA SNVs

detected across PT-CV axes within clonal patches best recapitulate our computational modelling scenario of portal progenitor contribution to expansion, as opposed to a system dominated by proliferation of randomly distributed hepatocytes.

In addition to determining the origins of hepatocyte clonal expansions, our experiments offer insights on the temporal nature of clonal expansions. Changes in methylation are realised only through cell replication over time (40). This technique, in theory, can then also inform on expansion dynamics. Should a patch expand, and stop, there is insufficient cell replication for the methylation pattern to appreciably alter, thus the diversity of patterns observed across the patch should remain low. Subsequent quiescence, as is generally observed in normal liver, would ensure that the methylation diversity remains largely unaltered. A continuous streaming scenario would involve a greater amount of replication, in theory allowing differences to be observed between the cell origin and terminus. Unfortunately, our methylation methodology did not have the resolution to distinguish these dynamics. This could be improved using methylation systems that allow a greater number of CpG sites to be sampled within a given target promoter. Regardless, our methylation technique was beneficial for its support of periportal clonal origins as mentioned.

By contrast, mitochondria will replicate in the absence of cell division and develop SNV diversity over time, even within quiescent cells. Patch expansion followed by lengthy quiescence allows even neighbouring clonal cells to largely individualise their mtDNA variant profiles, greatly reducing the proportion of public variants (those present across the whole PT-CV axis). Alternatively, a conveyor-like streaming liver model (19, 34, 48) serves to greatly increase the proportion of public variants over time. Both these hypotheses are confirmed by our simulations. Within our experimental data, the average patch has a low level of public variants with largely individualised variant pools, which is best matched by our long-term (~20 years) quiescence and slow patch expansion simulations. The slow patch expansion scenario translates to a periportal cell replication rate approximately 3 times greater than that of randomly distributed hepatocytes in the 1-dimensional model, a value broadly consistent with the Bayesian estimates from our 2-dimensional modelling. An untested but plausible scenario that may still explain our data would be punctuated streaming with long periods of intervening quiescence. Furthermore, limiting the simulations to 1 dimension will likely accentuate the effects of streaming dynamics. If hepatocytes do exhibit streaming *in vivo,*

they would not always strictly travel along the 1-dimensional PT-CV axis. Last, whilst lattice and off-lattice models of diseased livers have been previously compared, how well a lattice approximation recapitulates the dynamics of a highly ordered normal human liver, remains unexplored. **Supplementary Methods S1.3** contains a full discussion of such model limitations. These factors may explain some of the discordance between experimental measurements and the simulated data generated under streaming dynamics, as well as variance observed at an individual patch level.

Questions remain as to the cause of these patch expansions. Whilst we have observed differences in expression, mtDNA copy number and nuclear density in CCO-deficient clones compared to proficient counterparts, these are not expected to cause a selective advantage (35). As such, it is reasonable to assume that these clonal expansions are a liver-wide phenomenon; not restricted to regions made visible by CCO/SDH staining. Our methylation data indeed indicate that CCO-proficient and deficient hepatocytes are equally replicative. Additionally, CCO-proficient and deficient hepatocytes are indistinguishable in terms of Albumin, cytochrome P450 1A2 immunostaining and Ki67 positivity (31). Given we are analysing histologically normal tissue, these clonal expansions could arise from homeostatic processes and our data support long-term quiescence as is expected in normal tissue. Whilst histologically normal, our tissues could have been affected by the conditions that necessitated their resection. The majority, but not all of our tissues, were obtained from patients undergoing resection for metastases and not all patients received systemic therapy. Crucially, this allowed us to examine whether these factors had a significant influence on the biology of our hepatic tissues taken distally from the cause for resection. Our gene expression data indicated no appreciable contribution of these potentially confounding factors on the biology of our samples. It does, however, remain possible that clonal patches result from repeated bouts of acute injury from environmental or lifestyle exposures/stimuli however, chronic injury is usually required for new hepatocytes to extend far beyond their niche. Examining clonal patches from younger donors who have yet to experience repetitive hepatotoxic injury could in theory shed further light on this matter, however normal tissue from this age cohort is exceptionally rare. Additionally, our lineage tracing methodology, combined with our selection criteria of patches bridging the PT and CV, by its nature, predisposes our analyses to older patients who have more abundant patches, but are unlikely to have had a life free from

damage. Indeed, the youngest donor used in our NGS analyses was 62 years of age at time of resection.

Hepatocyte clonal dynamics in humans is understudied, especially within normal liver, yet it is essential as a baseline for comparison to disease dynamics. Furthermore, identification and characterisation of a progenitor niche could ultimately be therapeutically beneficial, particularly since hepatocyte transplantation in humans has met with only modest success (54). We have demonstrated the presence of a periportal progenitor niche in phenotypically normal human liver and have provided evidence that neo-hepatocyte expansions spread from this niche. Further work is required to understand what initiates these expansions and how their dynamics may be altered.

# Methods

## Samples

Human liver samples were obtained by informed consent from patients undergoing hepatic resection. Samples were obtained through the Queen Mary University of London Cancer Tissue Bank (REC reference 14/LO/2031 renewed 19/LO/1700, more information on *https://www.cancertissuebank.org*, and project approval 2015/01/QM/SM/E/FrozenTissue) and University of Cambridge (16/NI/0196) and University of Padova, Padova, Italy (IRB CESC 3312/AO/14). Additional clinical details can be found in **Supplementary Table 1.**

## Preparation of tissue

Normal liver, as identified by a pathologist, was taken from each resection, distal to the metastatic or diseased tissue. Resections were cut to size, submerged in isopentane and frozen in liquid nitrogen. Frozen samples were either sectioned onto UV-treated P.A.L.M. membrane slides (Zeiss, Cat # 415190-9041-000) or charged glass slides. Tissues and slides were stored at -80 °C until required.

## Immunohistochemistry

Dual immunohistochemistry was performed on frozen tissue sections using anti-SOX9 (ABCAM, Cat # ab185966) and anti-KRT19 (DAKO Cat # M0888) according to standard procedures. Anti-Ki67 (ABCAM, Cat # ab16667) immunostaining was performed using the Discovery Ultra Ventana apparatus. A detailed methodology can be found in **Supplementary Methods S3**.

## Enzyme histochemistry

Enzyme histochemistry to identify clonal cytochrome *c* oxidase (CCO) deficient patches/cells was performed as previously described (55). Enzymatic chromogen development of DAB and NBT for CCO and SDH activity were typically halted after 50 and 25 min respectively.

599

600 **CCO patch-structure association and digital cell counting**

601 Serial sections stained with H&E or for CCO/SDH, SOX9/KRT19 or Ki67, were scanned with a
602 NanoZoomer S210 (Hamamatsu). Slides were opened and analysed using detection tools and
603 scripts within QuPath. See **Supplementary Methods S4** for a full description of this
604 methodology.

605

606 **Laser capture microdissection**

607 Laser capture microdissection (LCM) was performed on 10 μm CCO/SDH-stained cryosections
608 cut onto UV-treated P.A.L.M. membrane slides. Prior to microdissection, slides were
609 completely thawed and air-dried, then tissue was captured in specialised laser-capture tubes
610 (Zeiss, Cat # 415190-9201-000) using a P.A.L.M. laser dissection microscope (Zeiss).

611

612 **Mitochondrial copy number analysis**

613 Tissue and DNA were extracted as per the **mitochondrial DNA NGS** methodology below, using
614 the same tissue cut sizes and quantities. The Human Mitochondrial DNA Monitoring Primer
615 Set (Takara Bio, Cat # 7246) was used to determine mitochondrial DNA (mtDNA) copy number.
616 Methodology and analysis was conducted as per the manufacturers' instructions.

617

618 **Mitochondrial Sanger sequencing**

619 Single hepatocytes or bulk stroma were subjected to LCM and tissue was lysed according to
620 the Arcturus® PicoPure® DNA Extration Kit protocol (Thermo Fisher Scientific, Cat # KIT0103).
621 Amplification of the mitochondrial genome was performed as previously reported (56). PCR
622 products were Sanger sequenced by Eurofins Genomics. Sequence files were aligned against
623 the revised Cambridge reference sequence to identify mutations. Visualisations and
624 comparisons were performed using the "A plasmid Editor" software available at
625 *https://jorgensen.biology.utah.edu/wayned/ape/*.

626

627 **DNA methylation**

628 DNA Methylation data for the CpG islands in the promoter region of Cardiac-specific
629 homeobox (*CSX*) were generated as described previously (40).

630

631 **3' RNA Seq**

632 Periportal, pericentral and centrilobular regions of the liver were laser-microdissected from
633 CCO/SDH-stained frozen liver sections. Libraries were prepared using the 'fresh-frozen'
634 variant of the Smart-3SEQ protocol (57), with slight modifications to the protocol including;
635 the use of 7xN Unique Molecular Identifiers (UMI), an 8 nt index for increased plexity, as well
636 as a 30 min lysis time and a 2 µM 1S primer. The libraries were pooled and single-end
637 sequenced on the Illumina NextSeq 2000 platform using a P2 kit for 100 cycles.

638 Adapter and polyA trimming were performed using Trimmomatic. Sequences were trimmed
639 to 75 nt then, using umi_homopolymer, UMIs were extracted and appended to the read
640 metadata and G-overhangs were discarded. Alignment was performed using STAR and
641 deduplication, using UMI_tools. Differential expression was determined using DESeq2 and
642 pathway enrichment was performed using FGSEA.

643

644 **Mitochondrial DNA NGS**

645 Four to five areas in a portal-central line were laser-capture microdissected from large
646 CCO-deficient patches that spanned from the portal tract to the central vein. DNA was
647 extracted, then PCR amplification of mitochondrial genomes and subsequent library
648 preparation was conducted according to the Human mtDNA Genome Guide (Illumina,
649 15037958). Libraries were sequenced on the Illumina MiSeq v2 or NextSeq 500 mid output
650 platforms. A full description of the mitochondrial DNA NGS library generation and data
651 processing can be found in **Supplementary Methods S5.1**.

652

653 **Mitochondrial DNA NGS data processing**

Data quality was assessed by FastQC and reads were aligned to the revised Cambridge Reference Genome (GenBank accession number NC_012920.1). Single-nucleotide variants (SNV) were identified using the deepSNV R package (58). DeepSNV variants with < 10 supporting reads in the test sample were filtered. Germ-line mutations, defined as those present in both test duplicate samples and in the control at a frequency of > 1%, were removed. Variants not called independently in both technical replicates were discarded. Public variants were defined as those present in all samples along the portal-central axis of a CCO-deficient patch. Conversely, private variants were defined as those which were called in a subset of samples along the portal-central axis. A full description of the DNA NGS data processing can be found in **Supplementary Methods S5.2**.

**Mathematical Modelling and Bayesian Inference to integrate model and data**

We investigate spatial patterns of mtDNA mutation in a simulated 2-dimensional liver lobule, comprised of "periportal "PT" cells and hepatocyte-like "non-PT" cells. We assume three alternative scenarios, where PT cells divide slower, faster and at the same rate compared to non-PT cells. The rates of PT and non-PT cell division are related by $r^{(PT)} = \beta \cdot r^{(non-PT)}$, where $\beta \geq 0$ is a model parameter. When $\beta=1$, both PT and non-PT cells divide at the same rate and there are no special niches for PT cells. The regime $0 \leq \beta < 1$ leads to "non-localised" expansion with the mutation accumulation over space dominated by the non-PT cells. Values of $\beta > 1$ lead to "localised" dynamics, where the mutation spatial pattern is driven by the expansion of PT cells in the lattice centre, which corresponds to the existence of periportal progenitor niches.

By modelling these alternative hypotheses under a range of β values and simulating the mtDNA mutations in cells located in space, we can predict different possible spatial patterns of mutation accumulation under different hypotheses. Finally, we compare our spatial patterns estimated by multiple simulations with our NGS data and use Bayesian statistics to infer the most likely range of β for each sample. Detailed methodology can be found in **Supplementary Methods S1**.

## Data availability

Sequence data has been deposited at the European Genome phenome Archive (EGA; *https://ega-archive.org*), which is hosted by the EBI and the CRG, CCO-deficient and proficient datasets are archived under the accession numbers EGAD00001007991 and EGAS00001006962 respectively. RNA-sequencing data is archived under accession number EGAS00001006984. Pipelines and code are available at *https://github.com/MagnusHaughey/liverMitoDNAPipeline*. Additional data that support the findings of this study are available from the author S.A.C.M. upon reasonable request.

## Statistical Analyses

Two-group individual gene expression statistics were performed with the non-parametric Wilcoxon test using the R package 'ggpubr'. All other two-group analyses were performed within Prism (Graphpad) using the unpaired, non-parametric Mann-Whitney U test. Correlation analyses were Statistics calculated using a Pearson correlation test. Gene set enrichment p-value statistics were Benjamini-Hochberg adjusted. Statistics relating to computational modelling are described within the supplemental materials.

## Acknowledgements

We are grateful to the Barts Health NHS Trust Cancer Tissue Bank (*https://www.cancertissuebank.org*) operational team (Amina Saad, Thomas Dowe, Ahmet Imrali, Abhirup Banerjee, Claude Chelala, Sameena Iqbal, Dayem Ullah, Rory Smith, Catherine Graham, Sarah Mueller) and Tissue Access Committee (Satyajit Bhattacharya, Chairman) for their support in this project. This study was funded through the Belgian Federal Science Policy Office (BELSPO) -HEPRO-2 (MRA), Medical Research Council (MRA/SM; G0901178), Cancer Research UK Barts Cancer Centre funding (C16420/A18066) and the Queen Mary University Life Science Initiative (HW).

# Figures

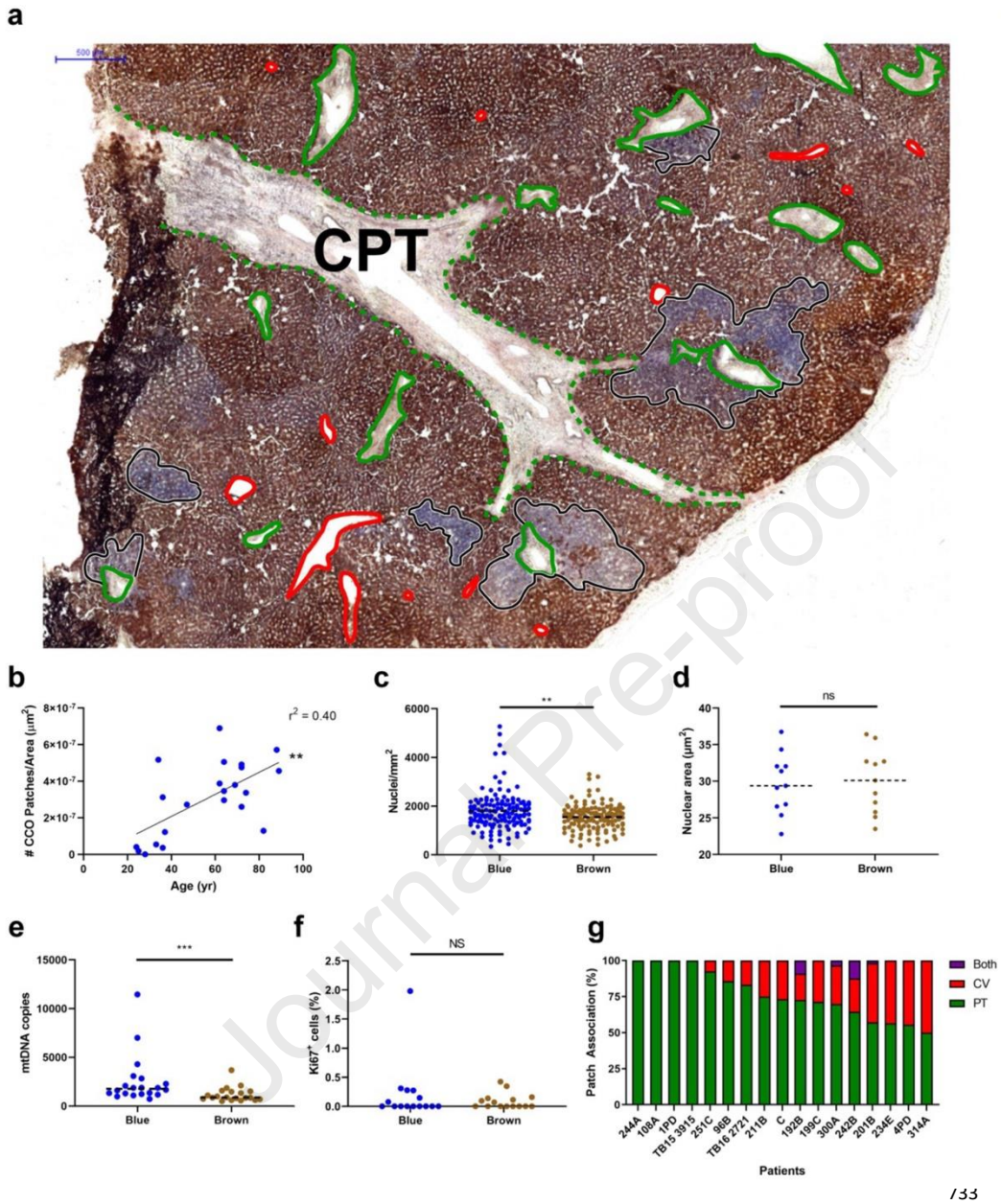
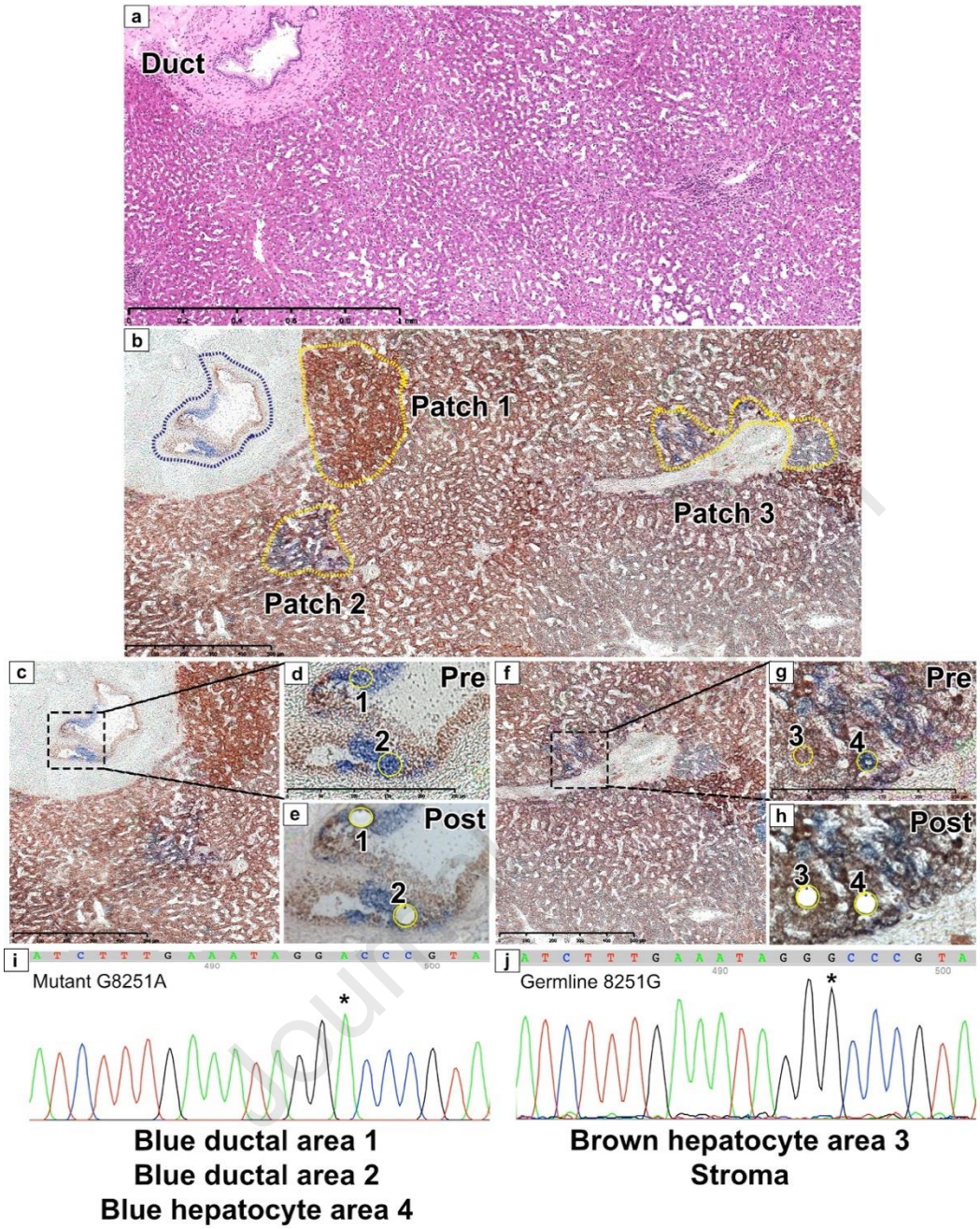



**Figure 1. Hepatocyte clonal expansion characteristics. (a)** A low-power image of a CCO/SDH-stained liver section with outlined examples of CCO-deficient cell 'patches' (black border) seemingly emanating from (green) that branch from the conducting portal tract (CPT; green dashed). Central veins are highlighted in red. **(b)** The number of patches normalised by tissue area vs patient age at resection; n=22 patients and 688 patches, *** $P$<0.001; testing by Pearson correlation coefficient. **(c&d)** The number and size of nuclei within CCO-deficient and proficient patches. Data in **(c)** are from 4 patients, n=130 patches, **$P$<0.01 and from 11 patients for **(d)**. **(e)** The relative number of mtDNA copies in CCO-deficient and proficient hepatocytes; n=20 microdissections, ***$P$<0.001. **(f)** The percentage of CCO-deficient and proficient cells that are Ki67$^+$; n=13 patients. **(g)** The % of patches in a given patient that were closer to PTs, CVs or both structures; n=17 patients. Statistical tests were performed using Mann-Whitney unless otherwise stated.




746

747 **Figure 2. An ancestral relationship between biliary epithelium and hepatocytes**. **(a)** An H&E
748 and **(b)** serial CCO/SDH-stained section highlighting three microdissected regions (Patch 1-3;
749 yellow outlines) and a partially CCO-deficient bile duct (blue outline). **(c)** The partially deficient
750 bile duct with corresponding high-power images pre **(d)** and post **(e)** laser microdissection. **(f)**
751 Patch 3 with corresponding high-power images pre **(g)** and post **(h)** laser microdissection. **(i&j)**
752 Sanger sequencing electropherograms from distal stroma and laser microdissected regions
753 outlined in **d&e** and **g&h**. ***** designates the location of the *m.8251G>A* mutant.

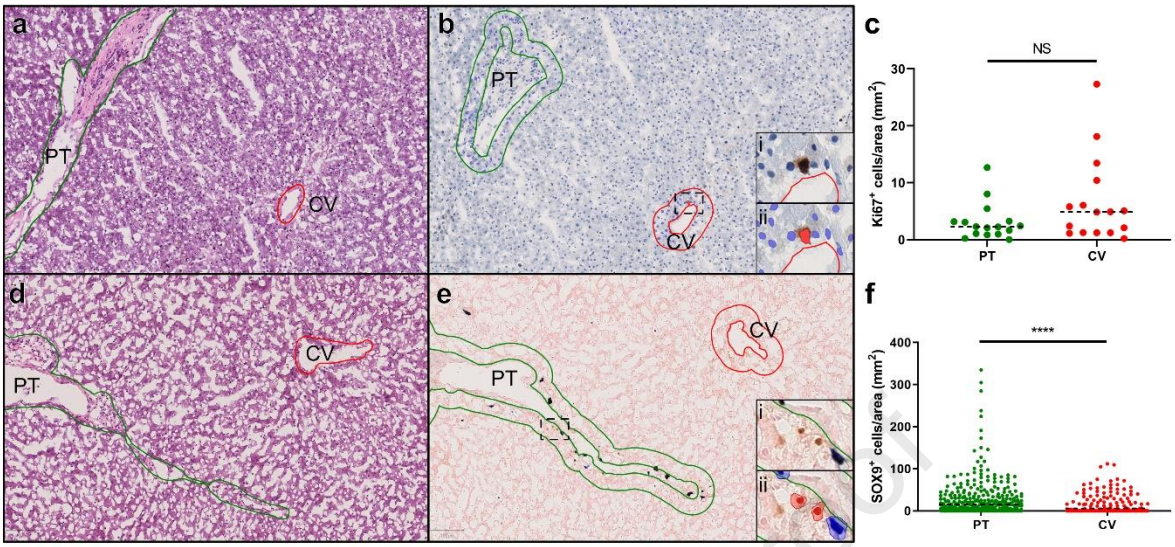



**Figure 3. SOX9+ hepatocytes are present in normal human liver and are predominantly periportal.** **(a&b)** A representative H&E and a serial section immunostained for Ki67. **(c)** Quantification for $Ki67^+$ cells. **(d&e)** A representative H&E and a serial section dual-immunostained for SOX9/KRT19. **(f)** Quantification for $SOX9^+KRT19^-$ hepatocytes. **(a-e)** Cell detections took place within 50 µm expansions of portal tracts (PT; green) and central veins (CV; red). **(bi)** Inset showing a single $Ki67^+$ cell and **(bii)** a red digital cell detection mask delineating this cell. $Ki67^-$ cells are digitally masked in blue. **(ei)** Inset of two $SOX9^+KRT19^-$ hepatocytes and two $KRT19^+SOX9^+$ biliary epithelial cells. **(eii)** Blue digital masks delineate non-hepatocyte $KRT19^+SOX9^+$ biliary epithelial cells and red masks highlight the $SOX9^+KRT19^-$ hepatocytes from **(ei)**. **(c)** Density of $Ki67^+$ cells within 50 µm PT and CV expansions from n=16 patients; NS=not significant. **(f)** Density of $SOX9^+$ hepatocytes within PT and CV boundaries from 6 patients for n=602 and 430 PT and CV respectively. Bars represent means and *P*****<0.0001; Mann-Whitney.

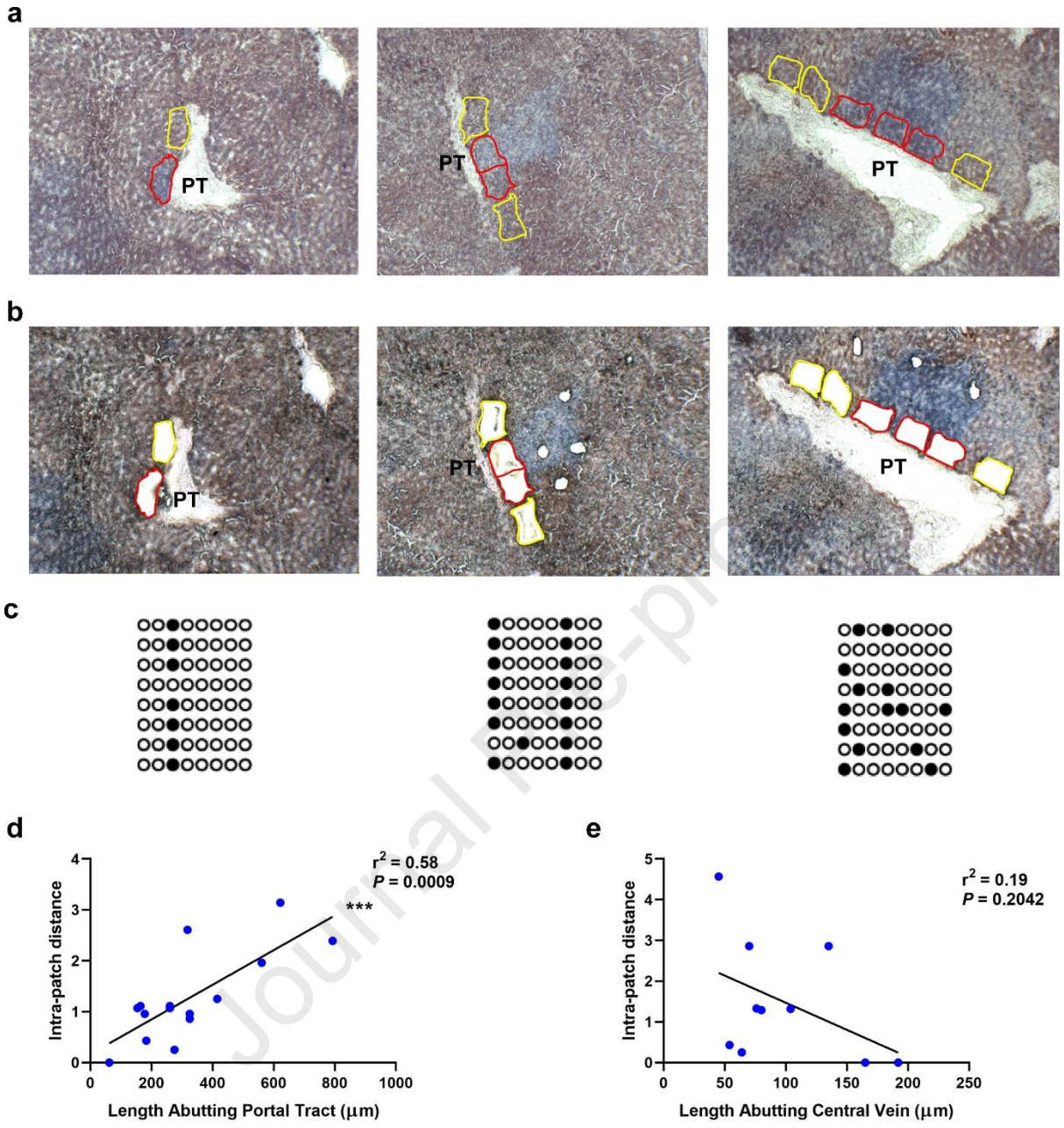



**Figure 4. Methylation diversity demonstrates a periportal hepatocyte origin. (a&b)** Left to right; representative cytochrome *c* oxidase-deficient patches of increasing abuttal length, before **(a)** and after **(b)** laser-capture microdissection. Red and yellow outlines highlight the equally-sized CCO-deficient and CCO-proficient regions captured. **(c)** Representative lollipop diagrams for small, medium and large patch-abuttal length. Columns represents a CpG site and row represents a sequenced cell. Solid circles denote CpG methylation. **(d&e)** Methylation diversity vs patch abuttal length correlations; n=15 portal tracts and 10 central veins. Statistics calculated using Pearson correlation test.

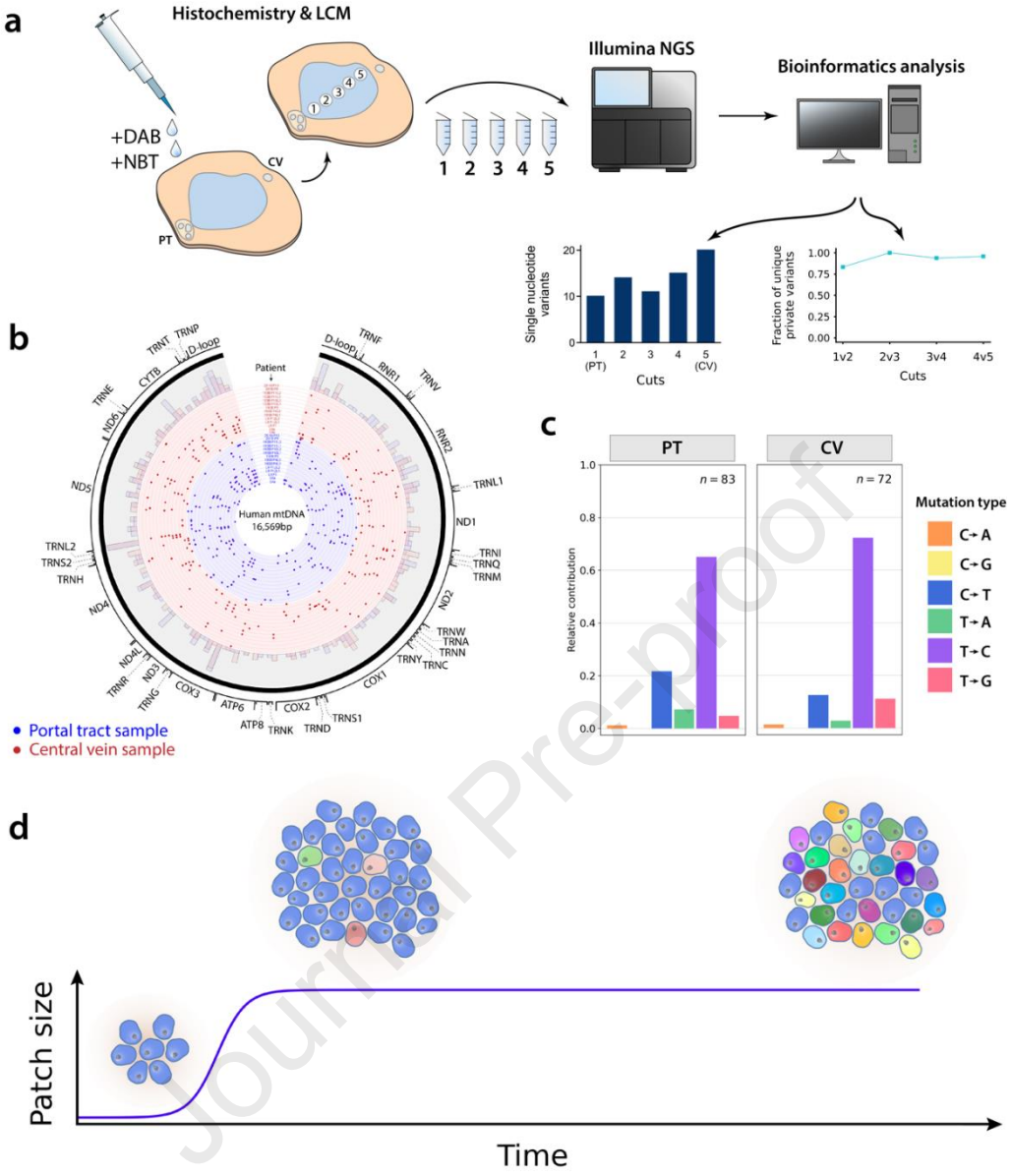



**Figure 5. Next-generation sequencing of CCO-deficient hepatocytes**. **(a)** CCO/SDH immunostained sections are laser microdissected across the portal tract (PT) to central vein (CV) axis. Samples are sequenced and analysed for single nucleotide variants (SNV). **(b)** Distribution of SNV across mt-genome. Blue and red data points represent PT and CV-abutting samples respectively. Outer histogram displays combined SNV data for all PT and CV samples (100bp bin width). **(c)** SNV distributions in PT and CV samples. **(d)** Proposed model of hepatocyte clonal expansion in normal human liver. A patch of clonal hepatocytes expands, then quiesces for extended periods. Over time, cells acquire unique mtDNA variants (different cell colours) independently of cell division.

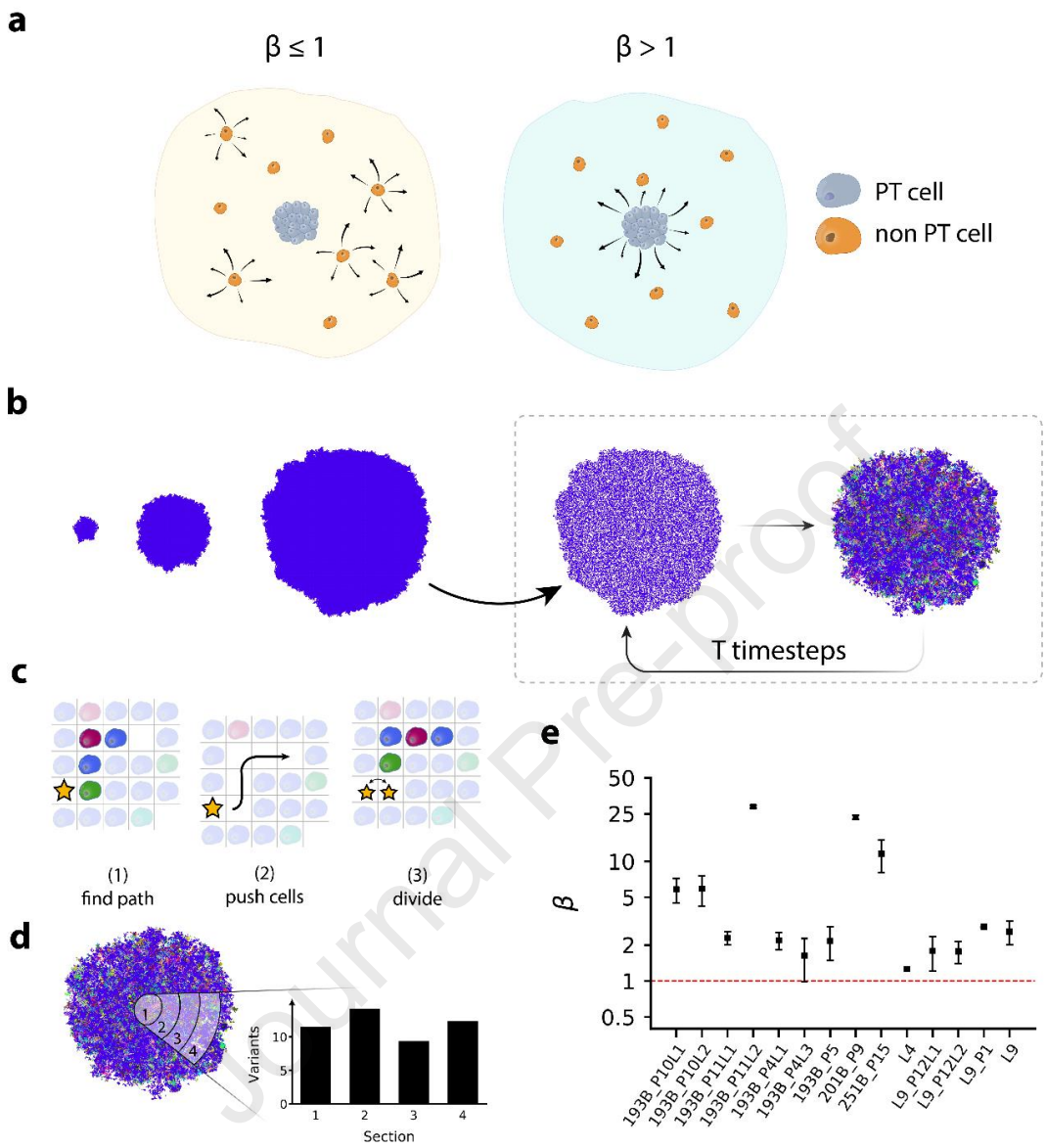





**Figure 6. Spatial modelling of homeostatic liver dynamics indicates the presence of a stem cell niche**. **(a)** Low (≤1) β regime dominated by non-localised non-PT dynamics. Regimes of β > 1 denote PT-cell dynamics localised at the centre of the system. **(b)** Simulated lobules. dynamics. After initialising the lobule, all cells carry zero mutations (single colour). Following time-steps of random killing and repopulation, cells develop mutations (different colours). **(c)** Cell pushing algorithm. Dividing cells (star) push neighbours towards a nearby empty lattice point to create an adjacent empty space, then divides and occupies the space with its progeny. **(d)** Number of unique variants measured along random radii outwards from system's centre. **(e)** Estimated β values for each portal-central axis assayed by mitochondrial NGS. Error bars are S.E.M.

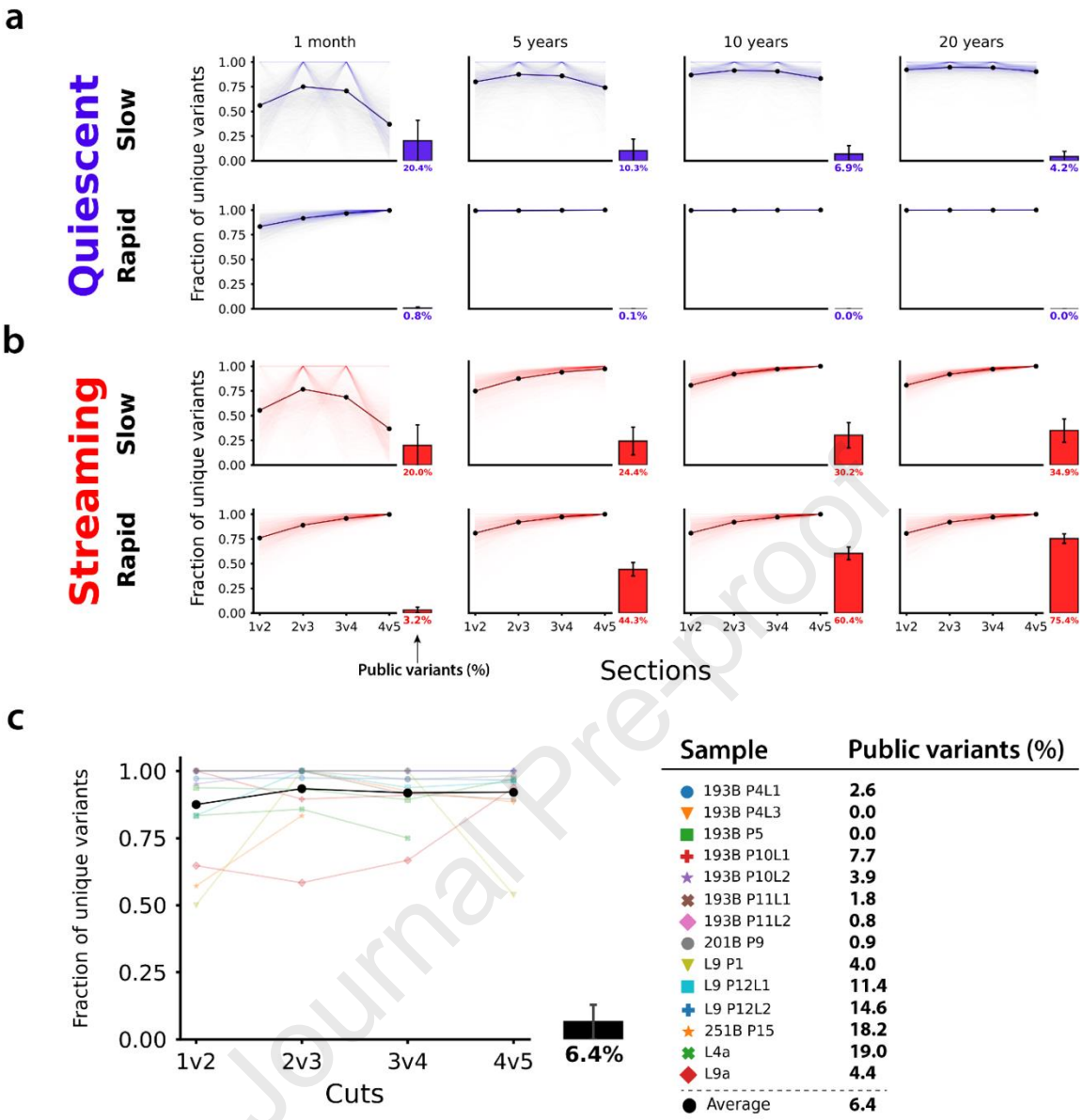



**Figure 7: Simulated and experimental public and unique private variant dynamics in spatially aligned microdissected areas of CCO clones. (a&b)** Unique private variants across sections spanning 1-dimensional simulated portal-central axes. **(a)** Quiescent and **(b)** streaming liver dynamics assuming either slow or rapid initial patch growth. Each panel displays data from 1000 simulated portal-central axes (transparent curves). Solid curves represent mean values across all simulations. Clonal hepatocytes are first generated at either a slow or rapid rate, followed by 1 month, 5, 10 and 20 years of quiescent or streaming dynamics. Proportion of public variants (those shared across all cuts) is displayed as a bar alongside each panel (mean ± SD). **(c)** Data for non-simulated experimental samples and a table of public variant percentages for each PT-CV axis.

1. Higgins GM. Experimental pathology of the liver; I. Restoration of the liver of the white rat following partial surgical removal. Arch pathol. 1931;12:186-202.
2. Rabes H, Wirsching R, Tuczek HV, Iseler G. Analysis of cell cycle compartments of hepatocytes after partial hepatectomy. Cell Proliferation. 1976;9(6):517-32.
3. Stöcker E, Wullstein H, Bräu G. Capacity of regeneration in liver epithelia of juvenile, repeated partially hepatectomized rats. Autoradiographic studies after continous infusion of 3H-thymidine (author's transl). Virchows Archiv B, Cell pathology. 1973;14(2):93-103.
4. Overturf K, Al-Dhalimy M, Ou C-N, Finegold M, Grompe M. Serial transplantation reveals the stem-cell-like regenerative potential of adult mouse hepatocytes. The American journal of pathology. 1997;151(5):1273.
5. Block GD, Locker J, Bowen WC, Petersen BE, Katyal S, Strom SC, et al. Population expansion, clonal growth, and specific differentiation patterns in primary cultures of hepatocytes induced by HGF/SF, EGF and TGF alpha in a chemically defined (HGM) medium. The Journal of cell biology. 1996;132(6):1133-49.
6. Tarlow BD, Pelz C, Naugler WE, Wakefield L, Wilson EM, Finegold MJ, et al. Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes. Cell Stem Cell. 2014;15(5):605-18.
7. Yimlamai D, Christodoulou C, Galli GG, Yanger K, Pepe-Mooney B, Gurung B, et al. Hippo pathway activity influences liver cell fate. Cell. 2014;157(6):1324-38.
8. Yanger K, Zong Y, Maggs LR, Shapira SN, Maddipati R, Aiello NM, et al. Robust cellular reprogramming occurs spontaneously during liver regeneration. Genes & development. 2013;27(7):719-24.
9. Michalopoulos GK, Barua L, Bowen WC. Transdifferentiation of rat hepatocytes into biliary cells after bile duct ligation and toxic biliary injury. Hepatology. 2005;41(3):535-44.
10. Passman AM, Strauss RP, McSpadden SB, Finch-Edmondson ML, Woo KH, Diepeveen LA, et al. A modified choline-deficient, ethionine-supplemented diet reduces morbidity and retains a liver progenitor cell response in mice. Dis Model Mech. 2015;8(12):1635-41.
11. Lázaro CA, Rhim JA, Yamada Y, Fausto N. Generation of hepatocytes from oval cell precursors in culture. Cancer research. 1998;58(23):5514-22.
12. Dorrell C, Erker L, Schug J, Kopp JL, Canaday PS, Fox AJ, et al. Prospective isolation of a bipotential clonogenic liver progenitor cell in adult mice. Genes & development. 2011;25(11):1193-203.
13. Raven A, Lu W-Y, Man TY, Ferreira-Gonzalez S, O'Duibhir E, Dwyer BJ, et al. Cholangiocytes act as facultative liver stem cells during impaired hepatocyte regeneration. Nature. 2017;547(7663):350.
14. Shin S, Upadhyay N, Greenbaum LE, Kaestner KH. Ablation of Foxl1-Cre-labeled hepatic progenitor cells and their descendants impairs recovery of mice from liver injury. Gastroenterology. 2015;148(1):192-202 e3.
15. Lu W-Y, Bird TG, Boulter L, Tsuchiya A, Cole AM, Hay T, et al. Hepatic progenitor cells of biliary origin with liver repopulation capacity. Nature cell biology. 2015;17(8):971.
16. Grompe M. Liver Stem Cells, Where Art Thou? Cell Stem Cell. 2014;15(3):257-8.
17. Wei Y, Wang YG, Jia Y, Li L, Yoon J, Zhang S, et al. Liver homeostasis is maintained by midlobular zone 2 hepatocytes. Science. 2021;371(6532):eabb1625.
18. Deng X, Zhang X, Li W, Feng R-X, Li L, Yi G-R, et al. Chronic liver injury induces conversion of biliary epithelial cells into hepatocytes. Cell Stem Cell. 2018;23(1):114-22. e3.
19. Zajicek G, Oren R, Weinreb Jr M. The streaming liver. Liver. 1985;5(6):293-300.
20. Blikkendaal-Lieftinck L, Kooij M, Kramer M, Den Otter W. Cell kinetics in the liver of rats under normal and abnormal dietary conditions: An autoradiographic study by means of [3H] thymidine. Experimental and molecular pathology. 1977;26(2):184-92.

21. Furuyama K, Kawaguchi Y, Akiyama H, Horiguchi M, Kodama S, Kuhara T, et al. Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. Nature genetics. 2011;43(1):34.
22. He L, Pu W, Liu X, Zhang Z, Han M, Li Y, et al. Proliferation tracing reveals regional hepatocyte generation in liver homeostasis and repair. Science. 2021;371(6532):eabc4346.
23. Wang B, Zhao L, Fish M, Logan CY, Nusse R. Self-renewing diploid Axin2+ cells fuel homeostatic renewal of the liver. Nature. 2015;524(7564):180.
24. Lin S, Nascimento EM, Gajera CR, Chen L, Neuhöfer P, Garbuzov A, et al. Distributed hepatocytes expressing telomerase repopulate the liver in homeostasis and injury. Nature. 2018;556(7700):244-8.
25. Planas-Paz L, Orsini V, Boulter L, Calabrese D, Pikiolek M, Nigsch F, et al. The RSPO–LGR4/5–ZNRF3/RNF43 module controls liver zonation and size. Nature cell biology. 2016;18(5):467-79.
26. Chen F, Jimenez RJ, Sharma K, Luu HY, Hsu BY, Ravindranathan A, et al. Broad Distribution of Hepatocyte Proliferation in Liver Homeostasis and Regeneration. Cell Stem Cell. 2020;26(1):27-33.e4.
27. Sun T, Pikiolek M, Orsini V, Bergling S, Holwerda S, Morelli L, et al. AXIN2+ Pericentral Hepatocytes Have Limited Contributions to Liver Homeostasis and Regeneration. Cell Stem Cell. 2020;26(1):97-107.e6.
28. Huch M, Gehart H, van Boxtel R, Hamer K, Blokzijl F, Verstegen MM, et al. Long-term culture of genome-stable bipotent stem cells from adult human liver. Cell. 2015;160(1-2):299-312.
29. Aizarani N, Saviano A, Mailly L, Durand S, Herman JS, Pessaux P, et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature. 2019;572(7768):199-204.
30. Brunner SF, Roberts ND, Wylie LA, Moore L, Aitken SJ, Davies SE, et al. Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. Nature. 2019;574(7779):538-42.
31. Fellous TG, Islam S, Tadrous PJ, Elia G, Kocher HM, Bhattacharya S, et al. Locating the stem cell niche and tracing hepatocyte lineages in human liver. Hepatology. 2009;49(5):1655-63.
32. Fellous TG, McDonald SAC, Burkert J, Humphries A, Islam S, De-Alwis NMW, et al. A Methodological Approach to Tracing Cell Lineage in Human Epithelial Tissues. STEM CELLS. 2009;27(6):1410-20.
33. Walther V, Alison MR. Cell lineage tracing in human epithelial tissues using mitochondrial DNA mutations as clonal markers. Wiley Interdisciplinary Reviews: Developmental Biology. 2016;5(1):103-17.
34. Font-Burgada J, Shalapour S, Ramaswamy S, Hsueh B, Rossell D, Umemura A, et al. Hybrid periportal hepatocytes regenerate the injured liver without giving rise to cancer. Cell. 2015;162(4):766-79.
35. Elson JL, Samuels DC, Turnbull DM, Chinnery PF. Random intracellular drift explains the clonal expansion of mitochondrial DNA mutations with age. Am J Hum Genet. 2001;68(3):802-6.
36. Greaves LC, Preston SL, Tadrous PJ, Taylor RW, Barron MJ, Oukrif D, et al. Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(3):714-9.
37. Durham SE, Samuels DC, Cree LM, Chinnery PF. Normal levels of wild-type mitochondrial DNA maintain cytochrome c oxidase activity for two pathogenic mitochondrial DNA mutations but not for m. 3243A→ G. The American Journal of Human Genetics. 2007;81(1):189-95.
38. McEnerney L, Duncan K, Bang B-R, Elmasry S, Li M, Miki T, et al. Dual modulation of human hepatic zonation via canonical and non-canonical Wnt pathways. Experimental & Molecular Medicine. 2017;49(12):e413-e.
39. Chu MW, Siegmund KD, Eckstam CL, Kim JY, Yang AS, Kanel GC, et al. Lack of increases in methylation at three CpG-rich genomic loci in non-mitotic adult tissues during aging. BMC medical genetics. 2007;8(1):50.

40. Graham TA, Humphries A, Sanders T, Rodriguez–Justo M, Tadrous PJ, Preston SL, et al. Use of Methylation Patterns to Determine Expansion of Stem Cell Clones in Human Colon Tissue. Gastroenterology. 2011;140(4):1241-50.e9.
41. Risques RA, Kennedy SR. Aging and the rise of somatic cancer-associated mutations in normal tissues. PLoS genetics. 2018;14(1):e1007108.
42. Werner B, Sottoriva A. Variation of mutational burden in healthy human tissues suggests non-random strand segregation and allows measuring somatic mutation rates. PLoS computational biology. 2018;14(6):e1006233.
43. Khrapko K, Coller HA, André PC, Li X-C, Hanekamp JS, Thilly WG. Mitochondrial mutational spectra in human cells and tissues. Proceedings of the National Academy of Sciences. 1997;94(25):13798-803.
44. Sandgren EP, Palmiter RD, Heckel JL, Daugherty CC, Brinster RL, Degen JL. Complete hepatic regeneration after somatic deletion of an albumin-plasminogen activator transgene. Cell. 1991;66(2):245-56.
45. Rhim JA, Sandgren EP, Degen JL, Palmiter RD, Brinster RL. Replacement of diseased mouse liver by hepatic cell transplantation. Science. 1994;263(5150):1149-52.
46. Alison MR. Cholangiocytes: no longer cinderellas to the hepatic regenerative response. Cell Stem Cell. 2017;21(2):159-60.
47. Russell JO, Lu WY, Okabe H, Abrams M, Oertel M, Poddar M, et al. Hepatocyte-Specific β-Catenin Deletion During Severe Liver Injury Provokes Cholangiocytes to Differentiate Into Hepatocytes. Hepatology. 2019;69(2):742-59.
48. Meineke FA, Potten CS, Loeffler M. Cell migration and organization in the intestinal crypt using a lattice-free model. Cell proliferation. 2001;34(4):253-66.
49. Sigal SH, Brill S, Fiorino AS, Reid LM. The liver as a stem cell and lineage system. The American journal of physiology. 1992;263(2 Pt 1):G139-48.
50. Reid LM, Fiorino AS, Sigal SH, Brill S, Holst PA. Extracellular matrix gradients in the space of Disse: relevance to liver biology. Hepatology. 1992;15(6):1198-203.
51. Schmelzer E, Zhang L, Bruce A, Wauthier E, Ludlow J, Yao H-l, et al. Human hepatic stem cells from fetal and postnatal donors. The Journal of experimental medicine. 2007;204(8):1973-87.
52. Zhang L, Theise N, Chua M, Reid L. Human hepatic stem cells and hepatoblasts: symmetry between liver development and liver regeneration. Hepatology. 2008;48(5):1598-607.
53. Han X, Wang Y, Pu W, Huang X, Qiu L, Li Y, et al. Lineage Tracing Reveals the Bipotency of SOX9+ Hepatocytes during Liver Regeneration. Stem Cell Reports. 2019;12(3):624-38.
54. Gramignoli R. New cell-based approaches for liver disorders: From experimental to validated clinical treatment. Journal of Hepatology. 2023;78(1):12-5.
55. Taylor RW, Barron MJ, Borthwick GM, Gospel A, Chinnery PF, Samuels DC, et al. Mitochondrial DNA mutations in human colonic crypt stem cells. The Journal of clinical investigation. 2003;112(9):1351-60.
56. Krishnan KJ, Blackwood JK, Reeve AK, Turnbull DM, Taylor RW. Detection of mitochondrial DNA variation in human cells. Genetic Variation: Springer; 2010. p. 227-57.
57. Foley JW, Zhu C, Jolivet P, Zhu SX, Lu P, Meaney MJ, et al. Gene expression profiling of single cells from archival tissue with laser-capture microdissection and Smart-3SEQ. Genome research. 2019;29(11):1816-25.
58. Gerstung M, Papaemmanuil E, Campbell PJ. Subclonal variant calling with multiple samples and prior knowledge. Bioinformatics. 2014;30(9):1198-204.

Highlights

- Hepatocyte clonal expansions emanate from periportal origins in normal human livers
- Clonal hepatocyte expansions are more numerous with age
- Biliary epithelium and hepatocytes share a common somatic ancestor
- Clonal expansion in normal human liver is slow and/or punctuated