Peer Review File

Manuscript Title: Convergent somatic mutations in metabolism genes in chronic liver disease

Editorial Notes:

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

Stanley Ng and collaborators analyzed 1202 genomes by selecting in 32 liver samples in normal liver tissues (n=5), alcohol related CLD (n=10) and NAFLD (n=17). 17 patients had synchronous HCC, 19 were cirrhotic, all normal livers came from patients with colorectal cancer. The analysis is based on laser-capture of 100-500 hepatocytes sequenced using WGS. They identified clonal evolutions, recurrent mutations in ACVR2A, CNCL5, NEAT1, GPAM, and FOXO1 with hotspot for the latter. They also identified shorter telomeres in clonal expansion and new mutational signatures.

General comment:

Even if these results are an extension of the previous study published by the same team, they are very interesting. My major concern is the overall limited number of patients with different conditions. Sometimes, it is difficult to understand which observation could be linked to the development of HCC or to the severity of the liver disease.

Specific comments

All 5 normal livers were sampled from patients with colorectal cancer, 5/6 with lung metastasis, 1/6 with liver metastasis. Are we sure that none of these patients received chemotherapy?

Since ACRV2A is a cancer driver gene, it is important to better understand the relation of mutation in non-tumor and the corresponding tumors. It is unclear in the text to understand how many patients and then clones/patients showed alterations. According to extended figure 3, 2 normal livers showed mutations and 6 patients with ARLD or NAFLD: is there a correlation with the presence or not of HCC? What is the significance of these mutations? What is the status of ACVR2A in the corresponding HCC? Could you extent on your interpretation?

As mentioned by the authors, functional consequences of mutations in NEAT1 and CNCL5 is questionable. They are accumulated in almost all patients, are they only tracer of clonality? Any relation with etiology or severity of the chronic liver disease? Please comment.

What is the exact number of patients with FOXO1 mutation? In the text it is mentioned 6 patients whereas in sup table 4 most of the patients showed a FOXO1 mutation, as for the other driver genes.
Does the conclusion of independent acquisition of FOXO1 mutations in patients very robust? On extended figure 6, only one sample located in the center of the others showed a different mutation with only few reads. Some of the trees are difficult to understand. Please comment.
Also, normal livers showed FOXO1 mutations, what is your interpretation?
In my view, FOXO1 hotspot mutations are very interesting, it should be validated in a larger series of patients to understand better the potential functional consequences in term of clonal expansion.
Since somatic S22W mutations and other are very rare in HCC, what is the significance in the non-tumor liver tissues regarding the risk of HCC or severity of the disease?

In the present data, it seems that the major determinant of telomere length was related to the

patient. Data should be shown according to age and presence of cirrhosis

Identification of the novel mutational signatures is intriguing. To which extend they are different from the known signatures in particular in comparison with aristolochic induced pattern of mutations? Their identification at the terminal branch is questionable. What could be the relation with pre-transplant treatments?

Referee #2 (Remarks to the Author):

A: Summary of the key results

The authors performed somatic whole genome sequencing of microdissected liver, identify clustering of somatic mutations in key metabolic genes and link this to the etiopathogenesis of NAFLD, hepatic insulin resistance and systemic metabolic dysfunction.

B: Originality and significance: if not novel, please include reference

The authors apply their successful analytical pipeline that was previously used in multiple other organs including liver to metabolic liver disease. The paper is based on an extension of a previously published dataset (N=14 patients), that has been expanded by 18 patients and a joint re-analysis of the data.

C: Data & methodology: validity of approach, quality of data, quality of presentation

Given the far-reaching conclusions and interpretation provided by the authors, the manuscript would benefit from clarification and additional experiments.

Clinical cohort and characterization:

The authors center their data interpretation and pathogenetic conclusions on non-alcoholic fatty liver disease and insulin resistance. Given that experimental goal, the choice of samples in this manuscript is unusual. The experiment includes 17 samples with a “NAFLD” etiology – and these come to a large part from a heterogenous and often very advanced stage liver disease cohort. The setting and the efficient re-use of existing data facilitates this approach and the authors need surgical material for their methods explain this strategy but a more typical NAFLD cohort at least a replication stage would improve the message. A few additional specific points:
1) Alcoholic liver disease has a distinct pathogenesis from non-alcoholic fatty liver disease.
2) The authors should improve their Suppl. Table 1: the NAS (Kleiner score) is not validated for alcoholic liver disease (but applied by the authors). Fibrosis usually is scored quantitatively from F1 to F4 and not qualitatively. Did a single pathologist score all the samples (the NAS score is known to vary substantially).
3) The diabetes in advanced cirrhosis (a lot of transplanted patients were investigated) has a very different etiology to “normal” diabetes. In cirrhosis, there is glucose metabolism instability due to the limited liver function, that is specific to cirrhosis rather than insulin resistance in type 2 diabetes.
4) Many patients with alcoholic liver disease have chronic pancreatitis and insulin deficiency. Was this present in the investigated patients? This would qualify more as “type 3” diabetes.
5) A typical set of patients (with surgical samples and “typical” type 2 diabetes) for NAFLD are bariatric patients. There are very large cohorts and biobanks around that could be used (also for validation – see below).

Main experimental finding (somatic variant):

I focus on the FOXO1 finding as this is the focus of the functional discussion by the authors.
1) The finding is based on 3 out of 17 patients (~17%) with NAFLD and 2 out of 10 samples with ALD (20%). One should aim to replicate the result, as it is the result of a complex and multi-stage experimental process (i.e. NG sequencing with its characteristic error signatures) and complex bioinformatic analyses.
2) Experimental validation: As >10% of genomic reads harbor the driver mutation, this could be validated using allelic real time PCR using DNA as a template to quantify the prevalence of the variant independently. This method is for instance used widely to quantify template before hybridizations of DNA or after FFPE extraction.
3) Such a quantitative allelic PCR could be used on the large liver biobanks of NAFLD and ALD even without microdissection and on FFPE sections to independently replicate the finding.
4) Why is this key somatic mutation not present in HCCs arising in the context of NAFLD and ALD? The authors quote the ICGC data and state that in over 1670 in only 0.18% of tumors carry a FOXO1 mutation. How does this look in the ALD/NAFLD subsets as compared to viral hepatitis? As reported by the authors, 15-20% of patients carry FOXO1 mutations in 5-40% of their hepatocyte clones (and these are further key pathogenetic drivers) one would expect 10-50 fold more positive HCCs. Are FOXO1 mutations selectively repaired in cancerogenesis or is FOXO1 specifically protecting clones from malignant transformation?
5) More general: As liver contains ~20% of non-parenchymal cells (endothelials cells, HSCs, Kupfer cells etc.) – have the authors checked their microdissections for hepatocyte content (e.g. by RNASeq)?

Functional exploration:

FOXO1 is proposed to be a major pathogenetic switch (in a subset of patients):
1) The authors are in the fortunate situation to have adjacent sections from these very patients available. With the large quantitative effects (up to 40% of observed reads with FOXO1 mutations in Figure 1) one could investigate the nuclear translocation directly in the patient sections harboring the mutations. One could perform RNASeq from microdissected tissue to look for characteristic downstream signatures of altered FOXO1 signaling, could perform immunohistochemistry or RNA in situ hybridization. Samples with a similar clinical profile but lack of FOXO1 somatic mutations that are already in the hand of the investigators would provide ideal controls.
2) The cellular experiments are not statistically analyzed. How many replicates were used? Can the effects be reproduced by Western blotting of isolated cellular fractions? Are the results stable in another cell line? Can we observe a typical FOXO1 downstream signature (e.g. by RNASeq)?
3) Do the (large) existing GEO liver and NAFLD expression datasets harbor a FOXO1 signature in the proposed ~15 percent of patients with this pathway? Do they form a distinct cluster?

D: Appropriate use of statistics and treatment of uncertainties

See above. In addition, the manuscript analyzed 32 liver samples from 27 patients (i.e. there were multiple samples from the same patient). The observations within one patient are likely not independent. It is not clear, how the authors controlled for this factor in their statistical testing.

E: Conclusions: robustness, validity, reliability

As the authors make fundamental contributions to the pathogenesis of insulin resistance in liver and NAFLD, the authors could enhance the study by a) the investigation of an appropriate cohort, b) a more complete characterization of existing samples, c) replication of the core finding in an independent cohort d) experimental validation of their mechanistic findings and e) a deeper functional exploration of the effects of their core finding.

Referee #3 (Remarks to the Author):

The article by Ng et al., aims to comprehensively characterize driver mutations in chronic liver disease. The analysis is particularly focused on identifying somatic mutations that may be important in the progression of chronic liver disease to hepatocellular carcinoma (HCC). The authors identified 4 significant genes ACVR2A, FOXO1, GPAM and CLCN5 and focus their attention on FOXO1 and GPAM. Of note, neither FOXO1 nor GPAM mutations were major contributors to HCCs (sequenced in ICGC).

The physiological implications for the most prevalent FOXO1 mutations were assessed in vitro and confirmed to impair nuclear export of FOXO1. In vitro analysis was not conducted on GPAM mutations though it was predicted to impair de novo lipogenesis. Both FOXO1 and GPAM mutations exhibited evidence for convergent acquisition of somatic mutations. As the authors discuss, these findings are suggestive of a novel concept that “somatic mutations might play a contributory role to systemic metabolic dysfunction.”

Ng and colleagues also quantified telomere lengths at clone-specific resolution. Interestingly, they showed that telomeres became progressively shorter as the size of a clone increased and that the effect sizes in chronic disease considerably outweighed the relatively minor shortening of telomere lengths with age.

In summary, the article provides a comprehensive and large genetic analysis of human liver in the context of 2 common chronic liver diseases. Mutations in FOXO1 and GPAM strongly suggest convergent acquisition of somatic mutations that may lead to dysregulated insulin signaling in the whole tissue. Despite the comprehensive genetics of this study, the implications and generalizability of the findings are limited.

Comments:

1. The study appears to be a logical progression from previous studies with increased sample numbers and conclusions drawn from this extended analysis are in line with, and build upon, prior observations. The novelty of the study lies in the observation of convergent somatic mutations in FOXO1 and GPAM. Expression changes in FOXO1 and GPAM are well understood to disrupt hepatic lipid and glucose metabolism. However, apart from limited in vitro characterization, the functional metabolic consequences of the accumulated mutations are unclear, and only a matter of speculation in the discussion.

2. Although large numbers of analyses have been performed, these have been applied a very limited number of liver samples from a single center with 2 distinct types of metabolic liver diseases (NAFLD and ARLD). Each of these diseases is distinctly heterogeneous from the standpoints of disease phenotype, genetic susceptibility, gender, age and ethnicity. Whether the mutational pressure that yielded the observed mutations is generalizable is not clear.

3. In this study, the control group was normal liver. It would be important to understand whether other chronic liver disease that may not ultimately exhibit phenotypes of insulin resistance (e.g. chronic cholestatic diseases or autoimmune liver disease) accumulate similar mutations at similar burdens.

4. There is no meaningful analysis of the overall histopathology (or accounting of patient characteristics) for either type of liver disease studied, nor are there provided the histopathologic features present at the sites of laser capture. Standardized methodologies are available for these characterizations.

5. For NAFLD, the authors fail to distinguish between simple steatosis (NAFL) and steatohepatitis. In many ways, these are quite distinct entities, with markedly different natural histories with respect to chronic liver disease.

6. The manuscript is extremely turgid and laden with technical jargon. For this to be of meaning to the broader hepatology and metabolic communities, it should be written in a much more accessible style.

Author Rebuttals to Initial Comments:

Referees' comments:

Referee #1 (Remarks to the Author):

Stanley Ng and collaborators analyzed 1202 genomes by selecting in 32 liver samples in normal liver tissues (n=5), alcohol related CLD (n=10) and NAFLD (n=17). 17 patients had synchronous HCC, 19 were cirrhotic, all normal livers came from patients with colorectal cancer. The analysis is based on laser-capture of 100-500 hepatocytes sequenced using WGS. They identified clonal evolutions, recurrent mutations in ACVR2A, CNCL5, NEAT1, GPAM, and FOXO1 with hotspot for the latter. They also identified shorter telomeres in clonal expansion and new mutational signatures.

General comment:

Even if these results are an extension of the previous study published by the same team, they are very interesting. My major concern is the overall limited number of patients with different conditions. Sometimes, it is difficult to understand which observation could be linked to the development of HCC or to the severity of the liver disease.

At 1202 whole genome sequences in the first submission, our manuscript already represented the most sequencing data generated to date for studying patterns of somatic mutations in any normal organ system – **with this revision, we have added a further 388 genomes across all 8 anatomical segments of two explanted livers**. Importantly, all the key messages from the first submission have been reproduced in the extended cohort – we see an additional 12 independent *FOXO1* S22W mutations, with multiple independent clones in both patients; we find further *GPAM* mutations; we identify another gene involved in lipid metabolism, *CIDEB*, as recurrently mutated in chronic liver disease, again showing the pattern of convergent evolution; and the novel mutational signature is recaptured in one of the new patients.

A key feature of our experimental design was the deliberate choice to perform multiple genome sequences on each sample (typically 30-50 per patient), which has the inevitable consequence that the number of patients is considerably lower than the number of genomes. We robustly defend this experimental design, based on the following arguments –

- *Normal tissue is polyclonal* – As we and others have demonstrated across all normal organ systems studied to date[1–9], including liver[10–13], normal tissue represents a patchwork of independently evolving clones. Indeed, for any given solid tissue, there are typically many millions of independent clones in an adult human. While in either cancer genome sequencing (where all tumour cells derive from a single ancestral cell) or germline genome analyses (where the interest is inherited genetic variation), a single genome per person is sufficient, studying somatic mutations in normal tissues benefits from analysing more than one of the millions of independent clones.
- *Liver disease generates considerable within-patient, clone-to-clone variation* – In our earlier study, we demonstrated that mutation burden and mutational signatures can vary by orders of magnitude between even neighbouring clones within the same patient[10]. In the current study, we show that this clone-to-clone variation extends to include telomere lengths and distribution of driver mutations. Two corollaries follow from this – (1) we fail to characterise a given patient's hepatic landscape

unless we assess many clones; (2) we continue to increase statistical power for detecting driver mutations by sequencing multiple samples per patient because they represent independent evolutionary experiments exposed to the same microenvironment.

- *Convergent evolution is a critical feature of the FOXO1*, *GPAM* and, now, *CIDEB story* – For a somatic mutation to generate a systemic disease, enough cells within the organ system must carry the mutation. This requires either that a single clone expands to a massive population size (as occurs in cancer) or that the mutation evolves many times independently within the same patient's organ. The finding that the *FOXO1*, *GPAM* and *CIDEB* mutations arise multiple times within a tissue sample representing <0.01% of the total liver mass[14] suggests that many hundreds of such clones likely existed within those patients' livers. As discussed below in our response to the 5th specific comment from this reviewer, definitive evidence of convergent evolution of a hotspot mutation, like those in *FOXO1*, can only be obtained with high-resolution reconstruction of phylogenetic trees from multiple whole genome sequences per patient.

In summary, then, characterising the landscape of normal tissues requires a different mindset than, say, cancer genomics studies. We do not sacrifice much statistical power for detecting drivers by sequencing multiple microdissections per patient, because clones evolve in parallel. Furthermore, we cannot comprehensively describe a given patient's tissue unless we measure clone-to-clone variation within that patient.

We have included data from the 388 additional genomes throughout the manuscript, especially in the new section reporting the *CIDEB* mutations (Results, 'Mutations in *CIDEB*', p. 8), as well as in sections on *FOXO1* and mutational signatures. Recognising that the above considerations on experimental design in the relatively new field of non-cancer mutagenesis would not be immediately obvious to the general readership of Nature, we have included a comprehensive section as Supplementary Note 1 on experimental design and analysis outlining these and other aspects (including statistical power for driver mutation discovery). This section is signposted in the main text, through a brief one-sentence summary of the key points (Results, 'Extended cohort of NAFLD patients', p. 4).

Specific comments

All 5 normal livers were sampled from patients with colorectal cancer, 5/6 with lung metastasis, 1/6 with liver metastasis. Are we sure that none of these patients received chemotherapy?

Clinical information was recorded prospectively on all patients with samples entered in the tissue bank. These cases were selected, in part, because none of them had received neoadjuvant systemic chemotherapy prior to resection of the hepatic metastases. At completion of our study, clinical

records for all patients were re-reviewed by a clinical hepatologist (M.H.) to ensure the accuracy of the data and fill in any missing information. As a result of this careful checking, we are certain none of the patients, including the normal liver controls, had received chemotherapy prior to sample acquisition.

The column reporting the lung metastases was included in error in the demographic details in the previous submission, but was detailing the clinical outcome of the patients after their hepatic resection. Although some patients did develop distant metastases later, staging CT immediately prior to hepatic resection showed only liver-localised disease.

We have included this information in the main text (Results, 'Extended cohort of NAFLD patients', p. 4) and more detail on clinical data acquisition and checking in the Supplementary Methods (pp. 1-2).

Since ACRV2A is a cancer driver gene, it is important to better understand the relation of mutation in non-tumor and the corresponding tumors. It is unclear in the text to understand how many patients and then clones/patients showed alterations. According to extended figure 3, 2 normal livers showed mutations and 6 patients with ARLD or NAFLD: is there a correlation with the presence or not of HCC? What is the significance of these mutations? What is the status of ACVR2A in the corresponding HCC? Could you extent on your interpretation?

We agree that this is an interesting question. We do not see any shared mutations between the normal liver regions we sequenced and that patient's HCC, so the *ACVR2A* mutations we observe have not directly contributed to HCC evolution. As the reviewer points out, we do find mutations or structural variants affecting *ACVR2A* in 2 normal liver controls, 3 patients with ARLD and 4 with NAFLD. There appears to be no correlation with the presence of HCC – several of the patients with these mutations in cirrhotic liver had not developed HCC.

Our data imply that these mutations confer a selective advantage on hepatocyte clones – the significant q value for point mutations ($q=7x10^{-9}$), as well as 4 structural variants specifically deleting the gene, suggest that hepatocytes gain a fitness advantage through mutations inactivating the gene (as seen in HCC). Since we see the mutations in normal liver, ARLD and NAFLD, it would seem that this fitness advantage is not specific to a particular microenvironment. This pattern of known cancer genes under positive selection in normal tissue resembles that seen in other organs, such as the recurrent *NOTCH1* and *TP53* mutations seen in normal squamous epithelium[1,4,5,7] or PI3K-pathway mutations in normal endometrium[6,15].

We have included more detail and interpretation of the *ACVR2A* mutations in the main text (Results, 'Driver mutations, p. 5). Also, we have moved the comparison of driver landscape between the ICGC HCC data and our cohort of chronic liver disease to a new section (Results, 'Comparison with selective landscape of HCC', pp. 9-10).

As mentioned by the authors, functional consequences of mutations in NEAT1 and CNCL5 is questionable. They are accumulated in almost all patients, are they only tracer of clonality? Any relation with etiology or severity of the chronic liver disease? Please comment.

We agree that the recurrent mutations in these two genes are of questionable significance. In fact, the additional 388 genomes sequenced for this paper's revision provided no new *CLCN5* mutations, and it no longer reaches significance using our dN/dS algorithm. We suspect *NEAT1* mutations arise from a specific mutational process that generates medium-sized indels in highly transcribed genes that operates across many tissues[16] – this hypothesis has also been proposed to explain their recurrence in many tumour types[17]. Therefore, it is likely the reviewer is correct that the *NEAT1* mutations are probably a marker of an underlying clonal mutation process.

We have removed mention of *CLCN5* from the manuscript text and figures.

What is the exact number of patients with FOXO1 mutation? In the text it is mentioned 6 patients whereas in sup table 4 most of the patients showed a FOXO1 mutation, as for the other driver genes.

We agree with the reviewer that because of our experimental design, with multiple microdissections per liver sample, we need to be scrupulous about what denominator we report for results. We were not sufficiently careful about this in the original submission, and this has the potential to lead to some confusion for the reader. Thus, for the *FOXO1* mutations, a total of 45 (out of 1590) microdissections carried the mutations – the individual microdissection counts of mutant / wild-type sequencing reads are what are shown in the pie charts in Figure 1A, for example. Many of these microdissections represented different cuts from the same original clone of hepatocytes, so with high-resolution reconstruction of the phylogenetic tree, we determined that 26 separate clones of hepatocytes carried *FOXO1* mutations – the clone level data are what are reported in Supplementary Table 4 and the phylogenetic trees in Figure 1. Finally, these 26 clones were distributed across 8 (out of 34) patient samples (noting that one clone in a normal control was a nonsense mutation in *FOXO1*, which is of uncertain significance compared to the hotspot S22W/R21L mutations) – this is the count that is reported in the abstract.

For the statistical analyses, we used clusters of mutations assigned to phylogenetic trees for inference of driver mutations and mutational signatures. This means that each mutation is only counted once in the statistical analysis, but that every clonally independent acquisition of a mutation is counted.

We are now more explicit about the counts and the denominator for *FOXO1*, *CIDEB* and *GPAM* mutations in the abstract, main text, figure legends and table legends.

Does the conclusion of independent acquisition of FOXO1 mutations in patients very robust? On extended figure 6, only one sample located in the center of the others showed a different mutation with only few reads. Some of the trees are difficult to understand. Please comment.

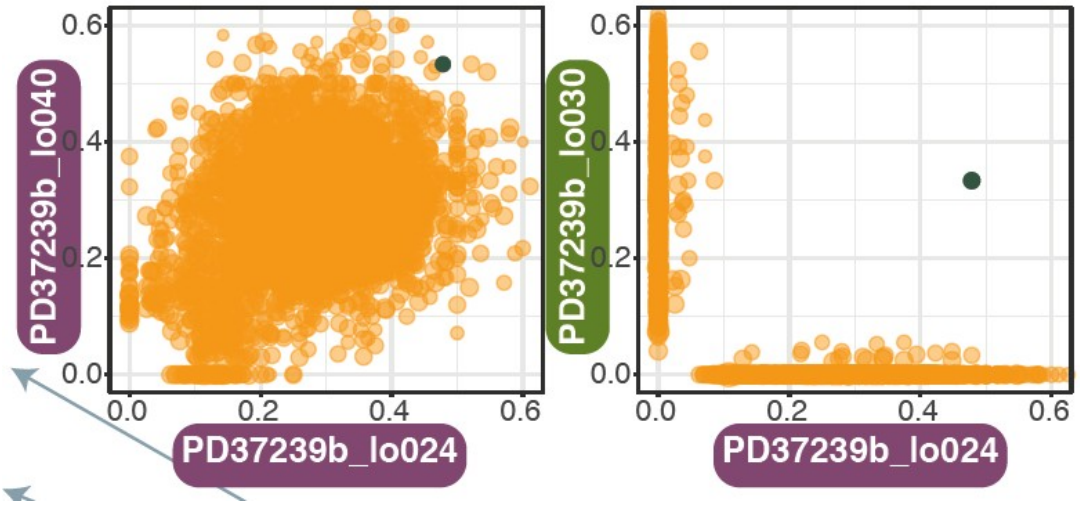
This is a critical component of our manuscript, so the reviewer is correct to check its robustness. The parallel evolution for *CIDEB* and *GPAM* mutations, like that for *NOTCH1* mutations in normal squamous epithelium[1,4,5,7,9], is easy to verify because the driver mutations are distributed throughout the gene and thus different clones will typically carry different specific base changes. For a hotspot mutation such as the S22W mutation in *FOXO1* (and, say, *PIK3CA* mutations in normal endometrium[6,15]), the same base change can arise independently in different clones, which could be mistaken for shared clonality – that is, the two microdissections are clonally related. The only means to distinguish these two possibilities is to use the somatic mutations in the rest of the genome (that is, passenger mutations).

This is what we are illustrating in Extended Figure 6, as highlighted by the reviewer (an insert from this figure is pasted below for reference, representing plots of variant allele fraction for all mutations called in individual pairs of microdissections). In a situation where two microdissections share a *FOXO1* S22W mutation (dark green point) because they are clonally related, we see many other passenger mutations shared between the two microdissections (pale orange points). This is shown in the scatterplot on the left in the insert below – many mutations present in the microdissection along the x-axis also have reads reporting that variant in the microdissection shown on the y-axis.

In contrast, when two clones acquire the same hotspot mutation independently, there will be no other shared mutations since they are clonally unrelated. This is exemplified by the pair of microdissections shown in the scatterplot on the right below – here the *FOXO1* S22W mutation (dark green point) is present at high variant allele fraction in both samples, but *all other mutations* (pale orange points) in the two microdissections lie along the x- or y-axes (allowing for occasional sequencing errors), indicating that the two microdissections are clonally unrelated.

Further evidence for the independent acquisition of *FOXO1* S22W mutations in multiple clones within the same patient comes from the sampling of all 8 anatomical segments of the liver in the 388 genomes added with the revised manuscript. We found *FOXO1* S22W mutations in 3 separate segments in one of these patients and in 4 separate segments in the other – given the many centimetres of liver separating these segments and their very early foetal divergence, this argues strongly for convergent acquisition of the mutations.

Since this is a critical aspect of our story, and involves a series of analyses and inferences that are new challenges arising in sequencing polyclonal normal tissues, we have included a Supplementary Note outlining the key principles in this clonal reconstruction (Supplementary Note 2), signposted in the main text together with the results from sampling all 8 anatomical segments of the liver in the two new patients (Results, '*FOXO1* hotspot mutations, p. 6).



Insert from Extended Figure 6 showing scatterplots of variant allele fraction for somatic mutations called in pairs of microdissections from the same patient. Passenger mutations are shown in pale orange, while the *FOXO1* S22W mutation is shown in dark green. The scatterplot on the left represents two clonally related microdissections, where the majority of mutations present in one microdissection have a high fraction of reads reporting the mutation in the other. In the scatterplot on the right, the only shared variant is the *FOXO1* mutation, indicating that the two microdissections are clonally unrelated.

Also, normal livers showed FOXO1 mutations, what is your interpretation?

To be clear, only one microdissection in one normal liver sample showed a *FOXO1* mutation and this was a nonsense mutation, not the S22W or R21L missense mutation. Our interpretation therefore is that the hotspot *FOXO1* mutations are specific to liver disease.

We have clarified the frequency and interpretation of the one *FOXO1* mutation in normal liver in the main text (Results, '*FOXO1* hotspot mutations, p. 6).

In my view, FOXO1 hotspot mutations are very interesting, it should be validated in a larger series of patients to understand better the potential functional consequences in term of clonal expansion. Since somatic S22W mutations and other are very rare in HCC, what is the significance in the non-tumor liver tissues regarding the risk of HCC or severity of the disease?

We are pleased that the reviewer agrees with us that the *FOXO1* mutations are very interesting. To demonstrate their reproducibility and their generalised distribution throughout the livers of patients with chronic liver disease, we have sequenced all 8 anatomical subunits of an additional two patients with NAFLD, adding 388 genomes to the cohort as discussed above. We found that both new patients carried *FOXO1* S22W mutations in their liver, and that these variants were seen in multiple segments (3 in one; 4 in the other). This confirms that *FOXO1* mutations are indeed widespread through the liver of patients with metabolic liver disease.

Given the absence of S22W mutations from HCC, we believe that their presence in normal liver may contain minimal information about a patient's future risk of developing HCC. Rather, we believe that they are more likely to reflect the selective pressures acting on the hepatocytes in the presence of metabolic disease such as sustained consumption of excess calories or alcohol. However, this will require considerable further clinical characterisation to assess fully, which we believe to be beyond the scope of this paper reporting their discovery.

We have included the additional 388 genomes from the two new patients throughout the manuscript, such as the analyses for driver mutations and mutational signatures.

In the present data, it seems that the major determinant of telomere length was related to the patient. Data should be shown according to age and presence of cirrhosis

We have included box-and-whisker plots of the raw telomere counts distributed by age and by disease status in Extended Figure 11A-B.

Identification of the novel mutational signatures is intriguing. To which extend they are different from the known signatures in particular in comparison with aristolochic induced pattern of mutations? Their identification at the terminal branch is questionable. What could be the relation with pre-transplant treatments?

We agree that the novel mutational signature is interesting. It is distinct from the aristolochic acid signature – while it also causes T>A mutations, and with transcriptional strand bias, the strong preference for mutation in a CpT context is somewhat to different to the broader distribution seen with aristolochic acid[18,19]. Furthermore, we did extract the classic aristolochic acid signature (SBS22) as a separate signature in the same run of the algorithm, suggesting that the novel signature is indeed different (cosine similarity <0.9). None of the patients who had the novel signature had known exposure to or risk factors for aristolochic acid. There was no clear association of the signature with pre-transplant treatments, albeit with relatively small numbers.

We interpret the preponderance of the new mutational signature on terminal branches of the phylogenetic tree as meaning that it emerges late in the disease process, supported by the fact that it became more pronounced over time in the patient in whom we studied liver samples from 3 time-points.

We have clarified the distinction between the novel signature and that of aristolochic acid in the main text (Results, 'Mutational signatures in chronic liver disease', p. 11).

Referee #2 (Remarks to the Author):

A: Summary of the key results

The authors performed somatic whole genome sequencing of microdissected liver, identify clustering of somatic mutations in key metabolic genes and link this to the etiopathogenesis of NAFLD, hepatic insulin resistance and systemic metabolic dysfunction.

B: Originality and significance: if not novel, please include reference

The authors apply their successful analytical pipeline that was previously used in multiple other organs including liver to metabolic liver disease. The paper is based on an extension of a previously

published dataset (N=14 patients), that has been expanded by 18 patients and a joint re-analysis of the data.

C: Data & methodology: validity of approach, quality of data, quality of presentation

Given the far-reaching conclusions and interpretation provided by the authors, the manuscript would benefit from clarification and additional experiments.

We agree that some of the analyses performed are complex, and we elided over some of the technical detail to improve readability in the main text.

We have included Supplementary Notes 1 and 2 to provide considerably more discussion about the reasoning behind the experimental design, the inference of convergent evolution and the robustness of the inferred driver mutations.

Clinical cohort and characterization:

The authors center their data interpretation and pathogenetic conclusions on non-alcoholic fatty liver disease and insulin resistance. Given that experimental goal, the choice of samples in this manuscript is unusual. The experiment includes 17 samples with a "NAFLD" etiology – and these come to a large part from a heterogenous and often very advanced stage liver disease cohort. The setting and the efficient re-use of existing data facilitates this approach and the authors need surgical material for their methods explain this strategy but a more typical NAFLD cohort at least a replication stage would improve the message.

To cement the statistical evidence for the driver mutations identified in our first submission and explore how widespread they are across the entire liver, we have now added a further 388 whole genomes. From two patients with NAFLD, we obtained samples from all 8 anatomical (Couinaud) segments of the liver, sequencing 22-28 microdissections from each segment. All the major findings from the first submission were replicated in the 388 additional genomes, and some new discoveries emerged with the increased statistical power:

- *A further 12 independent FOXO1 S22W clones were added* – Both patients carried multiple, independent *FOXO1* S22W clones, seen in 3 separate anatomical segments in one patient and 4 segments in the other. Even within a single segment, there were up to 4 independent clones with the mutation, such that one of the two patients had 9 independent *FOXO1* S22W clones discovered among the regions sampled.

- *A further 2 GPAM mutations added* – The two new *GPAM* mutations included one variant at an essential splice site, providing additional evidence that the drivers operate through loss of function effects.
- *The new mutational signature was observed in one of the new patients* – Intriguingly, the new mutational signature, comprising predominantly T>A mutations, was seen in all 8 segments of the liver. However, as reported in our first submission, we again saw that the activity of the new signature was remarkably variable across even neighbouring clones within the liver, with very high rates in some nodules but almost absent from others.
- *CIDEB now shows a significant excess of mutations* – *CIDEB* was just below the significance threshold with our first submission, and we therefore did not discuss it. However, with the additional 388 genomes, we uncovered many more mutations in the gene – in total, we have now identified 20 non-synonymous mutations in the gene, which encodes a protein of only 211 amino acids. The q-value for *CIDEB* is now $<10^{-16}$ even with correction for multiple hypothesis testing. In one of the new patients, we found no fewer than 15 non-synonymous mutations across 6 of the 8 segments. Intriguingly, *CIDEB* is involved in lipid metabolism in hepatocytes, and knock-out mouse models show resistance to dietary steatohepatitis and increased insulin sensitivity[20]. CIDE proteins regulate fusion of intracellular lipid droplets, mediated by the formation of homodimers between CIDE proteins on different droplets[21,22]. Homodimerisation occurs through electrostatic contacts between positively charged residues on the CIDE protein from one lipid droplet and negatively charged residues on the other. Interestingly, the missense mutations we observed were predominantly located in the two domains implicated in homodimerisation of CIDE proteins, and many of them either switched a charged residue for a neutral one (R45W, R45Q, K140N, R144P) or even reversed the charge (D42H, K62E, E78K). Previous *in vitro* mutagenesis studies have shown that altering the charge on these key conserved residues, including some of those mutated in our patients, abrogates homodimerisation, preventing fusion and growth of lipid droplets within the cell[21,22].

Our discovery that *FOXO1*, *CIDEB* and *GPAM* are recurrently mutated in chronic liver disease is absolutely robust. The false discovery rates (q-values) we report essentially measure the expected proportion of type 1 errors (false positive calls) given the multiple hypothesis tests we have performed. For our three key new genes, the q-values were $<1x10^{-16}$ (*FOXO1*), $<1x10^{-16}$ (*CIDEB*) and $1x10^{-5}$ (*GPAM*). Although we did not quote the q-value for the exact same base in *FOXO1* being mutated in 24 separate clones, it is $3x10^{-51}$. This tells us what is already obvious – that capturing 24 independent S22W variants from 1590 samples with a mutation density of <1 / million bases is preposterously improbable to happen by chance.

We accept that our cohort comes from patients with predominantly advanced stage disease – as the reviewer indicates, this is our best source of high-quality samples of liver tissue. While patients with earlier stage disease do on occasion have liver biopsy for clinical indications (although increasingly rarely), these biopsies tend to be fixed in formalin (incompatible with our microdissection methodology), and much of the biopsy is consumed for clinical diagnostic purposes.

There are, of course, many questions that remain about the clinical significance of these mutations, as all three reviewers point out. At what stage in the disease process do driver mutations emerge? What role do they play in systemic metabolic phenotypes? Does the burden of driver mutations correlate with disease severity? Can they be used to stratify patients for prognosis or even therapy choice?

We do not have the answers to these questions. Clinical phenotypes are notoriously difficult to relate to genetic variants, as evidenced by the paucity of verified gene-by-environment interactions for germline polymorphisms[23]. Indeed, we believe it will take 5-10 years, many hundreds, even thousands, of patients, and prospective cohort studies to address these questions with any degree of certainty. As the reviewers point out, they are undoubtedly important follow-on studies for our discovery, but we would argue that they are well beyond the scope of this paper. It would be a sorry Nature paper that did not prompt such far-reaching and profound questions, but it is unreasonable to expect all to be answered in the first report.

We have included the additional 388 genomes from the two new patients throughout the manuscript, such as the analyses for driver mutations and mutational signatures.

A few additional specific points:

1) Alcoholic liver disease has a distinct pathogenesis from non-alcoholic fatty liver disease.

This is true, and it is therefore perhaps surprising that we see *FOXO1* and *GPAM* mutations in both disease entities. However, there are numerous commonalities between the two diseases, and it may be these underpin the selective advantage of the driver mutations. The histological appearances are highly overlapping between the two conditions and the germline variation that predicts outcome in the two diseases, in terms of fibrosis progression, liver failure and HCC development are almost identical[24]. In particular, lipotoxicity is a feature of both NAFLD and alcoholic fatty liver[25], and it is plausible that it exerts such strong selective pressure to foster the emergence of clones with mutations affecting lipid handling.

In many ways, finding these mutations in both alcohol-related liver disease and NAFLD underscores their generalised importance – they are not restricted to a single disease process.

We have included cited references and discussed the importance of lipotoxicity and metabolic pathways to the pathogenesis of both ARLD and NAFLD in the Discussion (pp. 12-13).

2) The authors should improve their Suppl. Table 1: the NAS (Kleiner score) is not validated for alcoholic liver disease (but applied by the authors). Fibrosis usually is scored quantitatively from F1 to F4 and not qualitatively. Did a single pathologist score all the samples (the NAS score is known to vary substantially).

We agree that the Kleiner score is not validated for ARLD. In the absence of a validated and internationally agreed histological scoring system for ARLD we utilised the NAS / Kleiner system to allow comparison across the samples. Both specialist pathologists (SJA, SED), blinded to the sequencing data, individually reviewed all histological material for the samples in our cohort, with discussion for any discordances, generating the various semi-quantitative indices and qualitative comments on the slides.

We have placed the values for Kleiner score in Supplementary Table 1 for ARLD in brackets to emphasise they are there only for completeness, together with an explanatory note in the table legend. We have also provided the accepted quantitative scores of fibrosis, for both Kleiner and Ishak systems.

3) The diabetes in advanced cirrhosis (a lot of transplanted patients were investigated) has a very different etiology to "normal" diabetes. In cirrhosis, there is glucose metabolism instability due to the limited liver function, that is specific to cirrhosis rather than insulin resistance in type 2 diabetes.

We agree that the diabetes in advanced cirrhosis has a different mechanism and pathogenesis to the typical obesity-associated type 2 diabetes. **We have acknowledged the potentially different pathogenesis of metabolic disturbance in advanced stage cirrhosis in the Introduction (p. 3).**

4) Many patients with alcoholic liver disease have chronic pancreatitis and insulin deficiency. Was this present in the investigated patients? This would qualify more as "type 3" diabetes.

None of the patients we studied had clinical or imaging features of chronic pancreatitis, **data we now include in Supplementary Table 1.**

5) A typical set of patients (with surgical samples and "typical" type 2 diabetes) for NAFLD are bariatric patients. There are very large cohorts and biobanks around that could be used (also for validation – see below).

This is indeed an interesting set of patients, and one that could give interesting insights into the clinical relevance of these mutations. The NHS does not fund bariatric surgery, except in extreme cases, and no liver tissues are taken routinely from these cases. Therefore, we do not have access to these tissues within our patient cohorts. We would definitely be keen to examine such cases in follow-up studies.

Main experimental finding (somatic variant):

I focus on the FOXO1 finding as this is the focus of the functional discussion by the authors.
1) The finding is based on 3 out of 17 patients (~17%) with NAFLD and 2 out of 10 samples with ALD (20%). One should aim to replicate the result, as it is the result of a complex and multi-stage experimental process (i.e. NG sequencing with its characteristic error signatures) and complex bioinformatic analyses.

Massively parallel sequencing is, as the reviewer indicates, subject to characteristic error profiles. The algorithms we have used in this manuscript for mutation-calling are well-established, well-validated pipelines that have been the workhorse of cancer genome sequencing, and more recently normal tissue sequencing, at our Institute (and many others) for more than a decade. In the recent Pan-Cancer Analysis of Whole Genomes (PCAWG) consortium, published as a suite of papers in Nature earlier this year, a head-to-head validation of >10 variant-calling algorithms was performed at Washington University[26]. This independent, blinded analysis revealed that Sanger's algorithms had the highest precision for calling both somatic substitutions and indels of any algorithm (Figure 1A-B from the paper pasted in below). Precision here is defined as 1 – false positive rate; namely the fraction of variants called by our algorithm that were independently verified as true somatic mutations by the validation sequencing.

In our earlier publication describing somatic mutations in non-cancerous liver[10], we also undertook an extensive assessment of our mutation-calling approaches. It is less straightforward than for cancer because of the polyclonal nature of normal tissue, but we essentially used microdissections of the same x-y co-ordinates of adjacent z-sections from the same liver sample. These microdissections were then processed independently through library production, sequencing and mutation-calling. Because clones occupy regions in 3-dimensional space, this was equivalent to sequencing the same clone twice, and we could estimate precision of somatic mutation calls as >95%, even in this more challenging scenario[10]. We have equivalent data for the current study.

Finally, considering the reviewer's point concentrates on the *FOXO1* hotspot mutation specifically, we can evaluate the error profile of sequencing reads across this particular base in the genome. In our recently published paper using whole genomes to study normal endometrium, we used exactly the same microdissection, sequencing and mutation-calling pipeline as used here[6]. The reference base for the *FOXO1* mutation is a G; the S22W mutation changes this to a C. In the ~400 genomes from the endometrium paper, there were 10,803 reads reporting the G reference, versus zero reads reporting a C at this genomic position. Compared to the 45 separate microdissections reporting the S22W mutation, typically with variant allele fractions of 10-40%, it is clear that sequencing errors cannot explain our finding.

We have included the comparison with the endometrium data in Supplementary Note 2, describing the evidence for a recurrent hotspot mutation. We have referenced the high accuracy of our mutation calling in the main text (Results, 'Extended cohort of NAFLD patients', p. 4).

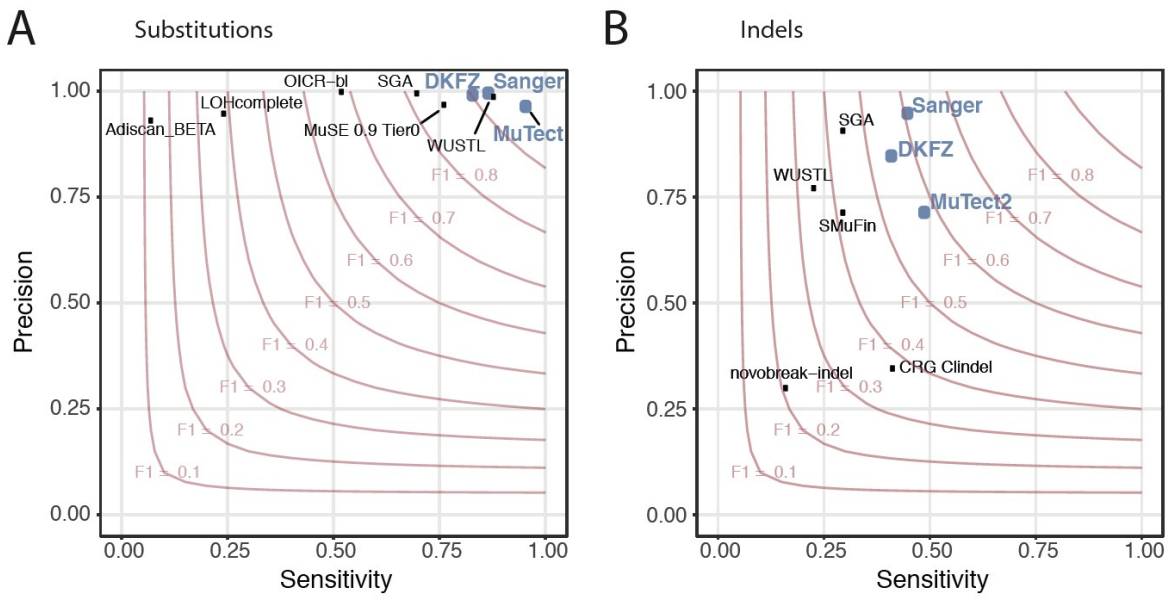
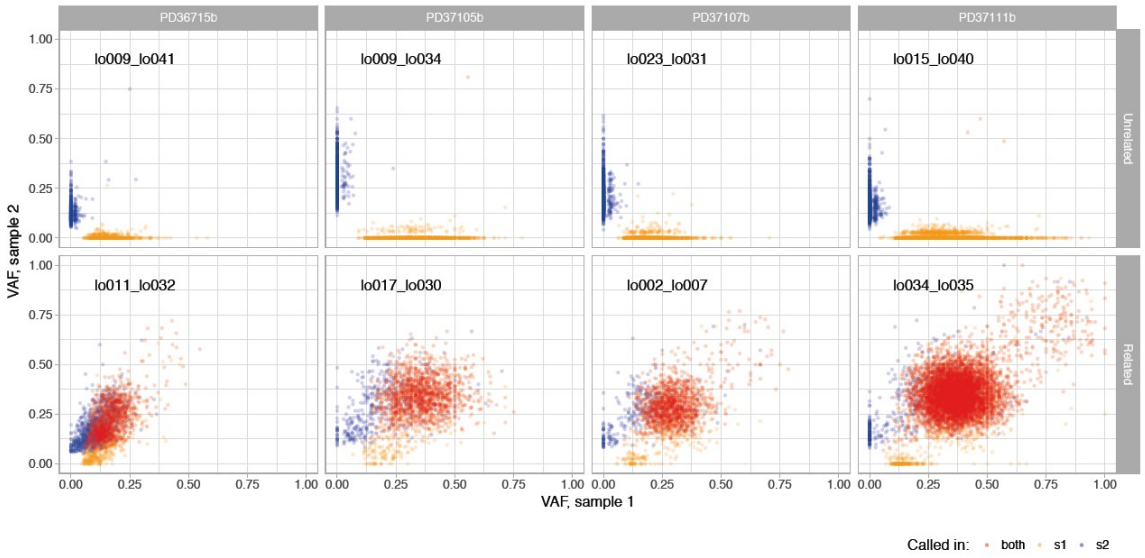


Figure 1A-B from the PCAWG marker paper[26]. Scatter plot of estimated sensitivity and precision for somatic SNVs (A) and somatic indels (B) across individual algorithms assessed in the validation exercise across n = 63 PCAWG samples. The algorithms used in the current manuscript are labelled 'Sanger'.



Extended Figure 1B from our original liver sequencing paper[10]. Examples of the variant allele fractions (VAFs) of variants from unrelated (top) and related (bottom) pairs of microdissection samples from four donors (left to right). The x axis represents the VAF of sample 1 from each pair and the y axis represents the VAF of sample 2. Each dot represents one variant. Red, variants called in both samples; yellow, variants called in sample 1; blue, variants called in sample 2.

2) Experimental validation: As >10% of genomic reads harbor the driver mutation, this could be validated using allelic real time PCR using DNA as a template to quantify the prevalence of the variant independently. This method is for instance used widely to quantify template before hybridizations of DNA or after FFPE extraction.

As discussed in the previous point, we have extensively validated our approach, using the method of independent replicates of x-y-plane microdissections from adjacent z-sections. This demonstrates the veracity of the *FOXO1* S22W mutations – we have a number of neighbouring microdissections from the same clone harbouring these mutations. This is indeed why 45 microdissections report the S22W mutation, but we count only 24 clones – an average of ~2 microdissections per clone report the same mutation. The microdissections reporting the variant are always clustered in physical space, as we show in Extended Figure 6 of the current manuscript. This is a much more accurate form of validation than a real-time PCR measurement of a single base, because of course we have the additional supporting evidence from the hundreds to thousands of other shared passenger mutations on the same branch of the phylogenetic tree. To spell the argument out, while it might be conceivable for some sequencing artefact to manifest as a shared false positive call at one genome position in adjacent microdissections, it is inconceivable that hundreds of such sporadic sequencing artefacts would so exactly match.

We have included these points in Supplementary Note 2 to provide further confirmation of the veracity of the *FOXO1* mutation calls.

3) Such a quantitative allelic PCR could be used on the large liver biobanks of NAFLD and ALD even without microdissection and on FFPE sections to independently replicate the finding.

While appealing, we are not convinced that this approach would give the accuracy required for large-scale clinical validation. It is true that within individual microdissections of our study, the *FOXO1* S22W mutation reaches VAFs of >10%, but within any whole liver sample, the overall VAF is considerably lower than this. There are several reasons for this: (1) not all clones within the sample carry the *FOXO1* mutation; (2) even within the microdissections up to 40% of cells are not hepatocytes, including Kupffer cells, endothelial cells and inflammatory cells; and (3) outside of hepatic nodules, there is extensive fibroblast infiltration generating the bands of fibrosis characteristic of cirrhosis.

The formaldehyde used in FFPE adds further complexity because it cross-links guanines, and the reference base we are interested in is a guanine. When we have studied sequencing (or allele-specific PCR approaches) on FFPE material, we find that error rates at guanines are elevated over other bases, meaning lower signal/noise ratios and impaired accuracy for calling variants at low VAF.

We do absolutely agree with the reviewer that a high-throughput assay to quantify the *FOXO1* mutations in large numbers of liver samples is required for detailed clinical characterisation of these mutations. Such an assay is non-trivial – we think the new wave of *in situ* transcriptomic assays might hold the key to this, but these are still under development.

We have included comment in the Discussion about the need for high-throughput, large-scale correlation of mutation burden with clinical phenotypes (Discussion, p. 13).

4) Why is this key somatic mutation not present in HCCs arising in the context of NAFLD and ALD? The authors quote the ICGC data and state that in over 1670 in only 0.18% of tumors carry a FOXO1 mutation. How does this look in the ALD/NAFLD subsets as compared to viral hepatitis? As reported by the authors, 15-20% of patients carry FOXO1 mutations in 5-40% of their hepatocyte clones (and these are further key pathogenetic drivers) one would expect 10-50 fold more positive HCCs. Are FOXO1 mutations selectively repaired in cancerogenesis or is FOXO1 specifically protecting clones from malignant transformation?

This is an absolutely fascinating question. We believe that the reviewer's final explanation is the most likely one – that, somehow, *FOXO1* mutations do not confer a selective advantage for malignant transformation (and may even be selected against). It is striking that of the 5 genes we find significantly mutated, *ACVR2A* and *TNRC6B* are also significant in HCC, with the same distribution of inactivating mutations. However, the three metabolism genes we identify, *FOXO1*, *CIDEB* and *GPAM*, are not significant in the 1670 HCCs. While mutating these genes might easily confer a selective advantage on hepatocyte clones living in a microenvironment with high levels of insulin signalling and caloric excess, malignant transformation engenders a microenvironment of local nutrient deficiency, hypoxia and metabolic stress. The Warburg effect, where cancer cells upregulate anaerobic metabolic pathways because their mitochondria cannot match the demands for energy production, is prominent in HCC[27] – in hepatocytes with impaired metabolism through say *FOXO1*, *CIDEB* or *GPAM* mutations, the resulting energy deficit may inhibit transformation. Clearly, understanding the prognostic implications of the novel mutations will require larger cohorts of patients with much longer follow-up, which will require new technologies to deliver at scale (see response to (3) above).

We have provided a new section in the Results containing a comparison of the rates of mutation in the genes reported here with those seen in HCC (Results, 'Comparison with selective landscape of HCC', p. 9).

5) More general: As liver contains ~20% of non-parenchymal cells (endothelials cells, HSCs, Kupfer cells etc.) – have the authors checked their microdissections for hepatocyte content (e.g. by RNASeq)?

One of the advantages of using a laser-capture microdissection approach is that we can directly visualise the regions that we cut out for sequencing. We chose these to be from regenerative nodules and so they were therefore enriched with hepatocytes. As the reviewer indicates, liver tissue does contain non-hepatocytic cells intermixed with hepatocytes – and of course we wind up sequencing the DNA from these cells as well. This manifests as a decrease in variant allele fraction for mutations arising in hepatocytes. If we consider some of the large monoclonal regenerative nodules we have studied here, for example, their monoclonal nature means that all hepatocytes within the nodule derive from a single ancestral clone, and therefore share a set of common mutations. The variant allele fraction of these shared mutations is typically 20-35% across multiple microdissections from the nodule, implying that this hepatocyte clone accounts for 40-70% of cells per microdissection. We believe the other 30-60% of cells include endothelial cells, Kupffer cells, inflammatory myeloid cells and lymphocytes. These interspersed cell types have different foetal origins and are therefore clonally unrelated to one another and to the hepatocytes, and therefore do not confuse the mutational landscape extracted (except by diluting the variant allele fraction).

We have now explained these points in Supplementary Note 1, discussing the rationale and features of our experimental design.

Functional exploration:

FOXO1 is proposed to be a major pathogenetic switch (in a subset of patients):

1) The authors are in the fortunate situation to have adjacent sections from these very patients available. With the large quantitative effects (up to 40% of observed reads with FOXO1 mutations in Figure 1) one could investigate the nuclear translocation directly in the patient sections harboring the mutations. One could perform RNASeq from microdissected tissue to look for characteristic downstream signatures of altered FOXO1 signaling, could perform immunohistochemistry or RNA in situ hybridization. Samples with a similar clinical profile but lack of FOXO1 somatic mutations that are already in the hand of the investigators would provide ideal controls.

We have attempted to perform RNA-sequencing by microdissection of adjacent sections on the patients characterised here. Unfortunately, we were unable to obtain sufficiently high-quality yields of sequencing reads from any given sample – the RNA was degraded for many sections and, in the RNA-sequencing reads we did obtain, it proved difficult to categorically match the mutations we expected to find from the genomes with those expressed in the transcriptome. Therefore, these data are regrettably not of sufficiently high standard for publication.

We have, however, performed more detailed functional evaluation of the cell lines transfected with wild-type or S22W *FOXO1-GFP*, including characterising levels of 105 metabolites in 5 replicates, with and without insulin. Overall, 32 metabolites were significantly different between S22W and wild-type constructs, with many intermediates in the gluconeogenesis pathway elevated in cells with mutant FOXO1-GFP, including malate, glycerol-3-phosphate, fructose-6-phosphate and glucose-6-phosphate. This suggests that the recurrent FOXO1 mutations do indeed induce changes in hepatocyte metabolism.

The new metabolomic data have been shown in Figure 2c and discussed in the main text (Results, 'Mutations impair nuclear export of FOXO1', p. 7).

2) The cellular experiments are not statistically analyzed. How many replicates were used? Can the effects be reproduced by Western blotting of isolated cellular fractions? Are the results stable in another cell line? Can we observe a typical FOXO1 downstream signature (e.g. by RNASeq)?

Three replicates for the cell line experiments have been performed, and they were performed in three HCC cell lines (HepG2, Hep3B and PLC/PRF5), each showing the same pattern of results (shown in Figure 2 and Extended Figure 8). We have now performed Western blotting using an antibody raised against phosphorylated T24 of the FOXO1 protein. This shows that the wild-type construct is phosphorylated upon insulin signalling, as expected, but the R21L and S22W mutant constructs do not show antibody binding, consistent with recent findings showing that the S22 AMPK phospho-site dominantly controls phosphorylation of the T24 Akt phospho-site[28]. We also have generated live cell imaging of the response of the construct to insulin exposure – this shows rapid nuclear export of the wild-type construct, as expected, but prolonged nuclear retention of the mutant constructs even with sustained insulin signalling. In particular, the live-cell imaging has now enabled us to quantify the time-course of changes in the nuclear-cytoplasmic ratio of mutant versus wild-type FOXO1-GFP after insulin exposure. Finally, as discussed in the previous point, we have now also performed metabolomic analysis in 5 replicates, with and without insulin, on cells with either S22W or wild-type *FOXO1-GFP* constructs.

The fluorescence images of the three cell lines are shown in Figure 2A and Extended Figure 9; the Western blotting is shown in Extended Figure 8B; the quantification of live-cell imaging data shown in Figure 2B; and the metabolite data in Figure 2C.

3) Do the (large) existing GEO liver and NAFLD expression datasets harbor a FOXO1 signature in the proposed ~15 percent of patients with this pathway? Do they form a distinct cluster?

We have assessed this, but the results are unclear. Existing studies have tended to use bulk RNA-sequencing of large samples of liver, with complex admixtures of multiple hepatocyte clones and other cells (fibroblasts, inflammatory cells, Kupffer cells, endothelial cells) – many of the changes observed in NAFLD and ARLD relate to changes in these other cellular populations. Also, the patchy nature of the *FOXO1* mutations may make it difficult to spot the expected transcriptional signatures. We believe that spatial transcriptomic studies coupled with genomics like the studies performed here will ultimately be the best way to perform these assessments, but these technologies are still under development.

D: Appropriate use of statistics and treatment of uncertainties

See above. In addition, the manuscript analyzed 32 liver samples from 27 patients (i.e. there were multiple samples from the same patient). The observations within one patient are likely not independent. It is not clear, how the authors controlled for this factor in their statistical testing.

The reviewer is absolutely correct that our design of multiple samples from the same patient (an average of 38 whole genomes per patient) means that sets of mutation calls from different microdissections will not necessarily be independent. To avoid false over-confidence when drawing statistical inferences, we must account for this dependence structure in the methods used – we did this for all statistical analyses reported in the paper. The specific details for each analysis are described below –

- *Reconstruction of phylogenetic trees* – Examples of high-resolution phylogenetic trees for each patient's liver sample are shown in Figures 1 and 3. To generate the trees, we used a multidimensional Dirichlet process clustering algorithm that essentially takes the observed VAF of every mutation in every microdissection from that patient, and identifies the unique clusters that emerge. From these clusters, we then built phylogenetic trees – the logic underlying the tree construction has been described and formalised in our previous papers[10,29], and has been adopted by other groups including the pan-cancer PCAWG consortium[30]. Having reconstructed the phylogenetic trees, every mutation was then placed onto a single branch through its most likely cluster membership. This means that any given mutation, which might be found in multiple microdissections because those regions shared a common ancestor, will be located in one, unique position in the phylogenetic tree, and therefore only 'counted' once.
- *Inference of recurrently mutated genes* – For this, we used an extensively tested and validated algorithm we originally built for inference of driver genes in cancer[17,31]. While normally this would take as input somatic mutation calls from whole genomes, we were concerned, as the reviewer highlights, that we would be double-counting mutations. The input we therefore used were the sets of mutations assigned to each branch of the phylogenetic tree – this means that each mutation is only counted once, in the ancestor in which it occurred, rather than the multiple times it is called in different, clonally related microdissections.
- *Inference of factors influencing telomere lengths* – Here, the clonal relationships among different microdissections make inference particularly challenging, largely because telomere lengths are heritable (the telomere length of the mother cell will be passed on to the daughter cells with some drift). To address this, we used a Bayesian linear mixed model specifically designed for inference on phylogenetic trees[32]. In these models, the random effect design matrix is a 'relatedness' matrix that captures the degree of shared ancestry between all pairs of microdissections. The analysis then estimates and controls for the extent to which variance in telomere lengths can be explained by clonal relationships, while simultaneously estimating the contributions of other effects such as age and presence of disease.
- *Inference of mutational signatures* – Here, we used Hierarchical Dirichlet processes[33]. As for inferring recurrently mutated genes, each mutation was placed onto a specific branch of the phylogenetic tree so that it was only counted once in the analysis. The 'hierarchical' feature of the hierarchical Dirichlet process then enabled us to group these individual branches as coming from

specific samples. The statistical analysis then explicitly allows for different liver samples to have different intensities of mutational signatures, and individual microdissections within each sample to vary around that patient's average distribution.

In summary, all of the statistical analyses performed in the manuscript rigorously model and correct for the dependence structure imposed by the phylogenetic tree. We are confident that the approach is robust and correct, and would be delighted for the paper to undergo statistical review if the editors deem this necessary.

We have briefly explained the measures for controlling non-independence in the main text (Results, 'Extended cohort of NAFLD patients', p. 4), with a more detailed explanation in the new Supplementary Note 1 dealing with the rationale for our experimental design.

E: Conclusions: robustness, validity, reliability

As the authors make fundamental contributions to the pathogenesis of insulin resistance in liver and NAFLD, the authors could enhance the study by a) the investigation of an appropriate cohort, b) a more complete characterization of existing samples, c) replication of the core finding in an independent cohort d) experimental validation of their mechanistic findings and e) a deeper functional exploration of the effects of their core finding.

We are pleased that the reviewer comments on the fundamental importance of these discoveries, and hope that the detailed responses above address the robustness of the finding, the details of the cohort studied and the functional consequences of the mutations.

Referee #3 (Remarks to the Author):

The article by Ng et al., aims to comprehensively characterize driver mutations in chronic liver disease. The analysis is particularly focused on identifying somatic mutations that may be important in the progression of chronic liver disease to hepatocellular carcinoma (HCC). The authors identified 4 significant genes ACVR2A, FOXO1, GPAM and CLCN5 and focus their attention on FOXO1 and GPAM. Of note, neither FOXO1 nor GPAM mutations were major contributors to HCCs (sequenced in ICGC).

The physiological implications for the most prevalent FOXO1 mutations were assessed in vitro and confirmed to impair nuclear export of FOXO1. In vitro analysis was not conducted on GPAM mutations though it was predicted to impair de novo lipogenesis. Both FOXO1 and GPAM mutations exhibited evidence for convergent acquisition of somatic mutations. As the authors discuss, these findings are suggestive of a novel concept that "somatic mutations might play a contributory role to systemic metabolic dysfunction."

Ng and colleagues also quantified telomere lengths at clone-specific resolution. Interestingly, they showed that telomeres became progressively shorter as the size of a clone increased and that the effect sizes in chronic disease considerably outweighed the relatively minor shortening of telomere lengths with age.

In summary, the article provides a comprehensive and large genetic analysis of human liver in the context of 2 common chronic liver diseases. Mutations in FOXO1 and GPAM strongly suggest convergent acquisition of somatic mutations that may lead to dysregulated insulin signaling in the whole tissue. Despite the comprehensive genetics of this study, the implications and generalizability of the findings are limited.

Comments:

1. The study appears to be a logical progression from previous studies with increased sample numbers and conclusions drawn from this extended analysis are in line with, and build upon, prior observations. The novelty of the study lies in the observation of convergent somatic mutations in FOXO1 and GPAM. Expression changes in FOXO1 and GPAM are well understood to disrupt hepatic lipid and glucose metabolism. However, apart from limited in vitro characterization, the functional metabolic consequences of the accumulated mutations are unclear, and only a matter of speculation in the discussion.

We divide our response to this point into three themes:

1. *In vitro* functional characterisation

We agree that in the original submission, the functional characterisation of the *FOXO1* mutations *in vitro* was relatively limited. To address this, we have performed extra experiments on the hepatocyte lines transfected with wild-type or S22W *FOXO1-GFP*. These include:

- *Live-cell imaging data* – We have extended the *in vitro* data with live cell imaging to show the rapid export of wild-type FOXO1 from the nucleus upon exposure of the cells to insulin. In

contrast, the S22W and R21L mutations prevent the FOXO1 construct from being exported from the nucleus, leading to sustained nuclear staining for protracted duration after insulin exposure. The live-cell imaging has also enabled us to quantify the nuclear-cytoplasmic ratio of mutant versus wild-type FOXO1-GFP over time after insulin exposure, showing striking differences.

- *Western blotting for phosphorylated T24* – We have now performed Western blotting using a commercially available, widely used antibody raised against phosphorylated T24 of the FOXO1 protein. This shows that the wild-type construct is phosphorylated upon insulin signalling, as expected, but the R21L and S22W mutant constructs do not show antibody binding.
- *Metabolite levels in mutant versus wild-type FOXO1-GFP* – We have characterised levels of 105 metabolites in 5 replicates of hepatocyte lines transfected with S22W versus wild-type FOXO1-GFP, with and without insulin. Overall, 32 metabolites were significantly different between S22W and wild-type constructs, with many intermediates in the gluconeogenesis pathway elevated in cells with mutant FOXO1-GFP, including malate, glycerol-3-phosphate, fructose-6-phosphate and glucose-6-phosphate.

The fluorescence images of the three cell lines are shown in Figure 2A and Extended Figure 9; the Western blotting is shown in Extended Figure 8B; the quantification of live-cell imaging data shown in Figure 2B; and the new metabolite data in Figure 2C.

2. Recurrence of driver mutations throughout the liver

We accept the reviewer's point that expression changes in *FOXO1* and *GPAM* are already known to disrupt hepatic glucose and lipid metabolism – indeed, this is one of the facts that makes the discovery of convergent somatic mutations in these genes so interesting. As the reviewer states, the major novelty in our data is the concept that not only are these genes somatically mutated in a reasonable fraction of patients with chronic liver disease, but they are *convergently* mutated in multiple independent clones within the same patient's liver. However, this conclusion was based on a single, small (<1cm^3) sample of liver per patient, and we could not be sure how representative the changes were of the whole liver.

To assess whether the changes from a single tissue sample are generalisable to the entire liver, we have now generated more sequencing data from 2 new patients with NAFLD in whom we took samples from each of the 8 Couinaud anatomical segments of the liver. In each of these, we then performed 22-28 microdissections, meaning that overall 388 new whole genomes were added to the study. The key insight from these additional genomes is that the *FOXO1* mutations are indeed generalised through the liver – one patient had the mutations in 3 separate anatomical segments, and the other in 4 segments (and up to 4 independent clones per segment, too). This argues that the changes seen in a single sample are indeed generalisable to the whole liver.

We note that a third gene involved in lipid metabolism, *CIDEB*, also emerged from the additional genomes. One of the patients had no fewer than 15 different mutations in the gene, again emphasising

the importance of convergent evolution of driver mutations in chronic liver disease. The gene is involved in the fusion and growth of lipid droplets in hepatocytes, and the mutations we observe are clustered in residues that are known to be important for homodimerisation of *CIDEB* during droplet fusion. This has obvious implications for why such mutations would confer a selective advantage on hepatocytes.

We have included discussion of the additional 388 genomes from the two new patients throughout the manuscript, such as the analyses for driver mutations, the new section on *CIDEB* and the characterisation of mutational signatures.

3. Systemic metabolic consequences of driver mutations

We absolutely agree with the reviewer that we have not studied the systemic metabolic effects of these mutations in our patients. There are many questions that remain about the clinical significance of these mutations, as all three reviewers point out, in particular what stage in the disease process driver mutations emerge and whether the burden of driver mutations correlates with disease severity.

We believe it will take 5-10 years, many hundreds, even thousands, of patients, and prospective cohort studies to address these questions with any degree of certainty. As the reviewers point out, they are undoubtedly important follow-on studies for our discovery, but we would argue that they are well beyond the scope of this paper. We believe that it is appropriate to speculate on the potential implications of our findings in the Discussion, as long as we are clear, first, that is a model / hypothesis and not conclusively proven here and, second, what studies would need to be performed to test the model.

We have therefore reframed the discussion more clearly to emphasise that it is only a model we are proposing and that much correlative clinical / metabolic follow-up needs to be performed (Discussion, pp. 12-13).

2. Although large numbers of analyses have been performed, these have been applied a very limited number of liver samples from a single center with 2 distinct types of metabolic liver diseases (NAFLD and ARLD). Each of these diseases is distinctly heterogeneous from the standpoints of disease phenotype, genetic susceptibility, gender, age and ethnicity. Whether the mutational pressure that yielded the observed mutations is generalizable is not clear.

As the reviewer indicates, our experimental design focused on analysing multiple whole genomes per patient liver sample – we averaged 33 whole genomes per sample across each of the 34 patient samples. The major reason for this experimental design was that our earlier work had shown that both healthy and diseased liver is polyclonal, with many independent clones carrying distinct catalogues of somatic mutations[10]. Not only is liver polyclonal, but there is striking clone-to-clone variation in the genomic changes seen – we find that even clones that are directly adjacent can exhibit 3-fold variation in mutation burden; mutational signatures that are completely absent from one accounting for >50% of mutations in the other; driver mutations present in one but not the other; widely varying telomere lengths. If we had chosen a design where we sequenced a single genome per sample, this remarkable within-patient, clone-to-clone variation would have passed undiscovered.

Beyond this, a critical component of our story is the parallel evolution of *FOXO1*, *GPAM* and, now, *CIDEB* mutations. For somatic mutations to play a role in a systemic disease, there must be a sufficient burden of cells carrying the mutations to cause organ-wide dysfunction. The fact that up to 5 clones within $1cm^3$ of liver can evolve the exact same mutation in the master regulator of insulin signalling is remarkable – it speaks to the strong selective pressures these clones are under. If we had analysed a single genome per sample, this parallel evolution would have been unappreciated (and we would have missed the mutation in many samples where it was actually present). Our new data, where we sequenced microdissections from samples across all 8 Couinaud anatomical segments of the liver, confirms that the findings from one tissue sample do indeed generalise to the whole liver.

However, as the reviewer indicates, the corollary to this experimental design is that we have analysed only moderate numbers of patients. Our cohort is indeed heterogeneous, and deliberately so – we believe this actually increases the generalisability of our discovery of *FOXO1* and *GPAM* mutations. For example, taking the factors the reviewer specifically comments on, we have:

- Disease phenotype – 2 patients with ARLD and 3 patients with NAFLD carried *FOXO1* mutations; 3 patients with ARLD and 3 with NAFLD carried *GPAM* mutations; and 3 patients with ARLD and 3 patients with NAFLD had *CIDEB* mutations;
- Genetic susceptibility – Considering the most penetrant inherited risk factor for cirrhotic liver disease[34], the *PNPLA3* polymorphism (rs738409), all three genotypes (CC, CG and GG) were present among patients with *FOXO1* mutations, all three were present among patients with *GPAM* mutations, and all three were present among patients with *CIDEB* mutations;
- Gender – 3 females, 4 males had *FOXO1* mutations; 2 females, 5 males had *GPAM* mutations; and 1 female, 5 males had *CIDEB* mutations;
- Age – Patients with *FOXO1* mutations ranged in age from 37 to 66 years; patients with *GPAM* mutations from 40-67 years; and with *CIDEB* from 40-72 years.
- Ethnicity – All our patients were white European, as expected for the local demographics in the Cambridge, UK, region.

Thus, despite the moderate cohort size, the *FOXO1* mutations, *CIDEB* mutations and *GPAM* mutations are seen across a wide range of patient characteristics. For sure, with such heterogeneity, it will take

cohorts of hundreds to thousands of patients to fully characterise the clinical associations, but we believe on the basis of the data collected to date that the mutations in all 3 genes will be widespread in chronic liver disease, across all the variables listed by the reviewer and other dimensions of disease. On this basis, we are confident that our discovery will be generalisable.

We have included all demographic and other clinical details on the patients with *FOXO1*, *CIDEB* and *GPAM* mutations in Supplementary Table 1.

3. In this study, the control group was normal liver. It would be important to understand whether other chronic liver disease that may not ultimately exhibit phenotypes of insulin resistance (e.g. chronic cholestatic diseases or autoimmune liver disease) accumulate similar mutations at similar burdens.

We agree with the reviewer that it will be interesting to assess whether these genes are also mutated in other causes of liver disease. We have not assessed these, and there are indeed many different entities to explore – cholestatic liver disease, autoimmune hepatitis, as the reviewer mentions, but also inherited causes of liver disease, viral liver disease and so on. We believe such studies are beyond the scope of the current study. We note that with the advent of effective therapies for hepatitis C, alcohol-related liver disease and NAFLD are now the two commonest causes of chronic liver disease in the West, with NAFLD showing alarming rises in prevalence as obesity becomes more widespread[35]. Therefore, any discoveries about fundamental mechanisms of disease in these two forms of cirrhosis are important, whether or not they apply in other liver diseases as well.

We have acknowledged the interest in assessing mutational landscape of other forms of liver disease in the Discussion (p. 13).

4. There is no meaningful analysis of the overall histopathology (or accounting of patient characteristics) for either type of liver disease studied, nor are there provided the histopathologic features present at the sites of laser capture. Standardized methodologies are available for these characterizations.

All liver samples studied here were independently assessed by a central histopathologist, and slides were scored for the variables captured in Supplementary Table 1. Variables scored include:

- Steatosis on a 0-3 scale

- Presence of lobular inflammation on a 0-3 scale
- Presence of hepatocellular ballooning (seen in NAFLD) on a 0-2 scale
- The NAS score, a measure of non-alcoholic steatosis, on a 0-8 scale
- Fibrosis stage, as a Kleiner score 0-4 and Ishak score 0-6
- Siderosis grade on a 0-4 scale
- Qualitative assessment of global dysplasia in the background liver

For those patients in whom an HCC developed, we have also included a range of histopathological metrics for the tumours, including:

- Size of largest HCC (mm)
- Number of HCCs
- Total tumour load diameter (mm)
- Presence of microvascular invasion
- Presence of large vessel invasion
- Worst differentiation grade

These are all standard histopathological measurements, using well-validated quantitative and ordinal scales. We accept that in the original submission, we did not highlight the extensive and careful review of histology that we had performed.

Furthermore, we show H&E sections for the patient studied at three timepoints in Extended Figure 1; the five patients with *FOXO1* mutations in Figures 1B-C, Extended Figure 7; 3 of the patients with *GPAM* mutations in Figure 3B, Extended Figure 9; and 4 samples with the new mutational signature in Figure 5B-C, Extended Figure 14. This gives a very representative overview of the histology in all of the most relevant samples to the stories we present.

We have signposted the central histological review and the complete data in Supplementary Table 1 more explicitly in the main text (Results, 'Extended cohort of NAFLD patients', p. 4).

5. For NAFLD, the authors fail to distinguish between simple steatosis (NAFL) and steatohepatitis. In many ways, these are quite distinct entities, with markedly different natural histories with respect to chronic liver disease.

We absolutely agree with the reviewer that simple steatosis is a very different entity to steatohepatitis in the early stages of disease. It is appreciated that as end stage liver disease develops and/or there is alcohol abstinence, the typical features of ongoing steatohepatitis and amount of steatosis can subside. All our patients had hard clinical endpoints relating to NAFLD or ARLD: the patients all had either HCC or liver failure or both. This was to ensure that if disease-related somatic mutations existed

we would find them in patients with more 'advanced' disease. One of the clear questions arising from our findings is the timepoint of acquisition of these mutations and its relationship to disease progression. As detailed in our responses above, this will form the basis of follow-on studies in earlier disease stages.

We now clarify that all our NAFLD patients had histological changes entirely in keeping with a NAFLD aetiology; this is even in the absence of steatohepatitis for which the presence of ballooning degeneration is required (Results, 'Extended cohort of NAFLD patients', p. 4).

6. The manuscript is extremely turgid and laden with technical jargon. For this to be of meaning to the broader hepatology and metabolic communities, it should be written in a much more accessible style.

We definitely agree with the reviewer that the core discoveries of the manuscript will have interest to the broader hepatology and metabolic communities, beyond those used to thinking about somatic mutations on a daily basis. We would very much like to make the manuscript accessible to all interested scientists, while retaining the detail necessary to satisfy the more exacting readers.

We have decreased the technical detail in the main text, while including expanded sections as two Supplementary Notes covering important aspects of the experimental design, considerations of statistical analyses and results that are too detailed for inclusion in the main text.

REFERENCES

1. Martincorena, I. *et al.* High burden and pervasive positive selection of somatic mutations in normal human skin. *Science (80-.).* **348**, 880–886 (2015).
2. Lee-Six, H. *et al.* Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).
3. Lee-Six, H. *et al.* The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* **574**, 532–537 (2019).
4. Martincorena, I. *et al.* Somatic mutant clones colonize the human esophagus with age. *Science (80-.).* **917**, 911–917 (2018).

5. Yokoyama, A. *et al.* Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature* **565**, 312–317 (2019).

6. Moore, L. *et al.* The mutational landscape of normal human endometrial epithelium. *Nature* **580**, (2020).

7. Yoshida, K. *et al.* Tobacco smoking and somatic mutations in human bronchial epithelium. *Nature* **578**, 266–272 (2020).

8. Blokzijl, F. *et al.* Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* **538**, 260–264 (2016).

9. Yizhak, K. *et al.* RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science (80-.).* **364**, eaaw0726 (2019).

10. Brunner, S. F. *et al.* Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. *Nature* **574**, 538–542 (2019).

11. Zhu, M. *et al.* Somatic Mutations Increase Hepatic Clonal Fitness and Regeneration in Chronic Liver Disease. *Cell* **177**, 608–621 (2019).

12. Kim, S. K. *et al.* Comprehensive analysis of genetic aberrations linked to tumorigenesis in regenerative nodules of liver cirrhosis. *J. Gastroenterol.* (2019). doi:10.1007/s00535-019-01555-z

13. Brazhnik, K. *et al.* Single-cell analysis reveals different age-related somatic mutation profiles between stem and differentiated cells in human liver. *Sci. Adv.* **6**, 1–11 (2020).

14. Tapper, E. B. & Lok, A. S.-F. Use of Liver Imaging and Biopsy in Clinical Practice. *N. Engl. J. Med.* **377**, 756–768 (2017).

15. Suda, K. *et al.* Clonal Expansion and Diversification of Cancer-Associated Mutations in Endometriosis and Normal Endometrium. *Cell Rep.* **24**, 1777–1789 (2018).

16. Imielinski, M., Guo, G. & Meyerson, M. Insertions and Deletions Target Lineage-Defining Genes in Human Cancers. *Cell* **168**, 460-472.e14 (2017).

17. Rheinbay, E. *et al.* Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature* **578**, 102–111 (2020).

18. Scelo, G. *et al.* Variation in genomic landscape of clear cell renal cell carcinoma across Europe. *Nat. Commun.* **5**, 5135 (2014).

19. Hoang, M. L. *et al.* Mutational Signature of Aristolochic Acid Exposure as Revealed by Whole-Exome Sequencing. *Sci. Transl. Med.* **5**, 197ra102-197ra102 (2013).

20. Li, J. Z. *et al.* Cideb Regulates Diet-Induced Obesity, Liver Steatosis, and Insulin Sensitivity by Controlling Lipogenesis and Fatty Acid Oxidation. *Diabetes* **56**, 2523–2532 (2007).

21. Barneda, D. *et al.* The brown adipocyte protein CIDEA promotes lipid droplet fusion via a

phosphatidic acid-binding amphipathic helix. *Elife* **4**, 1–24 (2015).

22. Sun, Z. *et al.* Perilipin1 promotes unilocular lipid droplet formation through the activation of Fsp27 in adipocytes. *Nat. Commun.* **4**, (2013).

23. Ioannidis, J. P. A., Thomas, G. & Daly, M. J. Validating, augmenting and refining genome-wide association signals. *Nat. Rev. Genet.* **10**, 318–329 (2009).

24. Eslam, M. & George, J. Genetic contributions to NAFLD: leveraging shared genetics to uncover systems biology. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 40–52 (2020).

25. Hirsova, P., Ibrabim, S. H., Gores, G. J. & Malhi, H. Lipotoxic lethal and sublethal stress signaling in hepatocytes: relevance to NASH pathogenesis. *J. Lipid Res.* **57**, 1758–1770 (2016).

26. Pan-cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature* **578**, 82–93 (2020).

27. Iansante, V. *et al.* PARP14 promotes the Warburg effect in hepatocellular carcinoma by inhibiting JNK1-dependent PKM2 phosphorylation and activation. *Nat. Commun.* **6**, (2015).

28. Saline, M. *et al.* AMPK and AKT protein kinases hierarchically phosphorylate the N-terminus of the FOXO1 transcription factor, modulating interactions with 14-3-3 proteins. *J. Biol. Chem.* **294**, 13106–13116 (2019).

29. Nik-Zainal, S. *et al.* The life history of 21 breast cancers. *Cell* **149**, 994–1007 (2012).

30. Gerstung, M., Jolly, C., Leshchiner, I. & Dentro, S. C. The evolutionary history of 2,658 cancers. *Nature* **578**, 122–128 (2020).

31. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029–1041 (2017).

32. Hadfield, J. D. & Nakagawa, S. General quantitative genetic methods for comparative biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* **23**, 494–508 (2010).

33. Neal, R. M. Markov Chain Sampling Methods for Dirichlet Process Mixture Models. *J. Comput. Graph. Stat.* **9**, 249–265 (2000).

34. Trépo, E., Romeo, S., Zucman-Rossi, J. & Nahon, P. PNPLA3 gene in liver diseases. *J. Hepatol.* **65**, 399–412 (2016).

35. Pimpin, L. *et al.* Burden of liver disease in Europe: epidemiology and analysis of risk factors to identify prevention policies. *J. Hepatol.* (2018). doi:10.1016/J.JHEP.2018.05.011

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

The authors added a huge number of new data and they have taken into account all my comments. I have no additional question. Congratulations.

Referee #2 (Remarks to the Author):

The authors have analyzed two additional patients. My main concerns have not been alleviated, however. The take home message for me is that interesting patterns of somatic mutation are observed. However it is at this point not possible to replicate them in larger numbers of patient samples or to decide if their specifics are biologically relevant.

1) For inference of mechanisms of general disease pathogenesis, replication of findings in larger patient cohorts is needed, because human disease pathogenesis is notoriously heterogenous. This is apparently not technically possible here (answer to question 3 somatic variant) because no assay other than whole genome somatic sequencing is currently feasible.

2) It is unclear if the findings of the authors are of functional relevance in human liver.
2a) The authors are unable to assess functional consequences in the explanted liversthey investigated by somatic mutation analysis. RNA-sequencing was attempted and was not possible because of RNA degradation (see answer to functional exploration 1). Other options like RNA in situ hybridization or immunohistochemistry could also be possible but have not been performed.
2b) Signatures of the newly identified pathways are not visible in the existing large liver expression datasets (see answer to functional exploration 3). The authors point to the use of bulk tissue sequencing by many of the previous studies as a reason for this lack of resemblance. However, there is now a growing number of single cell and microdissected RNASeq datasets on human liver published, which could have been used.

3) It is unclear if the mutation-carrying clones identified are relevant to disease progression or are an epiphenomenon of certain clones in cirrhotic liver:
3a) They are not relevant for progression to hepatocellular carcinoma (see answer 4 in the somatic mutation section). In fact, these clones appear to be selectively protected from cancer transformation. I do not fully understand their de-enrichment by a factor of 10-50 from HCCs and the authors explanation. Potentially, this would be a somewhat counter-intuitive but nonetheless very attractive anti-cancer pathway to look into.
3b) They are not relevant for progression of fatty liver disease to cirrhosis (which is one of the key questions in liver research) because firstly the authors did not have access to such tissue (i.e. NASH, ALD) but were restricted to end-stage disease explants in their sample access. Secondly, other groups with access to such samples did not see signatures of the proposed pathomechanisms in their data (see 3a above).

4) I think in vitro experiments in hepatoma cell lines can only partially help with the functional exploration. The metabolomics would have been most instructive if it was performed on the human liver samples harbouring the specific mutations. Human liver metabolomics is established (using cryo-microdissection too) but likely not possible with the tissue quality available to the authors. If there is already significant RNA degradation (see 2a), metabolomics seems not to be feasible.

Referee #3 (Remarks to the Author):

The authors have largely responded to the prior critiques, and that manuscript has been improved considerably, Acaveat that the authors really shoult not characterize what they have studies as “chronic liver disease”, but an affect that is applicable for 2 common etiologies of liver disease,

ALD and NAFLD. Although they make an argument that these are the most common chronic liver disease, hepatitis B, hepatitis C, as well as other liver diseases are also common enough, even in the West, and not characterized by the current approach.

Referee #4 (Remarks to the Author):

In their manuscript entitled “Convergent somatic mutations in effectors of insulin signalling in chronic liver disease”, Ng and colleagues analyze more than 1400 whole genome sequences obtained from laser-capture microdissections of hepatocytes from patients with non-alcoholic fatty liver disease (NAFLD) or alcohol-related liver disease (ARLD). They find clonal expansions harboring a number of recurrently mutated genes, most prominently hotspot mutations in the transcription factor FOXO1, which they explore with additional experiments. Furthermore, they describe a hitherto unknown mutational signature and shorter telomeres associated with clonal expansions.

The manuscript builds upon a dataset published by the same group in 2019, as well as a number of conceptually related papers on somatic evolution of normal tissues that have been coming out in recent years. While I appreciate the effort and cost behind this large-scale tissue collection and sequencing project and find the authors’ description of FOXO1 mutations in chronic liver disease very interesting, I cannot help but find that overall, this paper does not quite reach the quality and definitiveness of comparable top publications in the field.

The issue is not technical - I have no concerns about the driver gene identification methods used in this manuscript (see further below for details) – but a problem of narrative. The authors begin with a collection of clinically heterogeneous patients (some have ARLD and other NAFLD). They group them all together to identify mutated genes that drive clonal expansions in liver disease and discover – convincingly – strong selection for a particular hotspot mutation in FOXO1. Here is where, in my opinion, the problems begin. Instead of building a strong narrative around this discovery, the paper fragments into a number of loosely connected storylines that struggle to find a definitive common denominator. There is some very basic investigation of the FOXO1 mutation in cell lines which does not go beyond showing that it interferes with correct subcellular localization of the TF and (unsurprisingly) some resulting changes in metabolite concentration. The vital question if or how these changes contribute to increased hepatocyte survival in the context of liver disease is treated purely speculatively, and only in the discussion. There, the authors speculate that selective pressure is exerted by “toxic lipid species and oxidative stress in hepatocytes processing the chronic dietary excess of calories”. At this point, the grouping of ARLD and NAFLD patients becomes a particular problem (as some reviewers have pointed out), as the underlying disease mechanisms are different, and yet both groups have FOXO1 mutations. The authors try to make it all fit together by saying that “similar pathways probably operate in ARLD”, but this seems regrettably thin and preliminary, given that these FOXO1 mutants are the main discovery of the paper. The remaining storylines – CIDEB, GPAM, telomere lengths, a new mutational signature – remain largely separate and undeveloped. All in all, I find it difficult to identify a clear conceptual advance, although the FOXO1 mutations are certainly interesting.

Another issue that contributed to the “preliminary” impression the paper made on me is the authors’ repeated and insistent allusion to a possible role of FOXO1 mutants in metabolic dysfunction at the organismal level. This happens again and again, in the abstract, the introduction, and extensively in the discussion. I am afraid I cannot support even these tentative statements. The data in this paper offer no suggestion whatsoever that the mutant cells contribute to disease at an organismal level. Without showing with a formal calculation how they arrived at this conclusion, the authors claim that mutants “could affect hundreds of grams” of liver, a vague number that is not further justified but nonetheless prominently featured in the abstract. In my opinion, the authors are hurting themselves by alluding to this possibility repeatedly (even mentioning it in the abstract) without evidence or even a formal quantification of mutant burden

vis-à-vis total liver volume across different patients (the latter should be straightforward).

I think some of these issues could potentially be solved, but they would require a pretty extensive restructuring of the paper, which the authors may or may not be willing to undertake. I understand that figuring out whether the mutants are causing any harm is a non-trivial task, but the much more obvious and approachable (and frankly, in the context of this paper, more important) question as to how FOXO1 mutations confer a competitive advantage to hepatocytes in the setting of liver disease really would have to be explored at least a little bit in order for the story to come full circle.

Some comparatively minor thoughts:

1. I found that a solid overview figure showing all patients, disease conditions, microdissections, detected driver variants and their VAFs etc. was really missing. Such a figure would provide a firm footing and reference for the reader as he/she embarks upon the journey of "vignettes" that make up the majority of this paper. Instead, in figure 1A (which is usually reserved for such an overview), the reader finds himself presented with a very detailed representation of the FOXO1 sequence across species, without having a particularly clear sense of how he got there from the analysis of 1000+ genomes from 34 livers.

2. Somewhat related to point 1 above: One of the prior reviewers asked the authors to provide an analysis correlating genomic features and clinical variables, and the authors replied by pointing to Supplementary Table 1. I think that providing this table does not equate to analysis and agree with the reviewer that it would be important to provide this. Even if the patient numbers are small and it may be hard to achieve statistical significance, this is an important exercise.

3. As I mentioned above, I have no doubt about the accuracy of the driver gene identification methods used in this manuscript. They are state-of-the-art methods, and many of them were developed by the authors themselves in a long line of prior work. In particular, I am not at all worried about potential artifacts caused by interdependence between samples from the same patient (a point raised during review), as the phylogenetic tree structures clearly demonstrate that these mutations are independent somatic events. The authors nicely address the remaining possibility that FOXO1 mutations could be embryonic variants in their supplementary note, although it would be good to present those arguments in a more formal manner. (E.g. "The difficulty with this explanation is that in many of the pairs of microdissections that share a FOXO1 mutation, this is literally the only shared mutation" – please show the data instead of simply stating this).

Author Rebuttals to First Revision:

Referee #1 (Remarks to the Author):

The authors added a huge number of new data and they have taken into account all my comments. I have no additional question. Congratulations.

We thank the reviewer for this endorsement of our manuscript.

Referee #2 (Remarks to the Author):

The authors have analyzed two additional patients. My main concerns have not been alleviated, however. The take home message for me is that interesting patterns of somatic mutation are observed. However it is at this point not possible to replicate them in larger numbers of patient samples or to decide if their specifics are biologically relevant.

1) For inference of mechanisms of general disease pathogenesis, replication of findings in larger patient cohorts is needed, because human disease pathogenesis is notoriously heterogenous. This is apparently not technically possible here (answer to question 3 somatic variant) because no assay other than whole genome somatic sequencing is currently feasible.

Our manuscript reports the discovery of genes under positive selection in NAFLD and ARLD – the key finding is that 3 genes involved in coordinating the metabolic function of the liver are targets for recurrent and convergent somatic mutation in these patients. These findings have been robustly replicated in our current cohort:

- The excess of mutations in all 3 genes is highly statistically significant (*FOXO1* $q<2x10^{-16}$; *CIDEB* $q<2x10^{-16}$; *GPAM* $q=1x10^{-5}$) – this demonstrates that it is almost impossible for the number and distribution of somatic mutations in these genes to have occurred through the accumulation of random, selectively neutral mutations;
- For each of the 3 genes, there are several patients who show mutations in multiple independent clones within the same biopsy – this demonstrates that convergent evolution of driver mutations occurs commonly in chronic liver disease;
- For 2 patients, we replicated the genomic studies across specimens from all 8 Couinaud segments – this demonstrates that the findings from a single, random liver biopsy generalise to the whole liver;
- The clinical features of the patients with somatic mutations in *FOXO1*, *CIDEB* and *GPAM* span the full range of ages studied, both sexes, known germline predisposition alleles and both disease phenotypes studied – this demonstrates that the findings are seen across a wide range of patient characteristics, and therefore have generalised relevance;
- At 1590 whole genomes, this is by some margin the largest study of somatic mutations in a normal organ system published to date, and is indeed larger than most whole genome studies of specific types of cancer (except maybe breast cancer and acute leukaemia);

We wholeheartedly agree with the reviewer that human disease pathogenesis is heterogeneous, and therefore understanding the clinical implications of mutations in these genes (and others that will likely be discovered in the future) will require large cohorts. We believe that such clinical correlations, while they are undoubtedly important future studies, are well beyond the scope of this manuscript – they will require new technical methodologies to screen large numbers of patient samples for the genes identified here. These methods do not yet exist at the throughput required for such

characterisation, and we believe it would be inappropriate to hold back the discovery of these genes further.

We have now included a section in the main text ('Properties of clones and patients with drivers', p. 9), with accompanying Extended Figure 12 and Supplementary Code, reporting the diversity of clinical characteristics of patients with mutations in the three metabolism genes. We have rewritten the Discussion (pp. 11-12) to remove the emphasis on whether these findings have implications for systemic metabolic disease, instead focusing on the implications for the nature of selective pressure and clonal evolution in chronic liver disease.

2) It is unclear if the findings of the authors are of functional relevance in human liver.

2a) The authors are unable to assess functional consequences in the explanted livers they investigated by somatic mutation analysis. RNA-sequencing was attempted and was not possible because of RNA degradation (see answer to functional exploration 1). Other options like RNA in situ hybridization or immunohistochemistry could also be possible but have not been performed.

The major challenge for studies such as RNA *in situ* hybridisation and immunochemistry is defining the clonal structure and driver mutation distribution across the section used for the imaging. As our data clearly demonstrate, chronic liver disease generates clones with driver mutations interspersed with regions populated by wild-type clones. It will be impossible to interpret transcriptional data from *in situ* methods without knowing which regions carry the driver mutation of interest.

However, *in vitro* studies can provide some insights into the transcriptional consequences of somatic mutations. We have therefore now generated RNA-sequencing data from the HepG2 cell lines transfected with wild-type, S22W and R21L *FOXO1-GFP* constructs. This reveals that the *FOXO1* mutations do indeed generate coordinated alterations in the transcriptome of hepatocytes, and in the expected direction. For example, gene sets up-regulated in constructs with *FOXO1* mutations included several implicated in regulation of cell cycle and mitosis (eg 'CELL CYCLE, MITOTIC: REACTOME R-HSA-69278.4': $q<0.0001$); known targets of FOXO1-mediated transcription (eg 'FOXO-MEDIATED TRANSCRIPTION OF OXIDATIVE STRESS, METABOLIC AND NEURONAL GENES: REACTOME 9615017': $q=0.0084$); and lipid catabolism pathways (eg 'LIPID CATABOLIC PROCESS: GO 0016042': $q<0.0001$). Interesting gene sets that were down-regulated with constructs containing *FOXO1* mutations included glycolysis pathways (eg 'CANONICAL GLYCOLYSIS: GO 0061621': $q=0.0001$) and apoptosis pathways (eg 'APOPTOTIC PROCESS: GO 0006915': $q=0.0004$). Thus, there are transcriptional changes in metabolism pathways seen in these cells that mirror the metabolomic changes we reported in the earlier revision to our manuscript. In addition, the *FOXO1* driver mutations are associated with up-regulation of proliferative signals, such as cell cycle/mitosis gene sets, and down-regulation of apoptosis signals – this could potentially contribute to the survival benefit clones with these mutations have in diseased liver.

These new RNA-sequencing data have now been included in the main text ('Mutations impair nuclear export of FOXO1', p. 6), together with a new Extended Figure 10 and Supplementary Tables S6-S7.

2b) Signatures of the newly identified pathways are not visible in the existing large liver expression datasets (see answer to functional exploration 3). The authors point to the use of bulk tissue sequencing by many of the previous studies as a reason for this lack of resemblance. However, there is now a growing number of single cell and microdissected RNASeq datasets on human liver published, which could have been used.

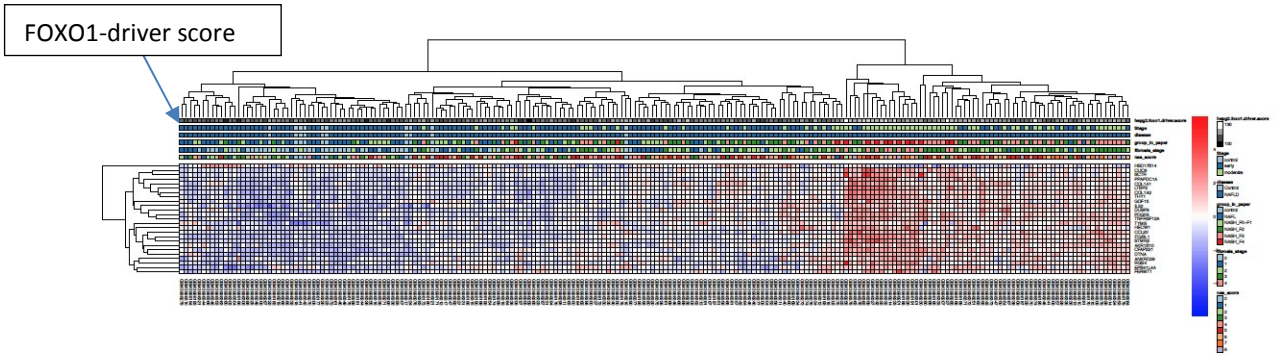
We agree that publications in single-cell transcriptome and laser-capture microdissection RNA sequencing have recently emerged. The single-cell transcriptome studies to date[1–4] have mostly used methods such as 10X Genomics, MARS-seq or CEL-Seq2, which unfortunately generate strong biases towards the 3' end of the transcript. This means that we get insufficient sequencing coverage of the S22 residue of *FOXO1*, and therefore cannot infer the mutational status of each hepatocyte for linking with its transcriptome.

The laser-capture microdissection datasets would potentially be more useful for our purposes, but mostly have been performed to answer different questions to the one we would pose (such as progression to cancer or stem cell niches or evaluating mouse models of liver disease). As a result, they typically have insufficient power to resolve clones with *FOXO1* (or other gene) mutations and their transcriptional differences.

As an example, consider a recent paper which used microdissection to isolate hepatic progenitor cells from patients with alcoholic steatohepatitis[5]. Here, 19 patients were studied, and the RNA-sequencing data we downloaded and reanalysed were of excellent quality, but illustrate the challenges of using published data for our questions. Considering the *FOXO1* hotspot, the median coverage of this position in the transcriptome across the 19 microdissections was 25 reads, so we only had power to detect large clonal fractions carrying the mutations. Of the 19 microdissections, one had evidence for the *FOXO1* S22W mutation – in our genome study, *FOXO1* S22W mutations were present in 45/1590 (2.8%) microdissections, so 1 in 19 is broadly as expected. For this one transcriptome, 3 reads reported the mutation compared to 22 reads reporting the reference base, suggesting that the mutant hepatocytes represented only a fraction of the hepatocytes sequenced or there was substantial contamination from stromal or inflammatory cells. Overall, then, it would be challenging to extract credible transcriptional signatures for the *FOXO1* mutation from this dataset.

Another approach we have tried, now that we have generated our own RNA-sequencing data from

the *in vitro* studies, is to identify the 25 most highly expressed genes in the *FOXO1*-mutant HepG2 cell lines ($p<0.01$, log fold-change > 0.5). Using this set of genes, we used regression to define a binary classifier (a weighted sum of expression levels for these 25 genes) to form a gene expression score that discriminates between *FOXO1*-mutant versus *FOXO1*-wild-type constructs in the HepG2 cell lines. We then applied this score to a published RNA-sequencing dataset from a large cohort of 216 patients with NAFLD[6,7] to see whether the score could predict severity of disease. Additionally, we also investigated whether the score was correlated with the expression of genes reported to be associated with aspects of NAFLD severity[6]. As shown in the panel below, our HepG2-derived *FOXO1*-driver score was found to be associated with the level of expression of 25 genes that were reported to be associated with steatosis, fibrosis, ballooning, and inflammation, as well as markers of severity of NAFLD, such as the NAFLD Activity Score (Spearman rho = 0.16, P=0.01), and fibrosis stage (Spearman rho = 0.37, $P=1.36\times10^{-8}$). However, while these data are suggestive that the expression changes associated with *FOXO1* mutations could predict clinical variables such as disease severity, the lack of information about whether the *FOXO1* mutations are actually present in the samples makes us reluctant to include the analysis in the manuscript.



Reviewer figure. Heatmap showing expression levels of 25 genes reported by Govaere et al[6] to be strongly associated with aspects of NAFLD severity. The 25 genes represent the rows of the heatmap and the 216 patients are shown in the columns. Our HepG2-derived *FOXO1*-driver gene expression scores across the cohort are displayed as the top annotation row, with white to light grey representing high scores and dark grey to black low scores. The NAFLD Activity Score (NAS) is shown as the bottom row of the classification tier (marked with an arrow), with blues to greens being low NAS scores and reds to purples being high NAS scores. Fibrosis staging information is displayed directly above the NAS score annotation row, with blues to light green indicating minimal fibrosis, while dark green and pink denote advanced fibrosis. Note that the patients with the most up-regulation of these 25 genes (predominantly red, to the right of the heatmap) also tend to have the most advanced fibrosis, and the highest HepG2-derived *FOXO1*-driver gene expression scores.

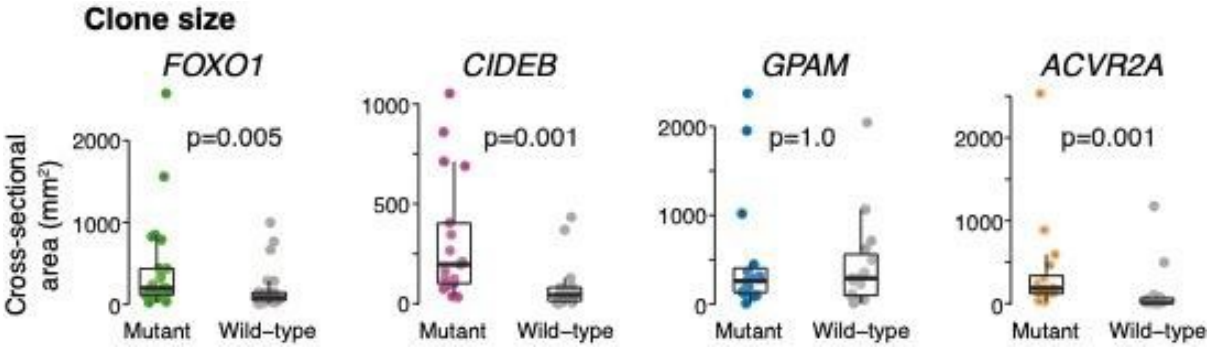
3) It is unclear if the mutation-carrying clones identified are relevant to disease progression or are an epiphenomenon of certain clones in cirrhotic liver:

The statistical analysis we base our inferences on identifies genes with an excess of non-synonymous mutations significantly above the number expected for the background mutation rate (estimated from the local synonymous mutation rate)[8,9]. For the vast majority of genes (>99.9%), both in cancer and in normal tissues (including this study of liver), the observed numbers of non-synonymous mutations are well within the range expected from this background rate. This implies, firstly, that the background models of mutation rate are well calibrated and accurate, and, secondly, that mutations in these genes are accumulating neutrally. Mutations in these genes are, to use the reviewer's term, 'epiphenomena': passenger mutations that have occurred randomly in the clones as a result of universal mutational processes active in our somatic cells throughout life.

Finding a significant excess of non-synonymous over synonymous mutations in a given gene implies that some of those mutations, known as driver mutations, confer a selective advantage on the clone – this has become the standard approach for discovering the genes driving cancer[10,11]. They cannot all be epiphenomena, because if they were all selectively neutral, they would be found at the equivalent density to the neutral synonymous mutations (after calibration). To illustrate this with a back-of-the-envelope calculation, we identified a total of 3,411,889 independent somatic mutations across the 1590 samples sequenced. With a genome size of ~3.1 billion bases, we would therefore expect an average of a little over 1/1000 mutations at any given base in our dataset (with maybe up to 10-fold variability across the genome through differential DNA repair etc[12,13]). Therefore, identifying 24 independent and identical variants at one base in *FOXO1* is extremely unlikely to happen by chance – they are enriched because clones that acquire this specific base change survive and expand better than clones with mutations spread randomly across the genome.

The genetics analysis, then, is strongly robust, and implies that the mutations in *FOXO1*, *CIDEB* and *GPAM* have conferred a selective advantage on clones in these patients; they are not epiphenomena. That mutations in these genes are not found in healthy liver suggests that the selective advantage they confer is stronger in the context of liver disease. What the genetics cannot tell us is what the cellular basis for that selective advantage is. Indeed, for many of the most frequently mutated genes in cancer, understanding the nature of their selective advantage remains only partially understood despite decades of *in vitro* experiments, mouse models and human studies. Our *in vitro* assays, transcriptomics and metabolomics provide some initial clues as to potential mechanisms, but clearly there are many further lines of investigation to pursue.

To provide further evidence for the claim that these driver mutations provide a selective advantage, we have now analysed the size of clones with driver mutations compared to clones without known drivers. This has shown that *FOXO1* (p=0.005), *CIDEB* (p=0.001) and *ACVR2A* (p=0.001) all generated clones larger than wild-type clones (see figure panel below). **These data have been included in the manuscript ('Properties of clones and patients with drivers', p. 9; Extended Figure 12B).**



Reviewer figure. Box and whisker plots showing the distribution of clone sizes for clones with and without mutations in each of the 4 genes.

3a) They are not relevant for progression to hepatocellular carcinoma (see answer 4 in the somatic mutation section). In fact, these clones appear to be selectively protected from cancer transformation. I do not fully understand their de-enrichment by a factor of 10-50 from HCCs and the authors explanation. Potentially, this would be a somewhat counter-intuitive but nonetheless very attractive anti-cancer pathway to look into.

We agree that the observation that these mutations seemingly do not promote cancer transformation is intriguing. Cancer is, in many ways, a rather specific phenotype for somatic cells, with its defining feature being tissue invasion. It is entirely conceivable that a given driver mutation may confer a selective advantage on a clone, enabling it to proliferate or evade cell death more successfully than its less-fit neighbours, but never promoting the cellular pathways that allow the clone to invade through tissue barriers. The set of genetic changes that enable a hepatocyte to evade the particular selective pressures in a local microenvironment defined by lipotoxicity and caloric excess would not necessarily be the same genetic changes as those promoting clonal expansion in the hypoxic, crowded and disordered architecture of a hepatocellular carcinoma. **We have elaborated these points in the revised Discussion (p. 11).**

3b) They are not relevant for progression of fatty liver disease to cirrhosis (which is one of the key questions in liver research) because firstly the authors did not have access to such tissue (i.e. NASH, ALD) but were restricted to end-stage disease explants in their sample access. Secondly, other groups with access to such samples did not see signatures of the proposed pathomechanisms in their data (see 3a above).

There have been four previous systematic sequencing studies analysing somatic mutations in regenerative nodules from chronic liver disease that we are aware of. As the reviewer states, none of these papers reported driver mutations in the three metabolism genes we discovered, but this is largely due to the different approaches used in the studies –

- In Zhu *et al*[14], published in Cell in 2019, the authors performed exome sequencing of regenerative liver regions from 53 patients, mostly at one sample per patient, identifying 389 somatic mutations overall. Importantly, for 70% of their patients, the cause of their cirrhosis was chronic viral hepatitis (mostly hepatitis C); only 5 patients had alcohol-related liver disease and 1 non-alcoholic steatohepatitis. They then performed a follow-up deep targeted sequencing study of 136 genes on mostly the same samples. *FOXO1*, *CIDEB* and *GPAM* were not included in this targeted gene set.
- In Nault *et al*[15], published in Hepatology in 2014, targeted gene sequencing was performed for genes mutated in HCC in 31 cirrhotic nodules (44% alcohol-related; 7% NAFLD-related; the rest viral). *FOXO1*, *CIDEB* and *GPAM* were not included in this targeted gene set.
- In Torrecilla *et al*[16], published in Journal of Hepatology in 2017, 31 dysplastic nodules (6% alcohol-related; none NAFLD; the rest viral) were sequenced with a gene panel comprising genes mutated in HCC. Again, *FOXO1*, *CIDEB* and *GPAM* were not included in this targeted gene set.
- In Kim *et al*[17], published in Journal of Gastroenterology in 2019, 205 regenerative nodules from 10 patients were sequenced (only 1 alcohol-related; 9 patients had a viral cause). Of these, 10 nodules had exome sequencing performed; the remainder were analysed by deep targeted sequencing on a gene panel that did not include *FOXO1*, *CIDEB* and *GPAM*.

It is clear, then, that we would not have expected these studies to report the driver mutations we have discovered. The extent of sequencing is considerably more limited that the 1590 whole genome sequences we report – indeed, across these studies there are ~70 exomes; the rest of the sequencing was performed on targeted gene panels that did not include the genes we report. Furthermore, the case mix of the published cohorts is strongly biased towards cases of viral hepatitis, which we would not necessarily expect to exhibit the same selective pressures as alcohol-related or non-alcoholic fatty liver disease. Therefore, the driver mutations that emerge in chronic viral hepatitis may be rather different, as indeed suggested in the data from Zhu *et al*[14].

We have now included these explanations for the differences between our data and published studies in the Discussion (p. 11).

4) I think in vitro experiments in hepatoma cell lines can only partially help with the functional exploration. The metabolomics would have been most instructive if it was performed on the human liver samples harbouring the specific mutations. Human liver metabolomics is established (using cryo-microdissection too) but likely not possible with the tissue quality available to the authors. If there is already significant RNA degradation (see 2a), metabolomics seems not to be feasible.

We agree that our *in vitro* experiments will need to be supplemented with studies using both mouse models and from human data. We believe these are beyond the scope of the current manuscript (and the human studies will require further technological development to enable sequencing combined with metabolomics).

We have supplemented the *in vitro* metabolomics with RNA-sequencing of the hepatoma cell lines, which provides further insights into the transcriptional changes associated with *FOXO1* mutations (as described in our response to point 2a above). These new RNA-sequencing data are discussed in the main text ('Mutations impair nuclear export of FOXO1', p. 7), together with a new Extended Figure 10 and Supplementary Tables S6-S7. In addition, we have emphasised the need for these *in vitro* data to be extended with larger clinical studies from human tissues in the Discussion (p. 12).

Referee #3 (Remarks to the Author):

The authors have largely responded to the prior critiques, and that manuscript has been improved considerably, Acaveat that the authors really shoult not characterize what they have studies as "chronic liver disease", but an affect that is applicable for 2 common etiologies of liver disease, ALD and NAFLD. Although they make an argument that these are the most common chronic liver disease, hepatitis B, hepatitis C, as well as other liver diseases are also common enough, even in the West, and not characterized by the current approach.

We thank the reviewer for the endorsement of our manuscript. The point about the focus of our cohort on alcohol-related liver disease and NAFLD is an important one – **we have adapted our description of the cohort throughout the paper to these entities, and included in the Discussion a suggestion that such studies should be replicated in other chronic liver diseases, where we may not expect the genomic landscape to mirror what we see in ARLD and NAFLD (p. 11).**

Referee #4 (Remarks to the Author):

In their manuscript entitled "Convergent somatic mutations in effectors of insulin signalling in chronic liver disease", Ng and colleagues analyze more than 1400 whole genome sequences obtained from laser-capture microdissections of hepatocytes from patients with non-alcoholic fatty liver disease (NAFLD) or alcohol-related liver disease (ARLD). They find clonal expansions harboring a number of recurrently mutated genes, most prominently hotspot mutations in the transcription factor FOXO1, which they explore with additional experiments. Furthermore, they describe a hitherto unknown mutational signature and shorter telomeres associated with clonal expansions.

The manuscript builds upon a dataset published by the same group in 2019, as well as a number of conceptually related papers on somatic evolution of normal tissues that have been coming out in recent years. While I appreciate the effort and cost behind this large-scale tissue collection and sequencing project and find the authors' description of FOXO1 mutations in chronic liver disease very interesting, I cannot help but find that overall, this paper does not quite reach the quality and definitiveness of comparable top publications in the field.

The issue is not technical - I have no concerns about the driver gene identification methods used in this manuscript (see further below for details) – but a problem of narrative. The authors begin with a collection of clinically heterogeneous patients (some have ARLD and other NAFLD). They group them all together to identify mutated genes that drive clonal expansions in liver disease and discover – convincingly – strong selection for a particular hotspot mutation in FOXO1. Here is where, in my opinion, the problems begin. Instead of building a strong narrative around this discovery, the paper fragments into a number of loosely connected storylines that struggle to find a definitive common denominator. There is some very basic investigation of the FOXO1 mutation in cell lines which does not go beyond showing that it interferes with correct subcellular localization of the TF and (unsurprisingly) some resulting changes in metabolite concentration. The vital question if or how these changes contribute to increased hepatocyte survival in the context of liver disease is treated purely speculatively, and only in the discussion. There, the authors speculate that selective pressure is exerted by "toxic lipid species and oxidative stress in hepatocytes processing the chronic dietary excess of calories". At this point, the grouping of ARLD and NAFLD patients becomes a particular problem (as some reviewers have pointed out), as the underlying disease mechanisms are different, and yet both groups have FOXO1 mutations. The authors try to make it all fit together by saying that "similar pathways probably operate in ARLD", but this seems regrettably thin and preliminary, given that these FOXO1 mutants are the main discovery of the paper. The remaining storylines – CIDEB, GPAM, telomere lengths, a new mutational signature – remain largely separate and undeveloped. All in all, I find it difficult to identify a clear conceptual advance, although the FOXO1 mutations are certainly interesting.

We believe that the "clear conceptual advance" and "definitive common denominator" here is the discovery that there are 3 genes involved in metabolism, specifically lipid metabolism, under positive selection in liver tissue affected by NAFLD and ARLD. Critical to the conceptual advance is that several common themes emerge across all 3 genes, suggesting important and unifying insights into the disease process:

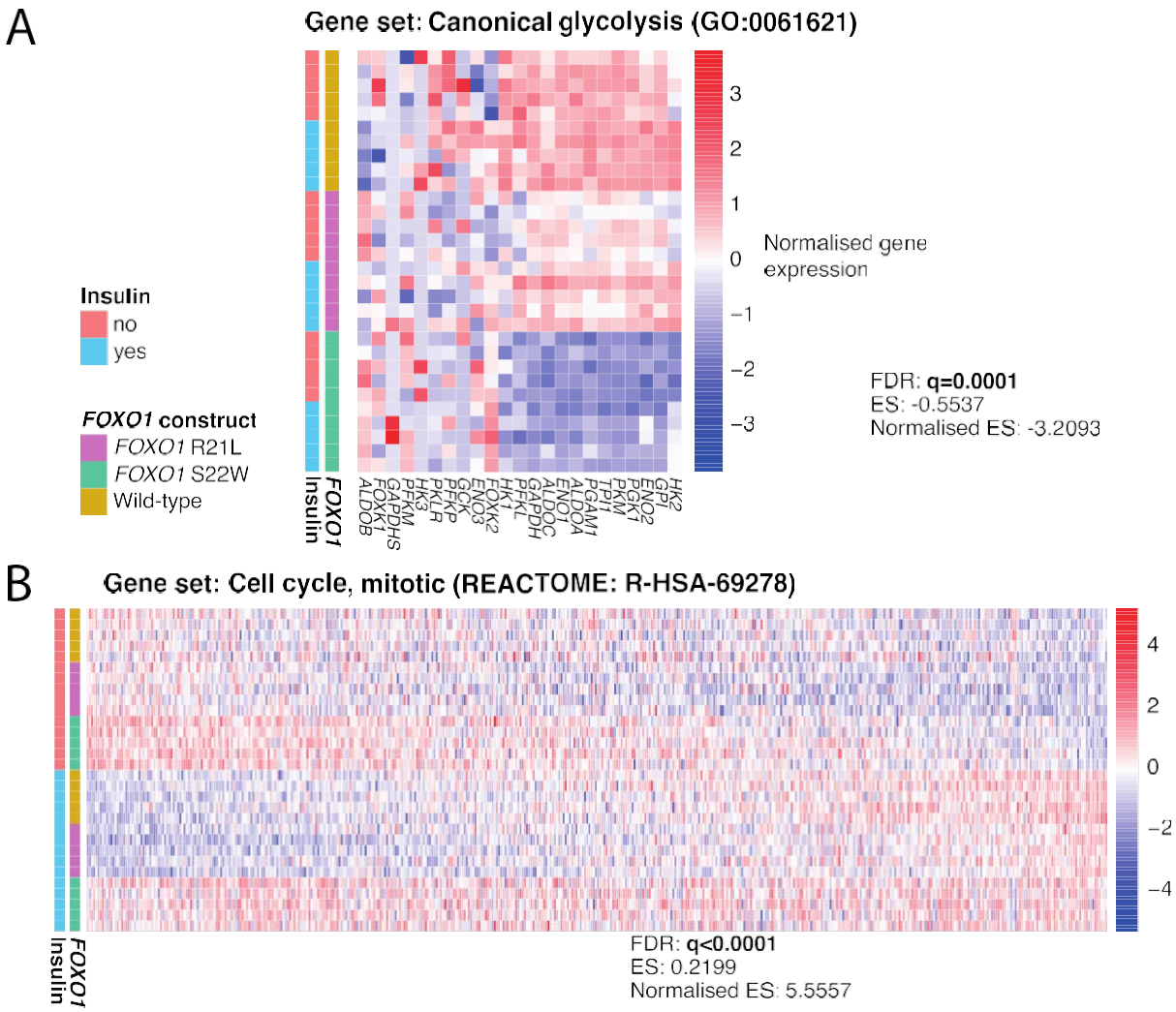
- *Convergent evolution of mutations in the same gene in independent clones was observed for multiple patients across each of the 3 genes* – this suggests that the selection advantage resulting from these mutations is strong, acts across the whole organ and is generalisable to many patients;
- *Driver mutations in these 3 genes are considerably less frequent or absent from both hepatocellular carcinoma and normal liver* – this suggests that the selective advantage from these mutations is context-dependent, not a general property of hepatocytes, thus distinguishing this finding from previous studies in, say, oesophagus, skin or lung[18–21] (where the drivers affect known cancer genes);

- *There is inter-individual variability in which genes are targets of driver mutations* – this suggests that the nature of the selection pressure on hepatocytes differs from patient to patient, since mutations emerge convergently in different genes across the patients in our cohort;
- *These 3 genes are critical regulators of different but interconnected aspects of lipid metabolism in hepatocytes* – *FOXO1* is the master transcription factor downstream of insulin signalling; *CIDEB* is the key regulator of lipid droplet fusion; and *GPAM* encodes the rate-limiting and initiating protein in the conversion of free fatty acids to storage triglycerides.

We accept that the generality of these messages across all 3 metabolism genes was not communicated strongly enough in the previous version of the manuscript. **To address this, we have (1) rewritten the Discussion to draw out the common themes more explicitly (pp. 11-12); (2) included a section on the distributions of clones across patients and clinical correlations ('Properties of clones and patients with drivers', p. 9; Extended Figure 12); and (3) shortened the analyses on mutational signatures and telomere lengths into a single section of the main text ('Other genomic analyses', pp. 10-11) and single Figure 4.**

To address the reviewer's point about the nature of the survival advantage in hepatocytes with *FOXO1* mutations, we have included two additional analyses. First, we analysed the geographical extent of clones with driver mutations compared to clones from the same patients without drivers, showing that the clones with drivers in *FOXO1*, *CIDEB* and *ACVR2A* are indeed significantly larger – **these data are now discussed in the Results ('Properties of clones and patients with drivers', p. 9) and shown in Extended Figure 12B.**

Second, we have performed RNA-sequencing on the cell line models transfected with wild-type, *FOXO1*^{R21L} and *FOXO1*^{S22W} constructs. This reveals that the *FOXO1* mutations do indeed generate co-ordinated alterations in the transcriptome of hepatocytes, with changes affecting both metabolic pathways and pathways involved in clone survival and proliferation. For example, gene sets up-regulated in constructs with *FOXO1* mutations included several implicated in regulation of cell cycle and mitosis ($q<0.0001$); known targets of FOXO1-mediated transcription ($q=0.0084$); and lipid catabolism pathways ($q<0.0001$). Interesting gene sets that were down-regulated with constructs containing *FOXO1* mutations included glycolysis pathways ($q=0.0001$) and apoptosis pathways ($q=0.0004$). Thus, there are transcriptional changes in metabolism pathways seen in these cells that mirror the metabolomic changes we observed. In addition, the *FOXO1* driver mutations are associated with up-regulation of proliferative signals, such as cell cycle/mitosis gene sets, and down-regulation of apoptosis signals – this could potentially contribute to the survival benefit clones with these mutations have in diseased liver. Some illustrative gene sets are shown below. **These new RNA-sequencing data have now been included in the main text ('Mutations impair nuclear export of FOXO1', p. 7), together with a new Extended Figure 10 and Supplementary Tables S6-S7.**



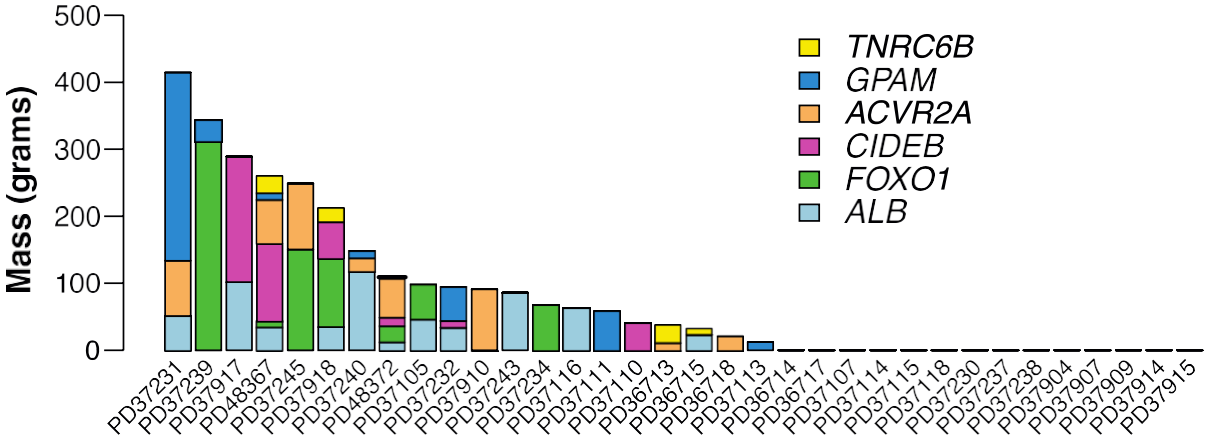
Reviewer figure. RNA-sequencing data from HepG2 cells transduced with either wild-type or mutant *FOXO1-GFP* constructs, and sampled with or without insulin. Each condition was assessed using 5 replicate samples. (A) The *FOXO1* S22W mutation leads to down-regulation (blue) of many genes in the 'Canonical Glycolysis' gene set. (B) The *FOXO1* S22W mutation leads to up-regulation (red) of many genes in the 'Cell cycle, mitotic' gene set.

Another issue that contributed to the "preliminary" impression the paper made on me is the authors' repeated and insistent allusion to a possible role of FOXO1 mutants in metabolic dysfunction at the organismal level. This happens again and again, in the abstract, the introduction, and extensively in the discussion. I am afraid I cannot support even these tentative statements. The data in this paper offer no suggestion whatsoever that the mutant cells contribute to disease at an organismal level. Without showing with a formal calculation how they arrived at this conclusion, the authors claim that mutants "could affect hundreds of grams" of liver, a vague number that is not further justified but nonetheless prominently featured in the abstract. In my opinion, the authors are hurting themselves by alluding to this possibility repeatedly (even mentioning it in the abstract) without evidence or even a formal quantification of mutant burden vis-à-vis total liver

volume across different patients (the latter should be straightforward).

We accept that the hypothesis that these mutations could contribute to metabolic dysfunction at an organismal level remains unproven. **We have therefore removed this speculation from the abstract, introduction and discussion, instead concentrating on discussing the common themes that emerge from our data and their implications for our understanding of metabolic liver disease at the cellular level.**

We have also now generated the formal estimates of how much liver tissue would be accounted for by driver mutations in each of the genes we identify by extrapolating from the biopsies we have sampled, see below – **this is included as Extended Figure 12A.**



Reviewer Figure. Estimates of the total liver mass accounted for by clones carrying driver mutations in each of the 6 significant genes across our patient cohort (based on an average total liver mass of 1500g).

I think some of these issues could potentially be solved, but they would require a pretty extensive restructuring of the paper, which the authors may or may not be willing to undertake. I understand that figuring out whether the mutants are causing any harm is a non-trivial task, but the much more obvious and approachable (and frankly, in the context of this paper, more important) question as to how FOXO1 mutations confer a competitive advantage to hepatocytes in the setting of liver disease really would have to be explored at least a little bit in order for the story to come full circle.

We agree that understanding the detailed metabolic and cellular consequences of these driver

mutations is a non-trivial task – this will require both *in vivo* mouse models of chronic liver disease and detailed genotype-phenotype correlations in large, well-characterised human cohorts, both of which we believe to be beyond the scope of the current paper (and the latter will depend upon further technological development to get to the sample sizes and breadth of characterisation required).

It seems likely that the selective benefit of the drivers we observe derives from the specific pathobiology of ARLD and NAFLD, because the mutations are not seen in normal liver or HCC. The mechanistic basis for hepatocyte death and inflammation in ARLD and NAFLD remains incompletely understood despite several decades of research. Hepatic lipotoxicity represents a common downstream feature of both diseases – in particular, there is strong overlap in which inherited variants predispose to NAFLD and ARLD, with many of the shared GWAS hits directly involved in lipid metabolism pathways[22] (including a recent paper reporting *GPAM* variants linked to both disease processes[23]). A number of lipotoxic lipid species have been implicated in the inflammation and hepatocyte cell death in both diseases, including diacylglycerols, ceramides and lysophosphatidyl choline species[22,24–27], potentially through up-regulation of apoptosis pathways linked to metabolic signalling pathways[28].

Our favoured hypothesis therefore is that it is the metabolic consequences of the driver mutations that result in their selective advantage to hepatocytes. This hypothesis is based on the following lines of reasoning –

- The confluence of somatic driver mutations in three master regulators of lipid processing and storage in these diseases, coupled with the distribution of inherited predisposition variants across these same pathways[22], suggests a common underlying genetic logic;
- Our *in vitro* studies of the *FOXO1* hotspot mutations (cellular localisation, metabolomics and now RNA-sequencing) show that they have profound consequences on lipid and glucose metabolism in hepatocytes;
- Published *in vitro* systematic mutagenesis studies of the CIDE family of genes, which includes *CIDEB*, have shown that altering the charge of key conserved residues, including several of those mutated in our patients, abrogates homodimerization of the proteins, preventing fusion and growth of lipid droplets within the cell[29,30]. Furthermore, *Cideb* knock-out mice show resistance to dietary steatohepatitis and increased insulin sensitivity[31] – the two nonsense and one stop-loss mutations we observe, and presumably also the missense mutations at key conserved residues, would likely have similar consequences.

There are likely to be many complex effects here, however. Our new RNA-sequencing on the *FOXO1* models suggests up-regulation of cell cycle/mitosis genes and down-regulation of apoptosis genes also occur – we do not know whether this is a downstream consequence of the metabolic changes or a pleiotropic effect. Likewise, CIDEB also has a direct role in triggering apoptosis through localisation to mitochondria[32], and the domain responsible for this localisation includes some of the residues mutated in our patients.

We accept the reviewer's point that our previous submission did not adequately address these points – we have therefore removed the speculation on the role of the mutations in causing systemic metabolic dysfunction from the paper, and focused the Discussion on exploring the potential complex selective environment of the liver in NAFLD and ARLD (pp. 11-12).

Some comparatively minor thoughts:

1. I found that a solid overview figure showing all patients, disease conditions, microdissections, detected driver variants and their VAFs etc. was really missing. Such a figure would provide a firm footing and reference for the reader as he/she embarks upon the journey of "vignettes" that make up the majority of this paper. Instead, in figure 1A (which is usually reserved for such an overview), the reader finds himself presented with a very detailed representation of the FOXO1 sequence across species, without having a particularly clear sense of how he got there from the analysis of 1000+ genomes from 34 livers.

This is an excellent suggestion, and **we have now included an overview Figure 1A to summarise the experimental design and cohort.**

2. Somewhat related to point 1 above: One of the prior reviewers asked the authors to provide an analysis correlating genomic features and clinical variables, and the authors replied by pointing to Supplementary Table 1. I think that providing this table does not equate to analysis and agree with the reviewer that it would be important to provide this. Even if the patient numbers are small and it may be hard to achieve statistical significance, this is an important exercise.

We agree that in the previous version of the manuscript, the clinical features of patients with the various driver mutations was not clearly presented. **To address this, we have introduced a new section in the manuscript to outline these analyses, together with the data showing estimates of liver mass affected by drivers across the cohort and comparisons of clone size between clones with and without driver mutations ('Properties of clones and patients with drivers', p. 9; Extended Figure 12; Supplementary Code).**

3. As I mentioned above, I have no doubt about the accuracy of the driver gene identification methods used in this manuscript. They are state-of-the-art methods, and many of them were developed by the authors themselves in a long line of prior work. In particular, I am not at all worried about potential artifacts caused by interdependence between samples from the same patient (a

point raised during review), as the phylogenetic tree structures clearly demonstrate that these mutations are independent somatic events. The authors nicely address the remaining possibility that FOXO1 mutations could be embryonic variants in their supplementary note, although it would be good to present those arguments in a more formal manner. (E.g. "The difficulty with this explanation is that in many of the pairs of microdissections that share a FOXO1 mutation, this is literally the only shared mutation" – please show the data instead of simply stating this).

We have now added panels showing (1) the distribution of VAFs for all pairs of samples carrying the *FOXO1* mutation within an example patient; and (2) the VAFs of that position in the genome across all microdissections for that patient (Supplementary Note 2).

References

1. Aizarani, N. *et al.* A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572**, 199–204 (2019).

2. Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017).

3. Ramachandran, P. *et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level*. *Nature* **575**, (Springer US, 2019).

4. MacParland, S. A. *et al.* Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat. Commun.* **9**, 1–21 (2018).

5. Ceulemans, A. *et al.* RNA-sequencing-based comparative analysis of human hepatic progenitor cells and their niche from alcoholic steatohepatitis livers. *Cell Death Dis.* **8**, (2017).

6. Govaere, O. *et al.* Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. *Sci. Transl. Med.* **12**, 1–18 (2020).

7. Pfister, D. *et al. NASH limits anti-tumour surveillance in immunotherapy-treated HCC*. *Nature* **592**, (2021).

8. Greenman, C., Wooster, R., Futreal, P. A., Stratton, M. R. & Easton, D. F. Statistical analysis of pathogenicity of somatic mutations in cancer. *Genetics* **173**, 2187–2198 (2006).

9. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029–1041 (2017).

10. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–58 (2013).

11. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour

types. *Nature* **505**, 495–501 (2014).

12. Schuster-Böckler, B. & Lehner, B. Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* **488**, 504–507 (2012).

13. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–8 (2013).

14. Zhu, M. *et al.* Somatic Mutations Increase Hepatic Clonal Fitness and Regeneration in Chronic Liver Disease. *Cell* **177**, 608–621 (2019).

15. Nault, J. C. *et al.* Telomerase reverse transcriptase promoter mutation is an early somatic genetic alteration in the transformation of premalignant nodules in hepatocellular carcinoma on cirrhosis. *Hepatology* **60**, 1983–92 (2014).

16. Torrecilla, S. *et al.* Trunk mutational events present minimal intra- and inter-tumoral heterogeneity in hepatocellular carcinoma. *J. Hepatol.* **67**, 1222–1231 (2017).

17. Kim, S. K. *et al.* Comprehensive analysis of genetic aberrations linked to tumorigenesis in regenerative nodules of liver cirrhosis. *J. Gastroenterol.* (2019). doi:10.1007/s00535-019-01555-z

18. Martincorena, I. *et al.* High burden and pervasive positive selection of somatic mutations in normal human skin. *Science (80-.).* **348**, 880–886 (2015).

19. Martincorena, I. *et al.* Somatic mutant clones colonize the human esophagus with age. *Science (80-.).* **917**, 911–917 (2018).

20. Yokoyama, A. *et al.* Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature* **565**, 312–317 (2019).

21. Yoshida, K. *et al.* Tobacco smoking and somatic mutations in human bronchial epithelium. *Nature* **578**, 266–272 (2020).

22. Bianco, C., Casirati, E., Malvestiti, F. & Valenti, L. Genetic predisposition similarities between NASH and ASH: Identification of new therapeutic targets. *JHEP Reports* **3**, 100284 (2021).

23. Hakim, A. *et al.* Genetic variation in the mitochondrial glycerol-3-phosphate acyltransferase is associated with liver injury. *Hepatology* hep.32038 (2021). doi:10.1002/hep.32038

24. Friedman, S. L., Neuschwander-Tetri, B. A., Rinella, M. & Sanyal, A. J. Mechanisms of NAFLD development and therapeutic strategies. *Nat. Med.* **24**, 908–922 (2018).

25. Liu, S. Y., Tsai, I. T. & Hsu, Y. C. Alcohol-related liver disease: Basic mechanisms and clinical perspectives. *Int. J. Mol. Sci.* **22**, (2021).

26. Clugston, R. D. *et al.* Altered hepatic lipid metabolism in C57BL/6 mice fed alcohol: A targeted lipidomic and gene expression study. *J. Lipid Res.* **52**, 2021–2031 (2011).

27. Puri, P. *et al.* A lipidomic analysis of nonalcoholic fatty liver disease. *Hepatology* **46**, 1081–

1090 (2007).

28. Zhao, P. *et al.* An AMPK – caspase-6 axis controls liver damage in nonalcoholic steatohepatitis. *Science (80-.).* **367**, 652–660 (2020).

29. Barneda, D. *et al.* The brown adipocyte protein CIDEA promotes lipid droplet fusion via a phosphatidic acid-binding amphipathic helix. *Elife* **4**, 1–24 (2015).

30. Sun, Z. *et al.* Perilipin1 promotes unilocular lipid droplet formation through the activation of Fsp27 in adipocytes. *Nat. Commun.* **4**, (2013).

31. Li, J. Z. *et al.* Cideb Regulates Diet-Induced Obesity, Liver Steatosis, and Insulin Sensitivity by Controlling Lipogenesis and Fatty Acid Oxidation. *Diabetes* **56**, 2523–2532 (2007).

32. Chen, Z., Guo, K., Shen Yon Toh, Zhou, Z. & Li, P. Mitochondria localization and dimerization are required for CIDE-B to induce apoptosis. *J. Biol. Chem.* **275**, 22619–22622 (2000).

Reviewer Reports on the Second Revision:

Referee #4 (Remarks to the Author):

I thank the authors for removing allusions to a potentially pathogenic role of FOXO1 mutants from the manuscript. The results are now communicated in a more rigorous fashion, and that is good.

The other changes that have been made are welcome but largely cosmetic. RNA sequencing of cell lines that have been transfected with mutant or wild type constructs adds very little to the meat of the story, in my opinion. As the authors point out, FOXO1 mutations are almost exclusively observed in the setting of ARLD/NAFLD. I am not sure how evaluating gene expression in cell lines outside of the context of these diseases answers the original question, namely: What kind of benefit do FOXO1 mutants enjoy in ARLD/NAFLD that they do not enjoy under normal physiological conditions? The authors' observation that mutants display a proliferation signature does not seem to help with the logic here, as a simple proliferation or survival advantage should lead to clonal expansion even under normal conditions or in viral hepatitis (i.e. increased proliferation is not an advantage that is specific to the disease condition of interest). The authors may want to make their argument a bit tighter here.

Overall, I don't see any substantial changes to the paper, although the streamlined text certainly does make a difference. I continue to think that the present narrative is incomplete and thus somewhat inconclusive, but I do not doubt any of the data presented in the paper.

With respect to the response to reviewer 2, my impression remains consistent. The authors make it clear that they believe that the sequenced genomes and their analysis are a sufficient contribution to warrant publication, and that further requests for experiments or mechanism are outside of the scope. I agree with the authors on some of their points, e.g. that replication of the mutation patterns in further patients is not necessary - the discovered FOXO1 mutations are very convincing already. I do not, however, fully understand why the authors say that "methods do not yet exist at the throughput required for such characterization" - surely targeted sequencing of FOXO1 across more samples is very feasible (although I agree not necessary). I agree with reviewer 2 that "in vitro experiments can only partially help with the functional exploration", as I believe that this criticism is related to my main point above, but the authors again conclude that such exploration is outside the scope of this study. It is necessary to point out that other comparable studies have managed to be much more comprehensive, e.g. Zhu et al. (Cell 2019) where careful functional follow-up on genes that were found to be recurrently mutated in the liver was provided via CRISPR screens and tailored mouse models. So it certainly is not impossible.

Minor comments:
Figure 1A is just a cartoon with no primary data. I do wonder whether this is really the best can be done to show the reader an overview of the data this study?

Author Rebuttals to Second Revision:

Referee #4 (Remarks to the Author):

I thank the authors for removing allusions to a potentially pathogenic role of FOXO1 mutants from the manuscript. The results are now communicated in a more rigorous fashion, and that is good.

The other changes that have been made are welcome but largely cosmetic. RNA sequencing of cell lines that have been transfected with mutant or wild type constructs adds very little to the meat of the story, in my opinion. As the authors point out, FOXO1 mutations are almost exclusively observed in the setting of ARLD/NAFLD. I am not sure how evaluating gene expression in cell lines outside of the context of these diseases answers the original question, namely: What kind of benefit do FOXO1 mutants enjoy in ARLD/NAFLD that they do not enjoy under normal physiological conditions? The authors' observation that mutants display a proliferation signature does not seem to help with the logic here, as a simple proliferation or survival advantage should lead to clonal expansion even under normal conditions or in viral hepatitis (i.e. increased proliferation is not an advantage that is specific to the disease condition of interest). The authors may want to make their argument a bit tighter here.

We agree with the reviewer that the increased expression of proliferation genes and down-regulation of pro-apoptotic pathways seen with *FOXO1* S22W mutations is only one component of

adaptation specific to the toxicity experienced by hepatocytes in the diseased microenvironment. It has been our consistent message that a major component of this adaptation is the metabolic consequences of the mutations – that is, the variants in *FOXO1*, *CIDEB* and *GPAM* protect hepatocytes from the lipotoxicity that is a common pathogenetic mechanism in NAFLD and ARLD. The RNA-sequencing data provide evidence for this, since the *FOXO1* S22W mutations also down-regulate canonical glycolysis genes and up-regulate lipid catabolism genes, alongside their effects on proliferation and apoptosis pathways. This is further evidenced by the metabolomics data, which also shows considerable differences in metabolite concentrations between cells transfected with the mutant versus wild-type *FOXO1*.

We accept that this argument was not made explicitly enough in the manuscript, and have revised the Discussion to make the point more clearly (p. 9)

Overall, I don't see any substantial changes to the paper, although the streamlined text certainly does make a difference. I continue to think that the present narrative is incomplete and thus somewhat inconclusive, but I do not doubt any of the data presented in the paper.

With respect to the response to reviewer 2, my impression remains consistent. The authors make it clear that they believe that the sequenced genomes and their analysis are a sufficient contribution to warrant publication, and that further requests for experiments or mechanism are outside of the scope. I agree with the authors on some of their points, e.g. that replication of the mutation patterns in further patients is not necessary - the discovered FOXO1 mutations are very convincing already. I do not, however, fully understand why the authors say that "methods do not yet exist at the throughput required for such characterization" - surely targeted sequencing of FOXO1 across more samples is very feasible (although I agree not necessary). I agree with reviewer 2 that "in vitro experiments can only partially help with the functional exploration", as I believe that this criticism is related to my main point above, but the authors again conclude that such exploration is outside the scope of this study. It is necessary to point out that other comparable studies have managed to be much more comprehensive, e.g. Zhu et al. (Cell 2019) where careful functional follow-up on genes that were found to be recurrently mutated in the liver was provided via CRISPR screens and tailored mouse models. So it certainly is not impossible.

We fully accept that *in vitro* experiments "only partially help with the functional exploration." We would emphasise, though, that mouse knockouts for *Cideb* and *Gpam* already exist and show phenotypes that strongly support our argument that the selective benefit of mutations in these genes accrues from their effects on lipid metabolism.

The mutations we observe in *CIDEB* and *GPAM* are consistent with loss-of-function effects. First, we see an excess of protein-truncating mutations (nonsense, splice site and frameshift indels) in both genes; second, the missense mutations are distributed across many residues with high levels of

evolutionary conservation (but are not hotspot variants); third, several of the residues with missense mutations in *CIDEB* have been already studied in *in vitro* mutagenesis screens, and shown to eliminate normal protein function[1,2]. Thus, mouse models in which these genes are knocked out would be expected to mimic the *in vivo* effects of the mutations we observe. The phenotypes of the knockout mice are revealing –

- *Cideb* knockout mice had lower levels of triglycerides and free fatty acids compared to controls[3]. Importantly, when fed a high-fat diet that generated fatty liver disease in the control mice, the *Cideb*-/- mice were resistant to dietary steatohepatitis. Furthermore, the knockout mice exhibited increased insulin sensitivity, a difference that was especially marked when challenged with a high-fat diet[3].
- *Gpam* knockout mice also showed a lower triglyceride content in liver than wild-type controls, with altered composition of the lipids that were present[4,5]. When challenged with a high-fat, high-sucrose diet, the knockout mice accumulated less hepatic triglyceride and had smaller lipid vacuoles than wild-type mice[6].

In conclusion, then, the mouse knockouts of *Cideb* and *Gpam* show reduced accumulation of hepatic triglyceride than wild-type controls, especially with a high-fat diet, and therefore less evidence of steatohepatitis.

We accept that the data from the existing mouse models were not extensively discussed in the manuscript, and have now provided more detail in the Discussion (p. 9).

Minor comments:

Figure 1A is just a cartoon with no primary data. I do wonder whether this is really the best can be done to show the reader an overview of the data this study?

We have revised Figure 1A to provide a quick-reference summary of the main findings on the distribution of mutations in the positively selected genes.

References

1. Barneda, D. *et al.* The brown adipocyte protein CIDEA promotes lipid droplet fusion via a phosphatidic acid-binding amphipathic helix. *Elife* **4**, 1–24 (2015).

2. Sun, Z. *et al.* Perilipin1 promotes unilocular lipid droplet formation through the activation of Fsp27 in adipocytes. *Nat. Commun.* **4**, (2013).

3. Li, J. Z. *et al.* Cideb Regulates Diet-Induced Obesity, Liver Steatosis, and Insulin Sensitivity by Controlling Lipogenesis and Fatty Acid Oxidation. *Diabetes* **56**, 2523–2532 (2007).

4. Hammond, L. E. *et al.* Mitochondrial Glycerol-3-Phosphate Acyltransferase-Deficient Mice Have Reduced Weight and Liver Triacylglycerol Content and Altered Glycerolipid Fatty Acid Composition. *Mol. Cell. Biol.* **22**, 8204–8214 (2002).

5. Yazdi, M. *et al.* The role of mitochondrial glycerol-3-phosphate acyltransferase-1 in regulating lipid and glucose homeostasis in high-fat diet fed mice. *Biochem. Biophys. Res. Commun.* **369**, 1065–1070 (2008).

6. Hammond, L. E. *et al.* Mitochondrial glycerol-3-phosphate acyltransferase-1 is essential in liver for the metabolism of excess acyl-CoAs. *J. Biol. Chem.* **280**, 25629–25636 (2005).