A genome-wide screen in the mouse liver reveals sex-specific and cell non-autonomous regulation of cell fitness

Heather R. Keys[1]† and Kristin A. Knouse[1*]†

[1]Whitehead Institute for Biomedical Research, Cambridge, MA

*Corresponding author. Email: knouse@wi.mit.edu

†These authors contributed equally to this work

ABSTRACT

Our ability to understand and modulate mammalian physiology and disease requires knowing how all genes contribute to any given phenotype in the organism. Genome-wide screening using CRISPR-Cas9 has emerged as a powerful method for the genetic dissection of cellular processes[1,2], but the need to stably deliver single guide RNAs to millions of cells has restricted its implementation to *ex vivo* systems. These *ex vivo* systems cannot reproduce all of the cellular phenotypes observed *in vivo* nor can they recapitulate all of the factors that influence these phenotypes. There thus remains a pressing need for high-throughput functional genomics in a living organism. Here, we establish accessible genome-wide screening in the mouse liver and use this approach to uncover the complete regulation of cellular fitness in a living organism. We discover novel sex-specific and cell non-autonomous regulation of cell growth and viability. In particular, we find that the class I major histocompatibility complex is essential for preventing immune-mediated clearance of hepatocytes. Our approach provides the first comprehensive picture of cell fitness in a living organism and highlights the importance of investigating cellular phenomena in their native context. Our screening method is robust, scalable, and easily adapted to examine diverse cellular processes using any CRISPR application. We have hereby established a foundation for high-throughput functional genomics in a living mammal, enabling unprecedented insight into mammalian physiology and disease.

INTRODUCTION

A complete understanding of the genetic determinants underlying cell fitness and function in mammalian tissues is limited by the availability of tools for high-throughput genetic dissection in living organisms. Genome-wide screening using CRISPR-Cas9 has emerged as a powerful approach for the comprehensive investigation of cellular phenotypes[1,2], but it requires stable delivery of a single guide RNA (sgRNA) library to tens of millions of cells at a low multiplicity of infection. Transposons and adeno-associated virus (AAV) can deliver sgRNAs *in vivo* with reasonable efficiency[3,4], but transposons often undergo multiple integrations per genome and AAV vectors do not integrate at all, limiting the utility of either of these approaches for genome-wide screening. Lentivirus is a preferable sgRNA delivery method as it allows for stable, single-copy integration in target cells. Unfortunately, genome-scale lentiviral delivery has only been achievable in cells *ex* vivo, thereby restricting genome-wide screening to cell culture systems or cellular transplantation models[5,6]. These *ex vivo* systems cannot always reproduce the cellular phenotypes desired for study and, even when they can, they cannot recapitulate the entirety of extracellular factors that influence these phenotypes *in vivo*. These limitations emphasize the need to bring high-throughput functional genomics into the organism. To satisfy this demand and achieve a comprehensive understanding of cellular fitness in its native context, we sought to enable effective genome-scale sgRNA delivery *in vivo* and establish accessible genome-wide CRISPR screening in a living mammal.

RESULTS

Genome-scale sgRNA delivery to the liver

54 The mouse liver, comprised of tens of millions of hepatocytes, is a promising
55 tissue for genome-wide screening because it provides cell numbers compatible with
56 genome-scale screening in a single mouse. Moreover, given the liver's diverse
57 metabolic functions and impressive regenerative capacity, hepatocytes exhibit a broad
58 range of phenotypes—from universal cellular processes to hepatocyte-specific
59 phenomena—that are ripe for genetic dissection. Previous attempts at delivering
60 lentivirus to liver have suffered from poor transduction efficiency and immune-mediated
61 clearance of transduced hepatocytes[7,8]. However, we hypothesized that intravenously
62 injecting highly concentrated lentivirus into neonatal mice might avoid these pitfalls and
63 achieve efficient, stable transduction[9]. To test this, we generated lentiviruses encoding
64 a non-targeting sgRNA alongside mCherry or mTurq2 reporters and injected varying
65 doses of an equal mixture of these two lentiviruses into postnatal day (PD) one mice
66 (Fig. 1a). We observed a dose-dependent increase in the percentage of transduced
67 hepatocytes, with a dose of 5×10^7 transduction units (TU) transducing over 75% of
68 hepatocytes (Fig. 1b,c). Using histologic measurements and the transduction frequency
69 of both fluorescent reporters, we estimated that a dose of 5×10^7 TU transduces
70 approximately 10 million hepatocytes per PD1 liver with an average of just two
71 integration events per cell (Fig. 1c and Extended Data Fig. 1a). This transduction
72 efficiency would afford >200-fold coverage of a 100,000 feature sgRNA library in a
73 single mouse. Importantly, these transduced hepatocytes were distributed uniformly
74 throughout the liver lobule and persisted into adulthood (Fig. 1b and Extended Data Fig.
75 1b,c). Our lentiviral approach therefore establishes an sgRNA delivery method that is
76 wholly compatible with genome-scale screening in the mouse liver.

77 We next asked whether we could use this sgRNA delivery approach as the basis
78 for temporally-controlled protein depletion in hepatocytes. We took advantage of
79 commercially available loxP-stop-loxP-Cas9 (LSL-Cas9) mice in which we could induce
80 Cas9 in nearly all hepatocytes by injecting an adeno-associated virus expressing Cre
81 recombinase from the hepatocyte-specific *Tbg* promoter[10] (AAV-Cre, Extended Data
82 Fig. 1d,e). To evaluate the efficiency and kinetics of protein depletion, we selected two
83 long-lived, non-essential proteins: the mitochondrial enzyme MAO-B (encoded by
84 *Maob*) and the nuclear lamin Lamin B2 (encoded by *Lmnb2*)[11,12]. After delivering
85 sgMaob-mCherry or sgLmnb2-mCherry lentivirus to PD1 mice, we injected PBS or
86 AAV-Cre at PD5 and harvested livers at various time points to evaluate protein levels in
87 individual hepatocytes (Fig. 1d). By two weeks after Cas9 induction, MAO-B and Lamin
88 B2 were depleted exclusively in mCherry-positive hepatocytes in mice injected with
89 AAV-Cre (Fig. 1e,f, and Extended Data Fig. 1f,g). Importantly, this combination of
90 lentiviral-mediated sgRNA delivery, AAV-Cre-mediated induction of Cas9, and resulting
91 gene targeting did not induce detectable liver inflammation nor did it affect hepatocyte
92 turnover (Extended Data Fig. 1h-k). This approach therefore offers an effective platform
93 for hepatocyte-specific protein depletion and genetic screening at any point in the
94 animal's lifetime without obvious collateral perturbation of hepatocyte fitness.
95

96 **A genome-wide screen in the liver**

97 To enable genome-wide screening for diverse hepatocyte phenotypes, we
98 generated an sgRNA library targeting all genes expressed in the developing, quiescent,
99 and regenerating mouse liver. We performed RNA sequencing on livers at various time

100 points during mouse development and after liver injury and determined that 13,266
101 protein-coding genes were expressed (FPKM > 0.3) at one or more time points (Fig. 2a,
102 Extended Data Fig. 2a,b, Supplementary Table 1). We generated an sgRNA library
103 targeting 13,189 of these genes (average of 5 sgRNAs per gene) alongside a previously
104 published set of 6,500 control sgRNAs (~2,000 non-targeting sgRNAs and ~4,500
105 sgRNAs targeting exonic and intronic regions of control genes) for a total of 71,878
106 unique sgRNAs[13] (Extended Data Fig. 2c, Supplementary Table 2).

107 With this method in hand, we undertook a genome-wide screen for hepatocyte
108 fitness (Fig. 2b). To screen for the ability of hepatocytes to both persist and proliferate,
109 we elected to screen over a three-week period in neonatal development when
110 hepatocytes undergo approximately three population doublings to increase liver mass.
111 We injected 5 x 10^7 TU of our lentiviral library into four female and four male LSL-Cas9
112 mice at PD1. At PD5, we harvested livers from two males and two females to evaluate
113 the initial library representation. The sgRNA representation in these four livers
114 correlated extremely well (Pearson r = 0.97) with the plasmid library and we detected all
115 sgRNAs, indicating that we can effectively deliver and recover a genome-scale sgRNA
116 library from the neonatal mouse liver (Fig. 2c, Supplementary Table 3). In the remaining
117 mice, we induced Cas9 at PD5 and harvested their livers at PD26 to evaluate the final
118 library representation. We used the Model-based Analysis of Genome-wide
119 CRISPR/Cas9 Knockout (MAGeCK) algorithm to identify enriched and depleted genes
120 based on statistical differences in their change in sgRNA abundance at PD26 relative to
121 PD5[14] (Supplementary Table 3). Using a false discovery rate cutoff of 0.05, we identified
122 0-6 significantly enriched and 40-386 significantly depleted genes in individual mice,

123 indicating that our method can detect enriched and depleted genes in a single mouse.
124 We also generated gene-level scores by calculating the median log2 fold change in
125 abundance of sgRNAs targeting a given gene and observed a strong correlation across
126 the four mice (Pearson r = 0.46 to 0.75, Fig. 2d, Extended Data Fig. 2d, Supplementary
127 Table 3). To improve our power to detect enriched and depleted genes, we combined
128 the data from all mice and tested whether the gene-level scores for each gene deviated
129 significantly from those for all genes across the four mice. Using false discovery rate
130 cutoffs of 0.05 and 0.25, we identified, respectively, 30 and 658 significantly enriched
131 genes and 661 and 1,482 significantly depleted genes across all mice (Fig. 2e,
132 Supplementary Table 3). Finally, we calculated a unified gene score representing the
133 median log2 fold change for each gene across all mice (Supplementary Table 3). These
134 gene scores were not biased by gene expression level or protein half-life, reaffirming
135 that long-lived proteins were effectively depleted[15] (Extended Data Fig. 2e,f). Together,
136 these data establish the technical feasibility of genome-wide screening in the mouse
137 liver. Importantly, while screening multiple mice in parallel increases the power to
138 discover significant hits, a single mouse is sufficient to identify significantly enriched and
139 depleted genes.

140 We next asked whether our screen reliably uncovered regulation of cell fitness.
141 We first confirmed that genes established to be essential in cell culture were highly
142 represented among the depleted genes across mice[16] (Fig. 2f). To evaluate whether our
143 screen could reveal regulation specific to hepatocyte fitness in the liver, we assessed
144 the gene scores for two sets of genes known to affect hepatocyte fitness *in vivo*: (1) a
145 set of 13 genes established as tumor suppressors in hepatocellular carcinoma

146 (expected to enrich) and (2) a set of seven genes required for hepatocyte viability
147 (expected to deplete)[17] (Supplementary Table 3). Among the tumor suppressor genes,
148 eight of the 13 genes were significantly enriched (FDR < 0.25, Fig. 2g). Among the
149 genes required for hepatocyte viability, all seven genes were significantly depleted
150 (FDR < 0.25, Fig. 2g). These results support our screen as a reliable platform for
151 uncovering genetic regulation of hepatocyte fitness in the liver.
152

153 **Sex-specific regulation of hepatocyte fitness**

154 Screening primary cells in their native context has the unique power to capture all
155 organismal regulation of cell fitness, something that cannot be achieved by screening
156 cell lines in culture. One such advantage is the ability to investigate the influence of
157 biological sex in an otherwise isogenic background. To determine whether genes can
158 shape hepatocyte fitness in a sex-dependent manner, we compared the gene scores for
159 all genes in males versus females. We identified three X-linked genes and nine
160 autosomal genes with sex-specific effects on fitness (Fig. 3a). Two X-linked genes,
161 *Ddx3x* and *Eif2s3x*, both involved in protein synthesis, exhibited the greatest differences
162 between the sexes and were exclusively essential in females. Both of these genes
163 escape X inactivation and have paralogs on the Y chromosome, *Ddx3y* and *Eif2s3y*,
164 with similar function[18]. Thus, it is possible that disruption of *Ddx3x* and *Eif2s3x* causes a
165 fitness defect in female hepatocytes while male hepatocytes are functionally
166 complemented by the Y chromosome paralogs. Among the remaining genes with sex-
167 specific effects, we were surprised to identify an uncharacterized open reading frame,
168 1810030O07Rik, also known as *Cxorf38*, that when disrupted conferred a fitness

169 advantage exclusively in females. The *Cxorf38* gene is also encoded on the X
170 chromosome and has been shown to escape X inactivation but does not have a known
171 Y chromosome paralog[19]. It is expressed at low levels across tissues (FPKM ranging
172 from 1-5), with a trend toward higher expression in embryonic versus adult tissues, and
173 is predicted to encode a 320 amino acid protein comprised of a domain of unknown
174 function[20]. To confirm that *Cxorf38* targeting indeed provides a female-specific fitness
175 advantage, we performed *in vivo* competition assays to evaluate the fitness of
176 hepatocytes depleted of Cxorf38 relative to control hepatocytes (Extended Data Fig. 3,
177 Fig. 3b, top panel). Consistent with our screen results, we observed a significant
178 expansion of Cxorf38-depleted hepatocytes relative to control hepatocytes exclusively
179 in female mice (Fig. 3b, bottom panel). We note that cancer genomics efforts have not
180 predicted *Cxorf38* to function as a tumor suppressor gene, suggesting that its function
181 might be limited to embryonic and neonatal development[21,22]. Our method thus provides
182 a unique opportunity to investigate how biological sex shapes cellular phenotypes and
183 reveals twelve genes that regulate hepatocyte fitness in a sex-specific manner.
184

185 **Negative and positive regulation of hepatocyte fitness**

186 Having established that the majority of genes influence hepatocyte fitness in a
187 sex-independent manner, we next sought to understand broad regulation of hepatocyte
188 fitness by performing gene set enrichment analysis using the unified gene scores
189 across the four mice. No gene sets were identified to be enriched in our screen.
190 However, we noted that over 50% of the top 25 most enriched genes in our screen have
191 been established to act as tumor suppressor genes in at least one context[21,22]

192 (Supplementary Table 3). Indeed, a set of the top 50 computationally predicted pan-
193 cancer tumor suppressor genes was significantly enriched in our screen[21], something
194 not typically observed in cell culture screens including those in mouse embryonic stem
195 cells or human hepatocellular carcinoma cell lines[23-26] (Fig. 3c,d, Supplementary Table
196 3). One possible explanation for this difference is that cell lines naturally harbor and/or
197 accumulate inactivating mutations in tumor suppressor genes, thereby compromising
198 identification of these genes in subsequent screening[27]. Our screen's unique ability to,
199 within only a few population doublings, reliably recover tumor suppressors highlights a
200 key advantage of genome-wide screening in primary tissue—the ability to utilize wild-
201 type cells and thereby identify all genetic perturbations that alter a chosen phenotype.
202 We next turned our focus to genes required for hepatocyte fitness. We identified
203 several gene sets that were significantly depleted in our screen (Fig. 4a, Supplementary
204 Table 4). The gene sets included those previously established as essential for fitness in
205 cell culture, including ribosome, proteasome, spliceosome, and RNA
206 polymerase[1,2,12,16,25,26]. However, we also identified several other gene sets not
207 documented to be essential for cells in culture, including N-glycan biosynthesis,
208 glycosaminoglycan biosynthesis/heparan sulfate, and antigen processing/presentation.
209 Notably, these pathways all play major roles in the presentation or secretion of proteins
210 at the cell surface[28,29]. Our screen thus uncovers requirements for cellular fitness
211 shared by cells in culture but indicates additional, possibly cell non-autonomous,
212 requirements for cell fitness in the organismal context.
213
214 **Cell non-autonomous control of hepatocyte fitness**

215 To identify factors uniquely important for cell fitness in the organismal context, we
216 compared our screen to screens of mouse embryonic stem (ES) cells in culture and
217 human hepatocellular carcinoma (HCC) cell lines in culture[23-26]. We specifically chose
218 ES cells and HCC cell lines in an effort to control for any species-specific and cell
219 lineage-specific requirements for cell fitness. We identified four gene sets that were
220 preferentially depleted in our screen relative to the cell culture screens: protein export,
221 SNARE interactions in vesicular transport, antigen processing/presentation, and
222 glycosaminoglycan biosynthesis/heparan sulfate (Fig. 4b, Supplementary Table 4). For
223 the protein export and SNARE interactions gene sets, the most depleted genes in each
224 set were depleted across all screens but to a greater extent in our screen (Extended
225 Data Fig. 4a,b). However, for the glycosaminoglycan biosynthesis/heparan sulfate and
226 antigen processing/presentation gene sets, some genes were depleted exclusively in
227 our screen (Fig. 4c,d). Within the glycosaminoglycan biosynthesis/heparan sulfate gene
228 set, *Hs2st1* and *Ndst1* were uniquely depleted in our screen (Fig. 4c). These two genes
229 encode enzymes involved in the biosynthesis of heparan sulfate, a glycosaminoglycan
230 typically conjugated to plasma membrane or extracellular matrix proteins. Heparan
231 sulfate interacts with a variety of extracellular proteins and thereby modulates both
232 mechanotransduction and signal transduction at the cell surface[30]. Within the antigen
233 processing/presentation gene set, *Tap1* and *B2m* were uniquely and dramatically
234 depleted in our screen, ranking as the 41st and 320th most depleted genes, respectively
235 (Fig. 4d). Both of these genes are involved in—and required for—presentation of
236 antigens at the cell surface by the class I major histocompatibility complex (MHC)
237 pathway[29]. Indeed, within the antigen processing/presentation gene set, eight of the 32

genes attributed to the class I MHC pathway were depleted in our screen whereas none of the 13 genes attributed to the class II MHC pathway exhibited depletion (FDR < 0.25, Extended Data Fig. 4c). Our screen thus uncovered two novel regulators of cell fitness—heparan sulfate and class I MHC—that are uniquely essential for cells in the organism and likely operate through interactions with the extracellular matrix and other cells.

The class I MHC pathway presents intracellular antigens at the cell surface. In this pathway, cytoplasmic proteins are degraded by the proteasome, resulting peptides are transported into the endoplasmic reticulum (ER) via a Tap1/Tap2 heterodimer, and these peptides are then loaded onto the class I MHC complex comprised of a heavy chain (H2) and β2-microglobulin (B2m) (Fig. 4e). At the cell surface, class I MHC can interact with both cytotoxic CD8 T cells and natural killer (NK) cells. The latter interaction can provide a pro-survival role by preventing NK cell cytotoxicity. However, loss of class I MHC alone should not be sufficient to induce NK cell cytotoxicity. Classically, NK cell activation requires both a loss of inhibitory signals, via loss of class I MHC, and presence of activating signals, expressed on the surface of transformed or infected cells[31]. Although our screening approach involves viral infection of hepatocytes, any inflammation resulting from lentiviral and AAV vectors has been shown to resolve within 72 hours and, consistently, we did not observe any inflammation one week after the combination of lentiviral and AAV-Cre infection[32,33] (Extended Data Fig. 1h,i). It was therefore surprising that loss of class I MHC alone would have such a dramatic impact on hepatocyte fitness. To validate this finding, we performed an *in vivo* competition between cells depleted of Tap1 and control cells (Extended Data Fig. 4d, Fig. 4f, top

261 panel). Consistent with our screen, hepatocytes depleted of Tap1 were lost over time
262 (Fig. 4f, bottom panel). To test whether this fitness defect occurs via NK cell-mediated
263 cytotoxicity, we repeated the competition in the setting of NK cell depletion and
264 observed reduced loss of Tap1-deficient cells (Extended Data Fig. 4e, Fig. 4f, bottom
265 panel). Our results confirm an essential role for class I MHC in hepatocyte viability that
266 acts cell non-autonomously via NK cells. That loss of class I MHC alone is sufficient to
267 induce hepatocyte clearance implies far more rigorous immune surveillance of somatic
268 cells than previously appreciated.

269

270 **DISCUSSION**

271 Herein, we successfully performed, to our knowledge, the first genome-wide
272 CRISPR-Cas9 screen for cell fitness in a living organism. In addition to identifying
273 requirements shared by cells in culture, we also uncovered sex-specific and cell non-
274 autonomous regulation of cell fitness. In particular, we have revealed an essential role
275 for class I MHC in hepatocyte viability. We note that expression of class I MHC at the
276 cell surface is exquisitely sensitive to cellular perturbations and that ER stress of various
277 etiologies can impair its expression[34,35]. Moreover, obesity and metabolic syndrome are
278 known to induce ER stress in hepatocytes[36]. We speculate that perhaps the rigorous
279 requirement for class I MHC in hepatocytes serves to identify and eliminate stressed
280 cells from the organ and that this surveillance may contribute to the inflammation
281 associated with metabolic syndrome. It will be interesting to determine if this

requirement extends to other somatic cell types and whether this serves as a stringent quality control mechanism for cells in an organism.

Our screen's ability to uncover regulation of cell fitness not previously identified in cell culture screens emphasizes the necessity and power of genome-wide screening in the living organism. Importantly, our lentiviral delivery approach coupled with inducible Cas9 can readily be applied to screen diverse hepatocyte phenotypes—from fundamental processes universal to all cells to specialized metabolic and regenerative phenomena unique to hepatocytes—and minimally requires just one mouse. Our method can also be adapted to other CRISPR-based approaches including CRISPR interference/activation and Perturb-Seq[37]. More broadly, our system establishes the feasibility of genome-wide screening in a living organism and inspires efforts to bring this technology to other organs. Collectively, these diverse applications will bring the experimental tractability once restricted to cell culture to the living organism, enabling unprecedented insight into mammalian physiology and disease.

REFERENCES

1. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343,** 84–87 (2014).
2. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343,** 80–84 (2014).
3. Xu, C. *et al.* piggyBac mediates efficient in vivo CRISPR library screening for tumorigenesis in mice. *Proc. Natl. Acad. Sci. U.S.A.* **114,** 722–727 (2017).

304 4. Wang, G. *et al.* Mapping a functional cancer genome atlas of tumor suppressors
305 in mouse liver using AAV-CRISPR-mediated direct in vivo screening. *Sci Adv* **4,**
306 eaao5508 (2018).
307 5. Chen, S. *et al.* Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth
308 and Metastasis. *Cell* **160,** 1246–1260 (2015).
309 6. Dong, M. B. *et al.* Systematic Immunotherapy Target Discovery Using Genome-
310 Scale In Vivo CRISPR Screens in CD8 T Cells. *Cell* **178,** 1189–1204.e23 (2019).
311 7. Pfeifer, A. *et al.* Transduction of liver cells by lentiviral vectors: analysis in living
312 animals by fluorescence imaging. *Molecular Therapy* **3,** 319–322 (2001).
313 8. Waddington, S. N. *et al.* Long-term transgene expression by administration of a
314 lentivirus-based vector to the fetal circulation of immuno-competent mice. *Gene*
315 *Ther.* **10,** 1234–1240 (2003).
316 9. Nguyen, T. H. *et al.* Therapeutic lentivirus-mediated neonatal in vivo gene therapy
317 in hyperbilirubinemic Gunn rats. *Molecular Therapy* **12,** 852–859 (2005).
318 10. Platt, R. J. *et al.* CRISPR-Cas9 knockin mice for genome editing and cancer
319 modeling. *Cell* **159,** 440–455 (2014).
320 11. Mathieson, T. *et al.* Systematic analysis of protein turnover in primary cells. *Nat*
321 *Commun* **9,** 689 (2018).
322 12. Wang, T. *et al.* Identification and characterization of essential genes in the human
323 genome. *Science* **350,** 1096–1101 (2015).
324 13. Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9-
325 mediated gene inactivation. *Nat Biotech* **32,** 1262–1267 (2014).

326 14. Li, W. *et al.* MAGeCK enables robust identification of essential genes from
327 genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* **15,** 554 (2014).
328 15. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression
329 control. *Nature* **473,** 337–342 (2011).
330 16. Hart, T. *et al.* Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout
331 Screens. *G3 (Bethesda)* **7,** 2719–2727 (2017).
332 17. Cancer Genome Atlas Research Network. Electronic address: wheeler@bcm.edu,
333 Cancer Genome Atlas Research Network & Saksena, G. Comprehensive and
334 Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* **169,**
335 1327–1341.e23 (2017).
336 18. Berletch, J. B. *et al.* Escape from X Inactivation Varies in Mouse Tissues. *PLoS*
337 *Genet* **11,** e1005079–26 (2015).
338 19. Zhang, Y. *et al.* Transcriptional profiling of human liver identifies sex-biased genes
339 associated with polygenic dyslipidemia and coronary artery disease. *PLoS ONE*
340 **6,** e23506 (2011).
341 20. Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse
342 genome. *Nature* **515,** 355–364 (2014).
343 21. Davoli, T. *et al.* Cumulative Haploinsufficiency and Triplosensitivity Drive
344 Aneuploidy Patterns and Shape the Cancer Genome. *Cell* **155,** 948–962 (2013).
345 22. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer.
346 *Nucleic Acids Research* **47,** D941–D947 (2019).

347 23. Shohat, S. & Shifman, S. Genes essential for embryonic stem cells are
348 associated with neurodevelopmental disorders. *Genome Research* **29,** 1910–
349 1918 (2019).
350 24. Tzelepis, K. *et al.* A CRISPR Dropout Screen Identifies Genetic Vulnerabilities
351 and Therapeutic Targets in Acute Myeloid Leukemia. *CellReports* **17,** 1193–1205
352 (2016).
353 25. Meyers, R. M. *et al.* Computational correction of copy number effect improves
354 specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nature Genetics*
355 **49,** 1779–1784 (2017).
356 26. Dempster, J. M. *et al.* Extracting Biological Insights from the Project Achilles
357 Genome-Scale CRISPR Screens in Cancer Cell Lines. *bioRxiv* **20,** 21–35 (2020).
358 27. Merkle, F. T. *et al.* Human pluripotent stem cells recurrently acquire and expand
359 dominant negative P53 mutations. *Nature* **545,** 229–233 (2017).
360 28. Cherepanova, N., Shrimal, S. & Gilmore, R. N-linked glycosylation and
361 homeostasis of the endoplasmic reticulum. *Curr. Opin. Cell Biol.* **41,** 57–65
362 (2016).
363 29. Peaper, D. R. & Cresswell, P. Regulation of MHC class I assembly and peptide
364 binding. *Annu. Rev. Cell Dev. Biol.* **24,** 343–368 (2008).
365 30. Ravikumar, M., Smith, R. A. A., Nurcombe, V. & Cool, S. M. Heparan Sulfate
366 Proteoglycans: Key Mediators of Stem Cell Function. *Front Cell Dev Biol* **8,**
367 581213 (2020).
368 31. Chan, C. J., Smyth, M. J. & Martinet, L. Molecular mechanisms of natural killer
369 cell activation in response to cellular stress. *Cell Death Differ* **21,** 5–14 (2014).

370 32. Brown, B. D. *et al.* In vivo administration of lentiviral vectors triggers a type I
371 interferon response that restricts hepatocyte gene transfer and promotes vector
372 clearance. *Blood* **109,** 2797–2805 (2007).
373 33. Arun Srivastava, R. W. H. Innate immune responses to AAV vectors. 1–10 (2011).
374 doi:10.3389/fmicb.2011.00194/abstract
375 34. de Almeida, S. F., Fleming, J. V., Azevedo, J. E., Carmo-Fonseca, M. & de
376 Sousa, M. Stimulation of an unfolded protein response impairs MHC class I
377 expression. *The Journal of Immunology* **178,** 3612–3619 (2007).
378 35. Granados, D. P. *et al.* ER stress affects processing of MHC class I-associated
379 peptides. *BMC Immunol* **10,** 10–15 (2009).
380 36. Ogawa, Y. *et al.* Palmitate-induced lipotoxicity is crucial for the pathogenesis of
381 nonalcoholic fatty liver disease in cooperation with gut-derived endotoxin. *Sci.*
382 *Rep.* 1–14 (2018). doi:10.1038/s41598-018-29735-6
383 37. Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR-Cas
384 technologies and applications. *Nat Rev Mol Cell Biol* **20,** 490–507 (2019).
385

386 **METHODS**

387 **Animals**

388 C57BL/6J mice (strain 000664) and LSL-Cas9 mice (strain 026175) were purchased
389 from the Jackson Laboratory. Mice were either singly- or group-housed with a 12-hour
390 light-dark cycle (light from 7 AM to 7 PM, dark from 7 PM to 7 AM) in a specific-
391 pathogen-free animal facility with unlimited access to food and water. To deliver
392 lentivirus, up to 100 μL of lentivirus in PBS was injected into the temporal vein of

393 postnatal day one mice. For protein depletion tests, mice were injected with 1.25 x 10^7
394 transduction units (TU) of sgRNA-mCherry lentivirus. For the screen, mice were injected
395 with 5 x 10^7 TU of sgRNA-mCherry lentiviral library. For validating *Cxorf38* and *Tap1*,
396 mice were injected with 1 x 10^7 TU of an equal mixture of sgAAVS1-mTurq2 and
397 sgCxorf38-mCherry or sgTap1-mCherry lentiviruses. To deliver AAV-Cre, a stock
398 solution of AAV8-TBG-Cre (Addgene 107787-AAV8) was diluted in PBS to a total
399 volume of 20 µL and injected intraperitoneally into postnatal day five mice. For protein
400 depletion tests, the screen, and validating hits, mice were injected with 2 x 10^{11} GC of
401 AAV-TBG-Cre. To deplete NK cells, 15 µg/g of anti-NK1.1 (clone PK136, BioXCell) was
402 injected intraperitoneally every three days beginning at postnatal day five. All animal
403 procedures were approved by the Massachusetts Institute of Technology Committee on
404 Animal Care.
405
406 **Cell lines**
407 The mouse hepatocyte cell line AML12 was purchased from the American Type Culture
408 Collection (ATCC) and cultured in DMEM/F12 medium supplemented with 10% fetal
409 bovine serum, 10 µg/mL insulin, 5.5 µg/mL transferrin, 5 ng/mL selenium, and 40 ng/mL
410 dexamethasone (ThermoFisher Scientific). HEK-293T cells were cultured in DMEM
411 supplemented with 10% fetal bovine serum, 100 units/mL penicillin, and 100 µg/mL
412 streptomycin (ThermoFisher Scientific).
413
414 **Vector construction**

415 The vector was produced through the following steps: 1) removal of the EFS-NS
416 promoter and Cas9 from the parental vector and insertion of a hepatocyte-specific
417 promoter driving dsRed expression, 2) replacement of dsRed with mCherry or mTurq2,
418 and 3) removal of the puromycin resistance cassette.
419
420 To produce pLCv2-opti-stuffer-dsRed-puro, 100 ng of a synthetic gblock encoding the
421 the HS-CRM8-TTRmin module[38] upstream of dsRed (Integrated DNA Technologies)
422 and 1 µg of sgOpti (gift from Eric Lander and David Sabatini, Addgene plasmid
423 #85681)[39], a lentiCRISPRv2 derivative containing an optimized scaffold (5'-
424 GTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTT-3')[40] were digested sequentially
425 with NheI and BamHI (New England Biolabs). The vector and fragment were purified
426 using the QIAquick Gel Extraction Kit (Qiagen) and ligated with T4 DNA Ligase (New
427 England Biolabs) in an 11 µL reaction to replace the EFS-NS promoter and Cas9 with
428 the gblock fragment. 2.5 µL of the ligation was used to transform Stbl2 cells (Invitrogen)
429 and DNA was isolated from ampicillin-resistant colonies with the QIAprep Spin Miniprep
430 Kit (Qiagen). Clones were verified by Sanger sequencing (Quintara Biosciences) prior to
431 retransformation and maxiprep using the ZymoPURE II Plasmid Maxiprep Kit (Zymo
432 Research).
433
434 HS-CRM8-TTRmin-dsRed:
435
436 GAATTCGCTAGCACCGGCGCGCCGGGGGAGGCTGCTGGTGAATATTAACCAAGG
437 TCACCCCAGTTATCGGAGGAGCAAACAGGGGCTAAGTCCACACGCGTGGTACCGT

438 CTGTCTGCACATTTCGTAGAGCGAGTGTTCCGATACTCTAATCTCCCTAGGCAAGG
439 TTCATATTTGTGTAGGTTACTTATTCTCCTTTTGTTGACTAAGTCAATAATCAGAATC
440 AGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGG
441 GTATAAAAGCCCCTTCACCAGGAGAAGCCGTCACACAGATCCACAAGCTCCTGAC
442 CGGTTCTAGAGCGCTGCCACCATGGTGCGCTCCTCCAAGAACGTCATCAAGGAGT
443 TCATGCGCTTCAAGGTGCGCATGGAGGGCACCGTGAACGGCCACGAGTTCGAGA
444 TCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCCACAACACCGTGAAACTGA
445 AGGTGACCAAGGGCGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCCCAGT
446 TCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCCGACTACAA
447 GAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGAC
448 GGCGGCGTGGTGACCGTGACCCAAGACTCCTCCCTGCAGGACGGCTGCTTCATCT
449 ACAAGGTGAAGTTCATCGGCGTGAACTTCCCCTCCGACGGCCCCGTAATGCAGAA
450 GAAGACCATGGGCTGGGAGGCCTCCACCGAGCGCCTGTACCCCCGCGACGGCGT
451 GCTGAAGGGCGAAATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCT
452 GGTGGAGTTCAAGTCCATCTACATGGCCAAGAAGCCCGTGCAGCTGCCCGGCTAC
453 TACTACGTGGACTCCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCG
454 TGGAGCAGTACGAGCGCACCGAAGGCCGCCACCACCTGTTCCTGGGATCCGGCG
455 CAACAAACTTCTCTCTGCTGAAACAAGCCGGAGATGTCGAAGAGAATCCTGGACC
456 GACCGAG
457
458 To construct pLCv2-opti-stuffer-mCherry-puro and pLCv2-opti-stuffer-mTurq2-puro,
459 mCherry and mTurq2 were amplified from pKL028 and mTurquoise2-CMV (gifts from
460 Iain Cheeseman), respectively, for 25 cycles with Q5 HotStart Polymerase (New

461 England Biolabs) using the following primers (underlined nucleotides are homologous to
462 the vector):
463
464 pLC_EBFP2_F: GGTTCTAGAGCGCTGCCACCATGGTGAGCAAGGGCGAGGAG
465 pLC_EBFP2_R: GCCGGATCCCTTGTACAGCTCGTCCATGCC
466
467 Amplicons and pLCv2-opti-stuffer-dsRed-puro were digested with XbaI and BamHI HF
468 (New England Biolabs) and purified, ligated, transformed, and DNA was isolated and
469 sequence verified as above.
470
471 To construct pLCv2-opti-stuffer-mCherry and pLCv2-opti-stuffer-mTurq2, a fragment
472 encompassing the WPRE and 3'LTR was amplified from pLCv2-opti-stuffer-mCherry-
473 puro as above using the following primers:
474
475 Puro_removal_F: TGAACGCGTTAAGTCGACAATCAACC
476 Puro_removal_R: TCGAGGCTGATCAGCGGGTTTAAAC
477
478 The amplicon and pLCv2-opti-stuffer-mCherry-puro (or -mTurq2-puro) were digested
479 with BsrGI-HF and PmeI (New England Biolabs) and purified as above. NEBuilder HiFi
480 DNA Assembly Master Mix (New England Biolabs) was used to assemble 25 ng each of
481 vector and fragment in a 20 µL reaction for 15 min at 50 °C. 50 µL of DH5-alpha cells
482 were transformed with 2 µL assembly mix, and DNA was isolated and sequence verified
483 as described above.

484

485 Individual sgRNAs were cloned as previously described[41], using the following

486 oligonucleotides (Integrated DNA Technologies):

487

488 sgAAVS1_F: CACCGGGGCCACTAGGGACAGGAT

489 sgAAVS1_R: AAACATCCTGTCCCTAGTGGCCCC

490 sgMaob_1_F: CACCGACGGATAAAGGATATACTTG

491 sgMaob_1_R: AAACCAAGTATATCCTTTATCCGTC

492 sgMaob_2_F: CACCGGGAAAATCATATGCCTTCAG

493 sgMaob_2_R: AAACCTGAAGGCATATGATTTTCCC

494 sgLmnb2_1_F: CACCGAGGTACGGGAGACCCGACGG

495 sgLmnb2_1_R: AAACCCGTCGGGTCTCCCGTACCTC

496 sgLmnb2_2_F: CACCGCTGCGCACCTACCTCACCGT

497 sgLmnb2_2_R: AAACACGGTGAGGTAGGTGCGCAGC

498 sgCxorf38_1_F: CACCGGTCAACCACAAAGTGATCAC

499 sgCxorf38_1_R: AAACGTGATCACTTTGTGGTTGACC

500 sgCxorf38_2_F: CACCGTCTGCGCATACCTGACACTG

501 sgCxorf38_2_R: AAACCAGTGTCAGGTATGCGCAGAC

502 sgTap1_1_F: CACCGGGTGCCAACGAGCCACTGAG

503 sgTap1_1_R: AAACCTCAGTGGCTCGTTGGCACCC

504 sgTap1_2_F: CACCGGGTAGAGAACGAATGAGACA

505 sgTap1_2_R: AAACTGTCTCATTCGTTCTCTACCC

506

507 **Lentivirus preparation, titration, and concentration**

508 HEK-293T cells were seeded at a density of 750,000 cells/mL in 20 mL viral production
509 medium (IMDM supplemented with 20% inactivated fetal serum, Thermo Fisher
510 Scientific). After 24 hours, media was changed to fresh viral production medium. At 32
511 hours post-seeding, cells were transfected with a mix containing 76.8 µL Xtremegene-9
512 transfection reagent (Thermo Fisher Scientific), 3.62 µg pCMV-VSV-G[42] (gift from Bob
513 Weinberg, Addgene plasmid #8454), 8.28 µg psPAX2 (gift from Didier Trono, Addgene
514 plasmid #12260), and 20 µg sgRNA plasmid in Opti-MEM (Thermo Fisher Scientific) to
515 a final volume of 1 mL. Media was changed 16 hours later to fresh viral production
516 medium. At 48 hours after transfection, virus was collected and filtered through a 0.45
517 µm filter, aliquoted, and stored at -80 °C until use.

518

519 To determine lentivirus titer, AML12 cells were transduced with a dilution series of
520 lentivirus in the presence of 10 µg/mL polybrene for 16 hours. After four days, cells were
521 harvested for flow cytometry analysis to determine percent of mTurq2- or mCherry-
522 positive cells. To concentrate lentivirus, lentiviral supernatant was ultracentrifuged at
523 23,000 RPM at 4 °C for 2 hours in an SW 32 Ti swinging bucket rotor (Beckman
524 Coulter). After centrifugation, media was decanted and pellets were air-dried at room
525 temperature for 15 minutes. Pellets were then resuspended in PBS at room temperature
526 for 30 minutes with gentle trituration. Concentrated lentivirus in PBS was stored for up
527 to one week at 4 °C prior to injection into mice.

528

529 **sgRNA validation**

530 To validate individual sgRNA sequences for the ability to deplete target proteins, AML12
531 cells stably expressing Cas9-EGFP were transduced with sgAAVS1-mTurq2 or sgRNA-
532 mCherry lentiviruses as described for lentivirus titration. After four days, transduced
533 cells were subjected to fluorescence activated cell sorting to purify cells expressing both
534 EGFP and mTurq2 or EGFP and mCherry and cultured for another seven days. Cells
535 were then harvested to prepare protein lysates and perform immunoblotting. To validate
536 sgRNAs targeting Tap1, cells were treated with 10 ng/mL of IFNγ 20 hours prior to
537 harvest to upregulate Tap1 expression such that it could be detected by
538 immunoblotting.
539
540 **Immunostaining**
541 Livers were harvested and fixed in 4% paraformaldehyde in PBS at room temperature
542 for 16-24 hours. Tissues were then washed with PBS and frozen in O.C.T. Compound
543 (Tissue-Tek). Tissue sections of 12 to 30 μm thickness were prepared using a cryostat
544 and adhered to Superfrost Plus Slides (Fisher Scientific). Slides were stored at -20 °C
545 until use. To visualize endogenous mCherry and mTurq2 fluorescence, slides were
546 dried at room temperature for 15 minutes, rehydrated in PBS for 5 minutes,
547 permeabilized with 1% Triton X-100 in PBS for 15 minutes, and counterstained with
548 Alexa Fluor 488 Phalloidin (ThermoFisher Scientific) diluted 1:500 in blocking buffer (3%
549 bovine serum albumin and 0.3% Triton X-100 in PBS). To immunostain for endogenous
550 proteins, slides were dried at room temperature for 4-24 hours and rehydrated in PBS
551 for 5 minutes. Antigen retrieval was then performed by pressure cooking slides in
552 sodium citrate buffer (10 mM tri-sodium citrate dihydrate, 0.05% Tween-20, pH 6.0) for

553 20 minutes in an Instant Pot (Amazon). Slides were rinsed in PBS for 5 minutes, dried
554 briefly, and sections outlined with an ImmEdge hydrophobic pen (Vector Laboratories).
555 Sections were permeabilized with 1% Triton X-100 in PBS for 15 minutes and blocked
556 with blocking buffer for one hour. Sections were then incubated in primary antibodies
557 diluted in blocking buffer at room temperature for 12-24 hours. Sections were washed
558 with blocking buffer three times for 10 minutes each. Sections were then incubated in
559 AlexaFluor secondary antibodies (ThermoFisher Scientific) diluted 1:1,000 in blocking
560 buffer at room temperature for 1-2 hours. In some cases, 5 µg/mL Hoechst 33342
561 (ThermoFisher Scientific) was added to the secondary antibody solution. Sections were
562 washed with blocking buffer twice for 10 minutes each followed by one wash with PBS
563 for 5 minutes. Slides were then mounted in ProLong Gold Antifade reagent
564 (ThermoFisher Scientific).
565
566 The following primary antibodies were used: Cas9 (1:200, clone 7A9-3A3, Abcam
567 ab191468), asialoglycoprotein receptor 1 (ASGR1) (1:500, clone 114, Sino Biological
568 50083-R114), mCherry (1:500, clone 16D7, ThermoFisher Scientific M11217),
569 monoamine oxidase B (MAO-B) (1:1,000, Novus Biologicals NBP1-87493), lamin B2
570 (1:1,000, clone EPR9701(B), Abcam ab151735), actin (1:250, clone AC-74, Sigma
571 Aldrich A2228), CD45 (1:500, Abcam ab10558), and Ki-67 (1:200, clone SP6, Abcam
572 ab16667). The Cas9 antibody was directly conjugated to AlexaFluor 647 using the
573 AlexaFluor 647 Antibody Labeling Kit (ThermoFisher Scientific). The actin antibody was
574 directly conjugated to DyLight 405 using the DyLight 405 antibody labeling kit
575 (ThermoFisher scientific).

576

577 **Image analysis**

578 Images were acquired using a CSU-22 spinning disc confocal head (Yokogawa) with
579 Borealis modification (Andor) mounted on an Axiovert 200M microscope (Zeiss) with
580 10X or 40X objectives (Zeiss), an Orca-ER CCD camera (Hamamatsu), and MetaMorph
581 acquisition software (Molecular Devices).

582

583 Images were analyzed using Volocity (Quorum Technologies). To measure MAO-B and
584 lamin B2 intensity, a single Z plane at the center of the cell was identified and the
585 cytoplasm or nucleus was outlined to measure the signal intensity per µm. A similar
586 procedure was done on sections stained only with secondary antibodies to calculate the
587 average background intensity. This average background intensity was subtracted from
588 each MAO-B and lamin B2 intensity measurement and the background-subtracted
589 measurements were then normalized within a given sample (mCherry-positive or –
590 negative hepatocytes within a single liver).

591

592 **Immunoblotting**

593 To prepare protein lysates from cultured cells, 500,000 cells were pelleted and
594 resuspended in 100 µL of 2X Laemmli sample buffer (100 mM Tris pH 6.8, 20%
595 glycerol, 4% SDS, 0.02% bromophenol blue, 5% β-mercaptoethanol). Lysates were
596 homogenized by pipetting followed by trituration through a 31-gauge needle and boiled
597 for 5 minutes. Samples were separated on homemade polyacrylamide gels and
598 transferred to Immobilon-FL membranes (Millipore) via wet transfer. Membranes were

599 blocked in 5% bovine serum albumin in TBST (50 mM Tris pH 8.0, 150 mM NaCl, 0.1%
600 Tween-20) for 1 hour at room temperature. Membranes were incubated in primary
601 antibody diluted in blocking solution at 4 °C with rocking overnight and washed with
602 TBST for five minutes five times. Membranes were incubated in HRP-conjugated goat
603 anti-rabbit secondary antibody (Abcam ab205718) diluted 1:50,000 in blocking solution
604 at room temperature with rocking for one hour and washed with TBST for five minutes
605 five times. Membranes were incubated in ECL Prime Western Blotting Detection
606 Reagent (GE Healthcare) for five minutes and exposed to CL-Xposure Film (Thermo
607 Fisher Scientific).
608
609 The following primary antibodies were used: Cxorf38 (1:1,000, Invitrogen PA5-62139)
610 and Tap1 (1:1,000, Cell Signaling Technology 12341).
611
612 **Flow cytometry**
613 To quantify NK cells in the liver by flow cytometry, livers were harvested, minced into 1
614 mm pieces, and incubated in digestion solution (1% fetal bovine serum, 40 U/mL
615 DNase, 125 U/mL collagenase type IV in Hank's balanced salt solution (HBSS)
616 containing 10 mM HEPES) at 37 °C for 30 minutes with shaking. Digested tissue was
617 passaged through a 70 μm strainer, combined with 10 mL of HBSS containing 10 mM
618 HEPES, and pelleted at 500 g for 5 minutes. Pellet was resuspended in 5 mL of 45%
619 Percoll (Sigma) in PBS and gently layered onto 3 mL of 55% Percoll in PBS. Gradient
620 was centrifuged at 900 g at 4 °C for 20 minutes with low acceleration and no braking.
621 The cell layer at the interface of the two Percoll layers was removed and transferred to

622 10 mL of HBSS containing 10 mM HEPES and pelleted at 500 g for 5 minutes. Cells
623 were diluted to <10^6 cells per 100 µL in flow cytometry buffer (1% fetal bovine serum
624 and 0.05% sodium azide in PBS) and incubated with conjugated primary antibodies on
625 ice in the dark for 20 minutes. Cells were washed twice with flow cytometry buffer,
626 resuspended in flow cytometry buffer, incubated with 500 ng/mL DAPI for 5 minutes,
627 and analyzed using a LSRFortessa flow cytometer with FACSDiva software (BD
628 Biosciences).
629
630 The following fluorophore-conjugated primary antibodies were used: CD3-Brilliant Violet
631 421 (1:20, clone 17A2, BioLegend 100227) and NK1.1-AlexaFluor 488 (1:200, clone
632 PK136, BioLegend 108717).
633
634 **RNA sequencing**
635 For surgical resection time points, partial hepatectomies were performed on 8 week-old
636 mice as previously described[43]. For toxic injury time points, 8 week-old mice were
637 injected intraperitoneally with 2 µL/gram of 25% carbon tetrachloride diluted in corn oil
638 (Sigma Aldrich). For all time points, livers from three male C57BL/6J mice were
639 harvested, flushed with PBS, immediately immersed in RNAlater (Qiagen), incubated at
640 room temperature for 24 hours, and stored at -20 °C until future use. To isolate RNA, 30
641 mg of each tissue was removed from RNAlater and homogenized in 700 µL of QIAzol
642 lysis reagent (Qiagen) using the TissueRuptor homogenizer (Qiagen). RNA was purified
643 using the miRNeasy Kit (Qiagen) according to kit instructions and eluted in 30 µL of
644 nuclease-free water. RNA sequencing libraries were prepared using KAPA mRNA

645 HyperPrep Kit (KAPA Biosystems) according to manufacturer instructions. Briefly, 0.1-1
646 ug of total RNA was enriched for polyadenylated sequences using oligo-dT magnetic
647 bead capture. The enriched mRNA fraction was then fragmented and first-strand cDNA
648 generated using random primers. Strand specificity was achieved during second-strand
649 cDNA synthesis by replacing dTTP with dUTP to quench the second strand during
650 amplification. The resulting cDNA was A-tailed and ligated with indexed adapters. The
651 library was amplified using a DNA polymerase that cannot incorporate past dUTPs to
652 quench the second strand during PCR. The libraries were quantified using a KAPA
653 qPCR Library Quantification Kit (KAPA Biosystems) as per manufacturer instructions.
654 The samples were sequenced on a HiSeq 2500 (Illumina) based on qPCR
655 concentrations. Base calls were performed by the instrument control software and
656 further processed using the Offline Base Caller version 1.9.4 (Illumina). Samples were
657 mapped with STAR version 2.6.1a[44] to the mouse genome release mm10, using a gtf
658 file from ENSEMBL version GRCm38.91, and setting the maximum intron length
659 ("alignIntronMax") parameter to 50000. We ran featureCounts version 1.6[45] to assign
660 reads to genes using the same gft file and setting "-s" parameter to 2. We normalized
661 gene counts with DESeq2 version 1.22.2[46]. FPKMs were calculated using the function
662 fpkm within the DESeq2 package. The FPKMs values for the three replicates were
663 averaged, and protein coding genes were selected based on the annotation in the gtf
664 file.
665
666 **sgRNA library preparation**

667 Genes with an average FPKM > 0.3 in any of the RNA sequencing time points were
668 chosen to build a liver transcriptome-wide library. sgRNA sequences were designed
669 using the Broad Institute GPP sgRNA Designer[13,47] using the Azimuth 2.0 rule set. For
670 genes which were not identified by the program, alternative gene names from
671 ENSEMBL versions GRCm38.76 - .93 were attempted. A small number of designed
672 sgRNAs targeted multiple genes; the sgRNA names and gene names were manually
673 annotated to indicate all targeted genes for these cases. Non-targeting and control-
674 gene-targeting sgRNAs from Doench et al. 2014[13] were also included. sgRNA
675 sequences from this control set that were identical to a sequence already in our library
676 were annotated according to the targeted gene; those that did not overlap with
677 sequences in our sgRNA library were annotated as control sgRNAs. The library
678 contains 71,878 sgRNAs targeting 13,189 genes.
679
680 For sgRNAs beginning with a nucleotide other than G, a G was prepended. The
681 following adapters were added to all sgRNA sequences:
682
683 Upstream: 5'-TATCTTGTGGAAAGGACGAAACACC-3'
684 Downstream: 5'-AAGAGCTATGCTGGAAACAGCATAGC-3'
685
686 Multiple rounds of cloning were combined to generate the final plasmid library. The
687 oligonucleotide library (Agilent Technologies) was amplified for 16 cycles using Q5
688 HotStart Polymerase (New England Biolabs) using a gradient annealing temperature
689 ranging from 50-62 °C across 8, 50 µL reactions. Reactions were pooled and purified by

690 DNA Clean and Concentrator 5 (Zymo Research). pLCv2-opti-stuffer-mCherry was
691 digested as described [41], and either gel purified using a Zymoclean Gel DNA Recovery
692 Kit (Zymo Research) followed by Ampure XP bead purification (Beckman Coulter) or
693 DNA Clean and Concentrator 5. The library was assembled using NEBuilder HiFi DNA
694 Assembly Master Mix (New England Biolabs) in 4 x 20 µL reactions at 50 °C for 1 hour
695 using 100 ng of vector per 5-10 ng of PCR amplicon. The reactions were combined and
696 2.5 µL of the assembly reaction or a control reaction without amplicon were used to
697 transform NEB5-alpha cells (New England Biolabs) to measure background assembly.
698 Subsequently, the assembly reactions were combined, concentrated using Ampure XP
699 beads, resuspended in 8 µL water, and used to electroporate 1-4 tubes of Endura DUO
700 electrocompetent cells (Lucigen) at 1.8 kV distributed over 2 cuvettes (0.1 cm gap
701 width) per tube using a Micropulser Electroporator (Bio-Rad Laboratories). 10-fold serial
702 dilutions of a 10 µL aliquot were plated on LB plates with ampicillin at 100 µg/mL to
703 assess electroporation efficiency, and the remainder of each electroporation (2
704 cuvettes) was plated on LB agar supplemented with 100 µg/mL ampicillin in 4 x 245 mm
705 square bioassay dishes (Corning). Plates were incubated overnight at 30 °C and
706 colonies were scraped the next morning. DNA was isolated using the ZymoPURE II
707 Plasmid Maxiprep Kit (Zymo Research). Plasmid DNA from multiple rounds of assembly
708 and electroporation were combined according to the measured electroporation
709 efficiency to achieve 25-fold coverage of the library. sgRNA representation was
710 measured by high-throughput sequencing as described below.
711

712 To improve coverage of some of the sgRNAs in the library, a second library containing
713 ~7,500 sgRNAs was synthesized and cloned as above, with the following modifications:
714 assembly was performed using NEB Gibson Assembly mix (New England Biolabs)
715 using a ratio of 200 ng vector : 10 ng sgRNA in each 20 µL reaction, and the final
716 combined and concentrated reaction was used to electroporate a single tube of Endura
717 DUO cells.
718
719 Subsequent propagation of the plasmid library was performed using 50 ng plasmid
720 library per single tube of Endura DUO cells.
721
722 All steps were performed according to manufacturer's instructions, except where noted.
723
724 **Genomic DNA isolation**
725 Livers were harvested from mice, separated into individual lobes, minced into 15 mg
726 pieces using a razor blade, snap-frozen in liquid nitrogen, and stored at -80 °C until use.
727 Genomic DNA (gDNA) was isolated from livers using the illustra blood genomicPrep
728 Mini Spin Kit (Cytiva) using one column for every 7.5 mg of tissue. The manufacturer's
729 protocol was used with the following modifications: 20 µL of 10 mg/mL Proteinase K
730 (Millipore-Sigma) solution in water was added per 7.5 mg of tissue. Tissue was
731 disrupted by thoroughly pipetting prior to adding lysis buffer, vortexing, and incubating at
732 56 °C overnight. Elution was performed using 25 µL of water pre-heated to 70 °C.
733 Samples were combined by lobe and concentration was measured using the Qubit
734 dsDNA HS Assay Kit (Invitrogen).

735

736 For the induction, equal amounts of gDNA from each lobe were combined within each
737 mouse, and equal inputs from four mice were combined to prepare a single sequencing
738 library. For the endpoint, gDNA from each lobe within a mouse was combined
739 proportionally to the average lobe mass across mice measured at liver harvest. A
740 sequencing library was prepared for each mouse individually using equal total gDNA
741 input per mouse.

742

743 **Sequencing library preparation and DNA sequencing**

744 All PCR reactions were performed in 50 µL reactions using ExTaq Polymerase (Takara
745 Bio) with the following program:

746

747 1 cycle	95 °C	5 min
748 14 or 28 cycles	95 °C	10 sec
749	60 °C	15 sec
750	72 °C	45 sec
751 1 cycle	72 °C	5 min
752 1 cycle	4 °C	hold

753

754 Using the following primers:

755

756 Forward: AATGATACGGCGACCACCGAGATCTACACCGACTCGGTGCCACTTTT

757 Reverse:

758 CAAGCAGAAGACGGCATACGAGATCnnnnnnTTTCTTGGGTAGTTTGCAGTTTT

759

760 Where “nnnnnn” denotes the barcode used for multiplexing.

761

762 10 ng of plasmid DNA was amplified for 14 cycles in 4 x 50 µL reactions. 1, 3, or 6 µg of
763 gDNA was initially amplified for 28 cycles in 50 µL test PCR reactions. Subsequently,
764 226 µg of gDNA (induction) was used in 38 reactions, or 75 µg of gDNA (endpoint) was
765 used in 25 reactions per mouse. All reactions were cleaned and concentrated using
766 Ampure XP beads prior to sequencing for 50 cycles on an Illumina Hiseq 2500 using the
767 following primers:

768

769 Read 1 sequencing primer:

770 GTTGATAACGGACTAGCCTTATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTA
771 AAC

772 Index sequencing primer:

773 TTTCAAGTTACGGTAAGCATATGATAGTCCATTTTAAAACATAATTTTAAAACTGCAA
774 ACTACCCAAGAAA

775

776 Base calls were performed by the instrument control software and further processed
777 using the Offline Base Caller (Illumina) v. 1.9.4.

778

779 **Screen analysis**

780 For initial measurement of sgRNA representation in the plasmid library or induction time
781 point, sequencing reads were mapped to the library, each sgRNA was given a
782 pseudocount of 1, and reads per million (RPM) was calculated as previously
783 described[12]. Raw counts were processed using MAGeCK for downstream analysis[14].
784 The plasmid library and induction timepoint were used as control samples and to
785 estimate variance, and each endpoint mouse was processed separately. For mouse 4,
786 sgLmnb2_1 was removed prior to MAGeCK analysis, as the high representation of this
787 sgRNA (an sgRNA used for development of the screening method) was likely due to
788 contamination during sequencing library preparation. Counts data from Shohat et al.
789 (day 18) and Tzelepis et al. (day 14) were processed individually using MAGeCK[23,24].
790 The corresponding plasmid libraries were used as control samples. For our screen,
791 Shohat et al., and Tzelepis et al., the size factor for normalization was estimated using
792 the control sgRNA set (Supplementary Table 2). For Tzelepis et al., the three replicate
793 day 14 samples were processed together to generate a single gene score, and those
794 three samples were used to estimate variance. For all screens, the gene test FDR
795 threshold was set to 0.05, the sgRNA p-value was FDR-adjusted, and the gene score
796 was calculated using the median. Twenty-four human hepatocellular carcinoma screens
797 from the CRISPR (Avana) Public 20Q4 release were downloaded from the Broad
798 DepMap portal using “Liver” as a lineage filter and “Hepatocellular Carcinoma” as a
799 lineage subtype filter[25,26].
800
801 All downstream analyses were performed in R version 3.6.0, and all plots were
802 generated in either base R, using the R corrplot package, or in GraphPad Prism Version

803 7.0d. For comparisons within our screen, the gene scores from individual mice were not
804 normalized across mice, as each mouse serves as a replicate screen. The gene score
805 for each gene across mice was tested against all gene scores using an unpaired two-
806 sample Wilcoxon test. The p-values from this test were adjusted using the Benjamini-
807 Hochberg (FDR) procedure. The median log2 fold change across mice was used as
808 input for pre-ranked GSEA using the c2.cp.kegg.v7.1.symbols.gmt gene sets. For
809 comparisons between screens from different sources, all screens from all the sources in
810 the specific comparison were quantile normalized to one another using the
811 preprocessCore R package prior to calculating the median log2 fold change within the
812 screens from each source. This normalized median log2 fold change was subtracted
813 from the normalized median log2 fold change of our screens to generate a differential
814 score used as input for pre-ranked GSEA using the c2.cp.kegg.v7.1.symbols.gmt gene
815 sets. For converting mouse gene symbols to human gene symbols, the
816 Mouse_Gene_Symbol_Remapping_to_Human_Orthologs_MSigDB.v7.1.chip was used
817 (note that this excludes genes that have multiple annotations in either human or
818 mouse).
819
820 Pearson correlation was used to compare gene effects between mice. Spearman
821 correlation was used to compare gene effects with liver mRNA expression and protein
822 half-life in mouse cells[15]. The TUSON dataset of predicted tumor suppressor genes was
823 sorted by ascending FDR q-value and the top 50 genes present in the compared
824 datasets were used[21]. Distribution differences were tested and p-values were calculated
825 using the Kolmogorov-Smirnov test. A one-sided test was used for gene sets for which

826 a phenotype could be predicted (core essential genes, tumor suppressor genes, and
827 control enriched and depleted genes); all other comparisons used a two-sided test.
828
829 For comparison of sex-specific fitness effects, only genes with an average of > 2
830 sgRNAs detected across mice were considered. For sex-specific enriched genes, a
831 median fold change (log2) > 0.5 across two mice of a given sex and an absolute median
832 fold change (log2) difference of > 0.25 compared to the other sex was required. To
833 identify true tumor-suppressor-like genes, a median fold change (log2) > -0.5 was
834 required in mice of the other sex. For sex-specific depleted genes, a median fold
835 change (log2) of < -0.5 across two mice of a given sex and an absolute median fold
836 change (log2) difference of > 0.75 compared to the other sex was required. To identify
837 true sex-specific essential genes, a median fold change (log2) > -0.5 was required in
838 mice of the other sex.
839
840 **Software information**
841 STAR version 2.6.1a
842 Conda version 4.9.2
843 MAGeCK-RRA version 0.5.9.2
844 R version 3.6.0
845 featureCounts version 1.6
846 DESeq2 version 1.22.2
847 preprocessCore version 1.48.0
848 corrplot version 0.84

849 GSEA version 4.1.0

850 Mouse_Gene_Symbol_Remapping_to_Human_Orthologs_MSigDB.v7.1.chip

851 Human_Symbol_with_Remapping_MsigDB.v7.1.chip

852 c2.cp.kegg.v7.1.symbols.gmt

853 GraphPad Prism version 7.0d

854

855 **DATA AVAILABILITY**

856 Sequencing data generated in this manuscript is in the process of being uploaded to

857 publicly accessible online repositories. All other data are included within the manuscript

858 figures or as supplementary information.

859

860 **METHODS REFERENCES**

861 38. Chuah, M. K. *et al.* Liver-specific transcriptional modules identified by genome-

862 wide in silico analysis enable efficient gene therapy in mice and non-human

863 primates. *Mol. Ther.* **22,** 1605–1613 (2014).

864 39. Fulco, C. P. *et al.* Systematic mapping of functional enhancer-promoter

865 connections with CRISPR interference. *Science* **354,** 769–773 (2016).

866 40. Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an

867 optimized CRISPR/Cas system. *Cell* **155,** 1479–1491 (2013).

868 41. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide

869 libraries for CRISPR screening. *Nat Meth* **11,** 783–784 (2014).

870 42. Stewart, S. A. *et al.* Lentivirus-delivered stable gene silencing by RNAi in primary

871 cells. *RNA* **9,** 493–501 (2003).

872 43. Knouse, K. A., Lopez, K. E., Bachofner, M. & Amon, A. Chromosome Segregation
873 Fidelity in Epithelia Requires Tissue Architecture. *Cell* **175,** 200–211.e13 (2018).
874 44. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–
875 21 (2013).
876 45. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose
877 program for assigning sequence reads to genomic features. *Bioinformatics* **30,**
878 923–930 (2014).
879 46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
880 dispersion for RNA-seq data with DESeq2. *Genome Biol* **15,** 550 (2014).
881 47. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize
882 off-target effects of CRISPR-Cas9. *Nat Biotech* **34,** 184–191 (2016).
883

884 **ACKNOWLEDGMENTS**

885 We thank Mehreen Khan and Keya Viswanathan of the Whitehead Institute Functional
886 Genomics Platform for technical assistance, Amanda Chilaka and Sumeet Gupta of the
887 Whitehead Institute Genome Technology Core for high-throughput sequencing, Inma
888 Barrasa of the Whitehead Institute Bioinformatics and Research Computing Core for
889 RNA sequencing analysis, the Whitehead Institute Flow Cytometry Core for cell sorting
890 and the Whitehead Institute Keck Microscopy Facility for microscopy. We thank Agilent
891 Technologies for their commitment to supporting academic research. We thank Leah
892 Schmidt for helpful discussions and Iain Cheeseman, Jonathan Weissman, Peter
893 Reddien, Raghu Chivukula, Sharon Grossman, Tobiloba Oni, and members of the
894 Knouse laboratory for their comments on the manuscript. This work was supported by

the National Institutes of Health NIH Director's Early Independence Award (DP5OD026369) to K.A.K. K.A.K. is also supported by the Scott Cook and Signe Ostby Fund.

AUTHOR CONTRIBUTIONS

H.R.K. and K.A.K. conceived the project, performed all experiments, analyzed the data, and wrote the manuscript.

COMPETING INTERESTS

H.R.K. and K.A.K. are co-inventors on a patent filed by the Whitehead Institute related to work in this manuscript.

Supplementary information is available for this paper

Correspondence and requests for materials should be addressed to K.A.K. All plasmids and the plasmid library are in the process of being deposited at Addgene.

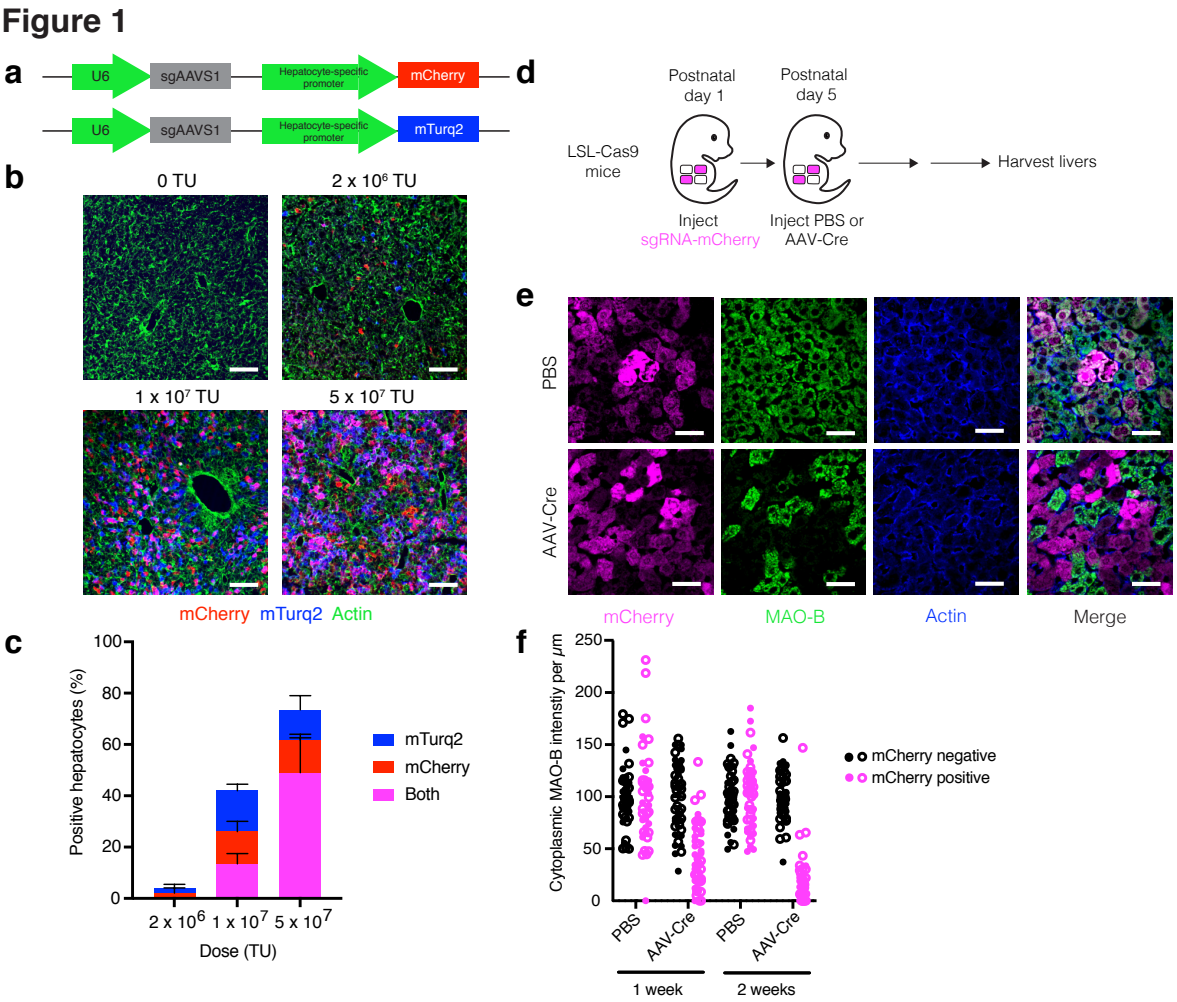
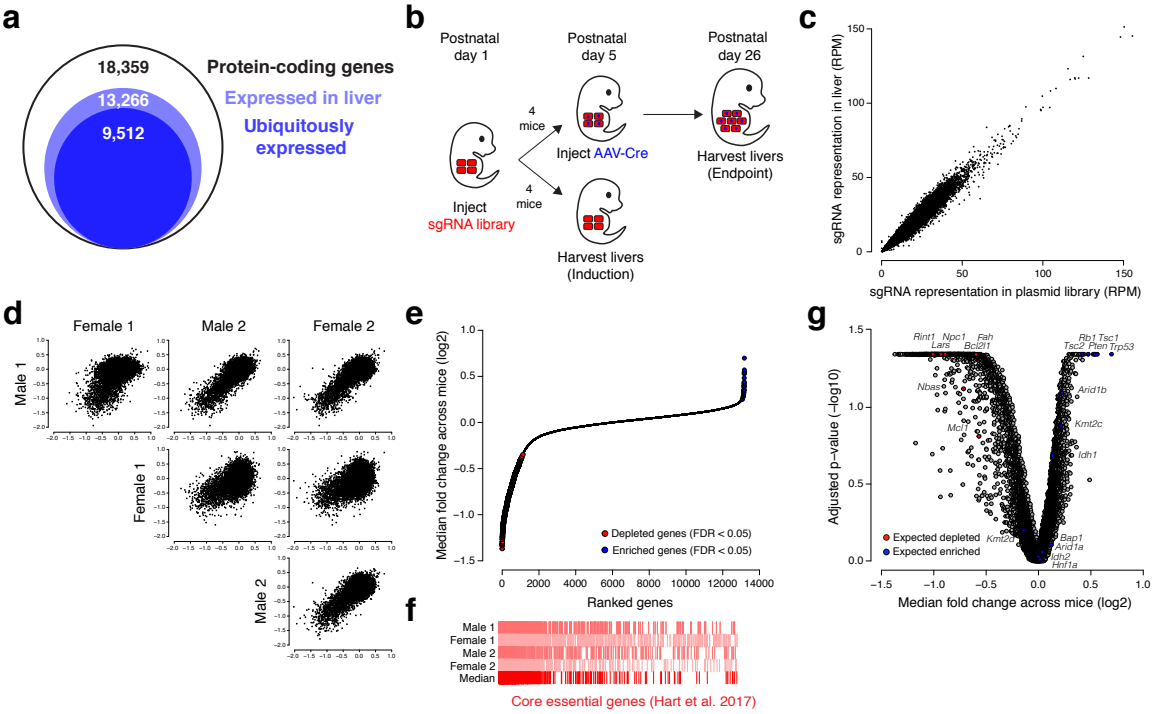


Fig. 1 | Efficient delivery of sgRNAs and depletion of long-lived proteins in the liver. a, Lentiviral vectors for U6-driven expression of an sgRNA and hepatocyte specific-expression of a fluorescent reporter (mCherry or mTurq2). **b,** Images of endogenous mCherry and mTurq2 fluorescence in livers from mice four days after injection of an equal mixture of sgAAVS1-mCherry and sgAAVS1-mTurq2 lentiviruses. Livers were counterstained with phalloidin (green) to label actin. Scale bars, 100 µm. **c,** Percent mCherry-, mTurq2-, and double-positive hepatocytes in livers from mice four days after injection with an equal mixture of sgAAVS1-mCherry and sgAAVS1-mTurq2 lentiviruses. n = 3 mice per dose and 200 hepatocytes per mouse. Error bars indicate

920 standard deviation. **d,** Scheme for inducing protein depletion in LSL-Cas9 mice. **e,**
921 Images of livers from LSL-Cas9 mice injected with sgMaob-mCherry followed by PBS or
922 AAV-Cre immunostained for mCherry (magenta), MAO-B (green), and actin (blue).
923 Scale bars, 45 μm. **f,** Cytoplasmic MAO-B intensity per μm in mCherry-positive and
924 mCherry-negative hepatocytes from LSL-Cas9 mice injected with sgMaob-mCherry
925 followed by PBS or AAV-Cre. n = 1 male and 1 female mouse per condition and 25 cells
926 per mouse. Closed and open circles represent values from male and female mouse,
927 respectively.

Figure 2



928

929 **Fig. 2 | A genome-wide screen for hepatocyte fitness in the neonatal mouse liver.**

930 **a,** Number of protein-coding genes expressed in the liver as determined by RNA

931 sequencing of livers at various time points. **b,** Scheme for performing a genome-wide

932 screen for hepatocyte fitness in neonatal mice. **c,** Representation of sgRNAs in livers

933 four days after injection with lentiviral library relative to the sgRNA representation in the

934 plasmid library expressed as reads per million (RPM). n = 2 male and 2 female mice.

935 Pearson correlation r = 0.97. **d,** Pairwise comparisons of median fold change (log2) for

936 each gene for each mouse at the endpoint of the screen. **e,** Genes ranked by median

937 gene score across mice (log2) with significantly depleted genes denoted by red points

938 and significantly enriched genes denoted by blue points (FDR < 0.05 by two-sample

939 Wilcoxon test). **f,** Core essential genes (red bars) positioned based on gene rank as in

940 (**e**) revealing their significant depletion in individual mice and across mice. $p < 2.2 \times 10^{-}$

941 ^{16} for each mouse and across mice by one-sided Kolmogorov-Smirnov test. **g,**
942 Benjamini-Hochberg-adjusted Wilcoxon p-value (-log10) versus the median fold change
943 of gene scores across mice (log2) for each gene in the screen. Highlighted are control
944 gene sets consisting of tumor suppressor genes in hepatocellular carcinoma (expected
945 to enrich, blue) and genes required for hepatocyte viability (expected to deplete, red).
946 Expected depleted $p = 3.5 \times 10^{-6}$ by one-sided Kolmogorov-Smirnov test, expected
947 enriched $p = 4.7 \times 10^{-4}$ by one-sided Kolmogorov-Smirnov test.

Figure 3

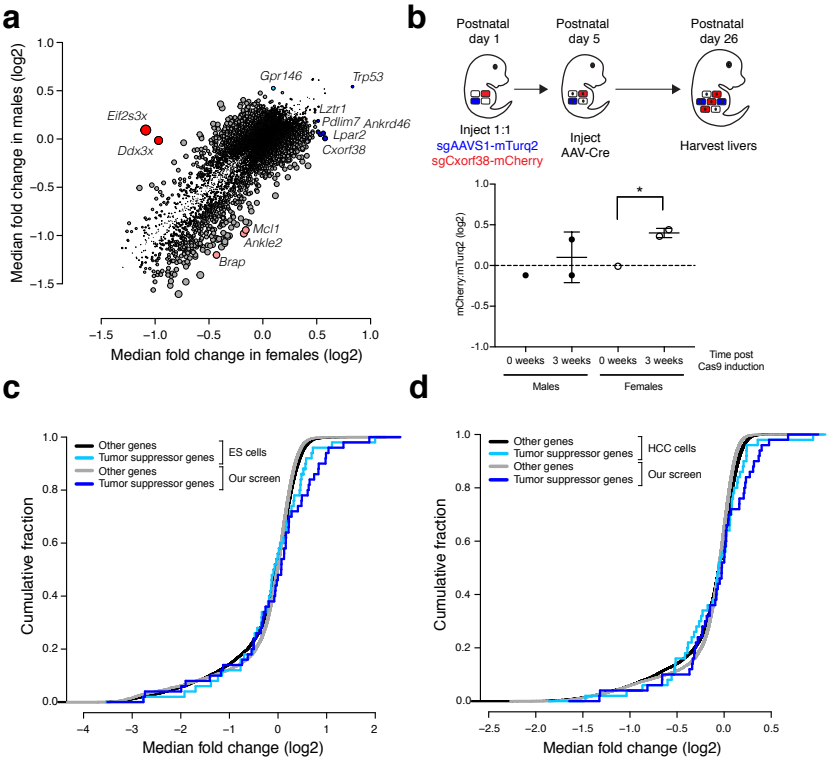


Fig. 3 | An *in vivo* screen uncovers sex-dependent fitness effects and canonical tumor suppressor genes. a, Median fold change (log2) across males versus median fold change (log2) across females for each gene. Highlighted are genes uniquely enriched in females (blue), genes uniquely enriched in males (cyan), genes uniquely depleted in females (red), and genes uniquely depleted in males (pink). Point size is proportional to the absolute difference in median log2 fold change between females and males. **b,** Scheme for *in vivo* competition between hepatocytes expressing sgAAVS1-mTurq2 and sgCxorf38-mCherry (top panel). Ratio (log2) of sgCxorf38-mCherry-positive to sgAAVS1-mTurq2-positive hepatocytes in livers from male and female mice 0 or 3 weeks after Cas9 induction (bottom panel). n = one mouse per gender and >200 hepatocytes per mouse for 0 week time point and two mice per gender and >350

960 hepatocytes per mouse for 3 week time point. *p = 0.016 (female 0 weeks versus
961 female 3 weeks) by one-tailed Fisher's exact test. **c,** Cumulative fraction of tumor
962 suppressor genes (cyan, blue) and other genes (black, grey) based on quantile-
963 normalized median fold change (log2) of their gene scores across screens in mouse
964 embryonic stem (ES) cells and our screen. ES cells $p > 0.05$ by one-sided Kolmogorov-
965 Smirnov test, our screen $p = 6.5 \times 10^{-4}$ by one-sided Kolmogorov-Smirnov test. **d,**
966 Cumulative fraction of tumor suppressor genes (cyan, blue) and other genes (black,
967 grey) based on quantile-normalized median fold change (log2) of their gene scores
968 across screens in human hepatocellular carcinoma (HCC) cell lines and our screen.
969 HCC cells $p > 0.05$ by one-sided Kolmogorov-Smirnov test, our screen $p = 0.0028$ by
970 one-sided Kolmogorov-Smirnov test.

Figure 4

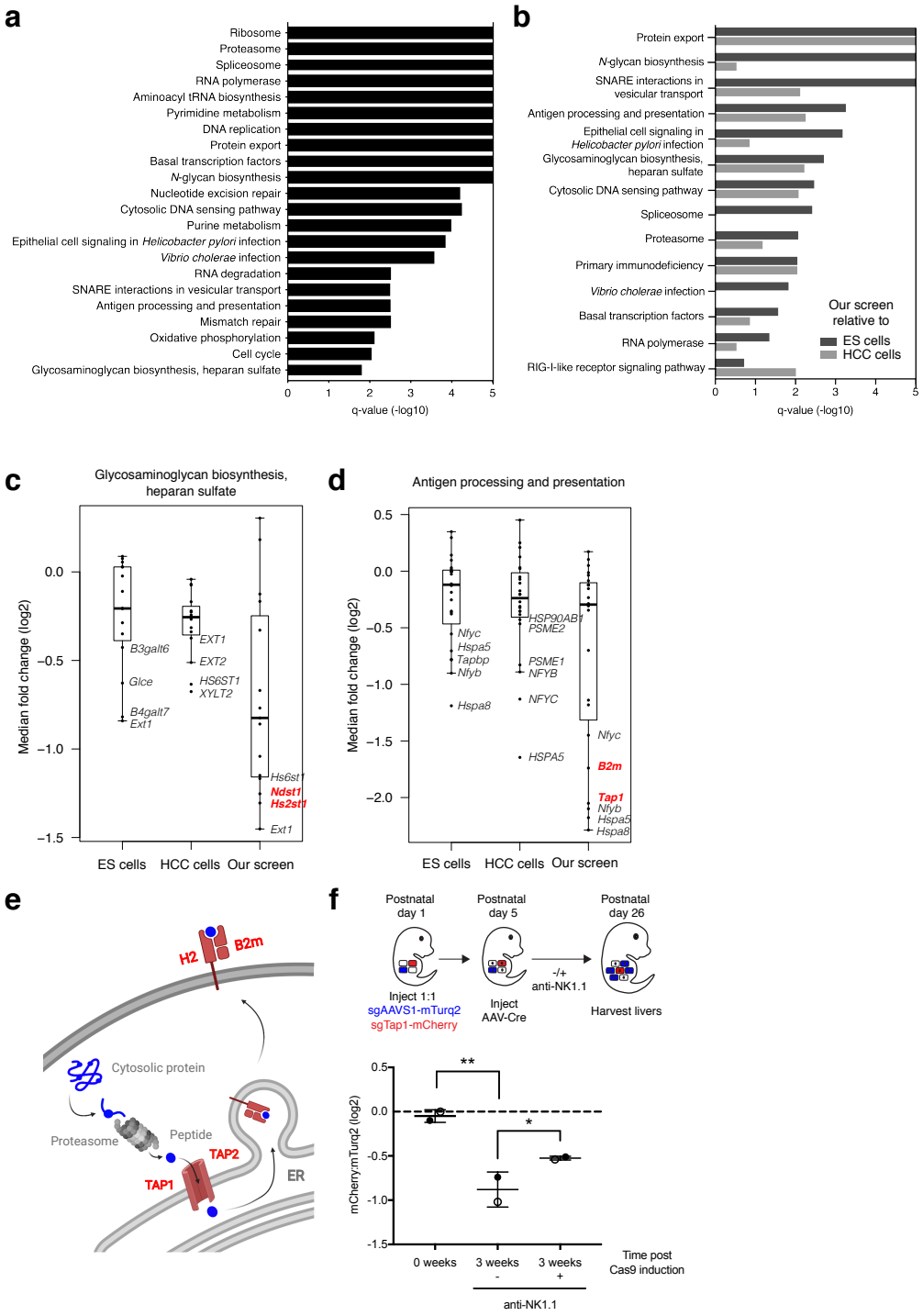


Fig. 4 | Class I MHC is essential for hepatocyte fitness.

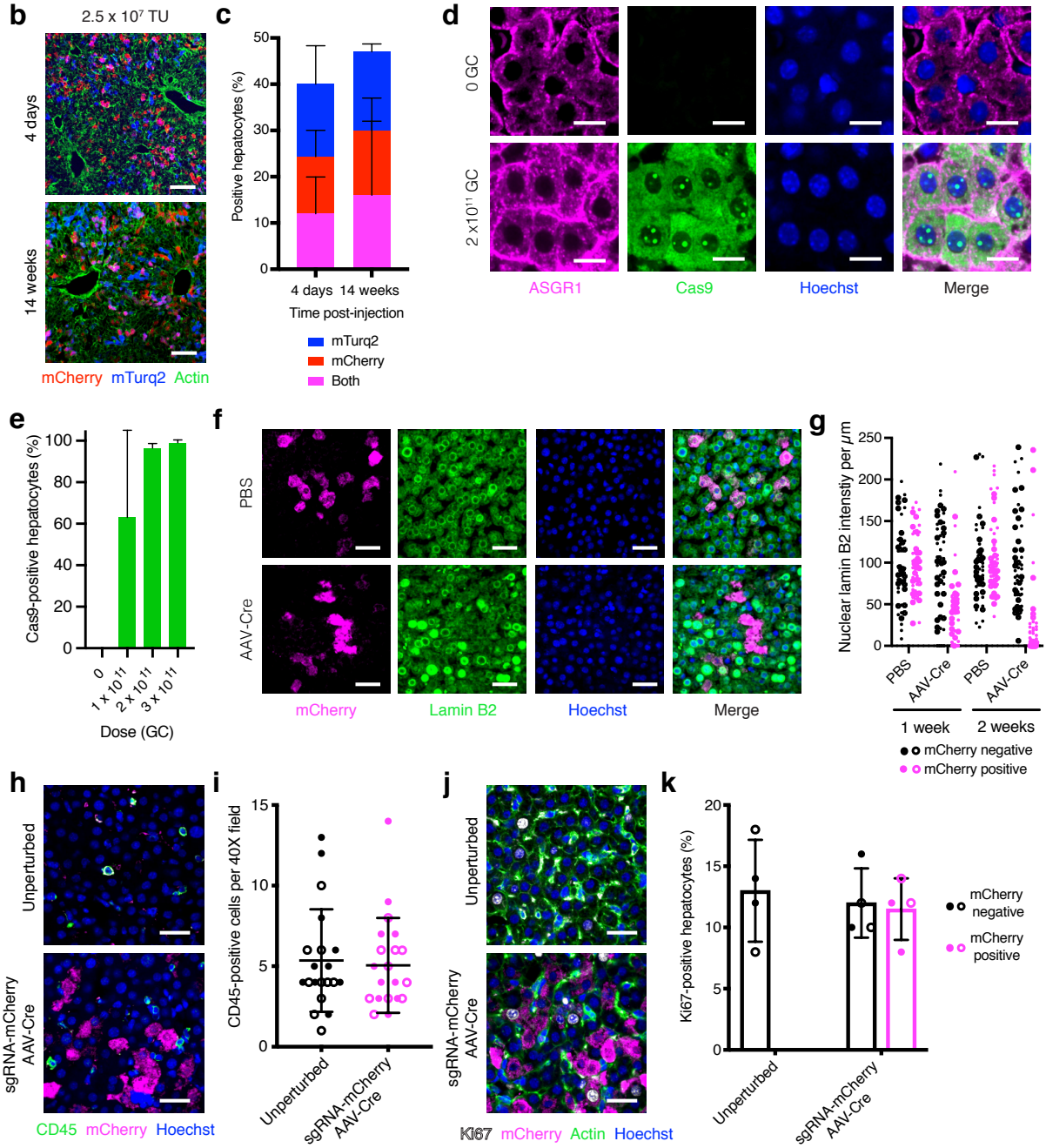
973 **a,** KEGG gene sets exhibiting significant depletion (FDR q-value < 0.05) at the endpoint
974 of the screen ranked by FDR q-value (-log10). Bars extending to the end of the plot
975 indicate an FDR q-value of 0. **b,** KEGG gene sets exhibiting significant depletion (FDR
976 q-value < 0.05) in our screen relative to screens in either mouse embryonic stem (ES)
977 cells (dark grey bars) or human hepatocellular carcinoma (HCC) cell lines (light grey
978 bars) ranked by FDR q-value (-log10). Bars extending to the end of the plot indicate an
979 FDR q-value of 0. **c,** Quantile-normalized median fold change (log2) for genes in the
980 KEGG gene set for glycosaminoglycan biosynthesis and heparan sulfate in ES cell
981 screens, HCC cell line screens, and our screen. Genes uniquely depleted in our screen
982 are highlighted in red. The bounds of the box indicate the first and third quartiles, and
983 the whiskers extend to the furthest data point that is within 1.5 times the interquartile
984 range. **d,** Quantile-normalized median fold change (log2) for genes in the KEGG gene
985 set for antigen processing and presentation in ES cell screens, HCC cell line screens,
986 and our screen. Genes uniquely depleted in our screen are highlighted in red. The
987 bounds of the box indicate the first and third quartiles, and the whiskers extend to the
988 furthest data point that is within 1.5 times the interquartile range. **e,** Pathway for
989 presentation of peptide antigens on class I MHC complexes. Created with
990 BioRender.com. **f,** Scheme for *in vivo* competition between hepatocytes expressing
991 sgAAVS1-mTurq2 and sgTap1-mCherry (top panel). Ratio (log2) of sgTap1-mCherry-
992 positive to sgAAVS1-mTurq2-positive hepatocytes in livers from mice at 0 or 3 weeks
993 after Cas9 induction without or with NK cell depletion (bottom panel). Closed and open
994 circles represent values from male and female mice, respectively. n = one male and one
995 female mouse per time point and >100 hepatocytes per mouse for 0 week time point

996 and >200 hepatocytes per mouse for 3 week time point. **p = 0.0031 (0 weeks versus 3
997 weeks without NK depletion) and *p = 0.026 (3 weeks without NK depletion versus 3
998 weeks with NK depletion) by one-tailed Fisher's exact test.

Extended Data Figure 1

a

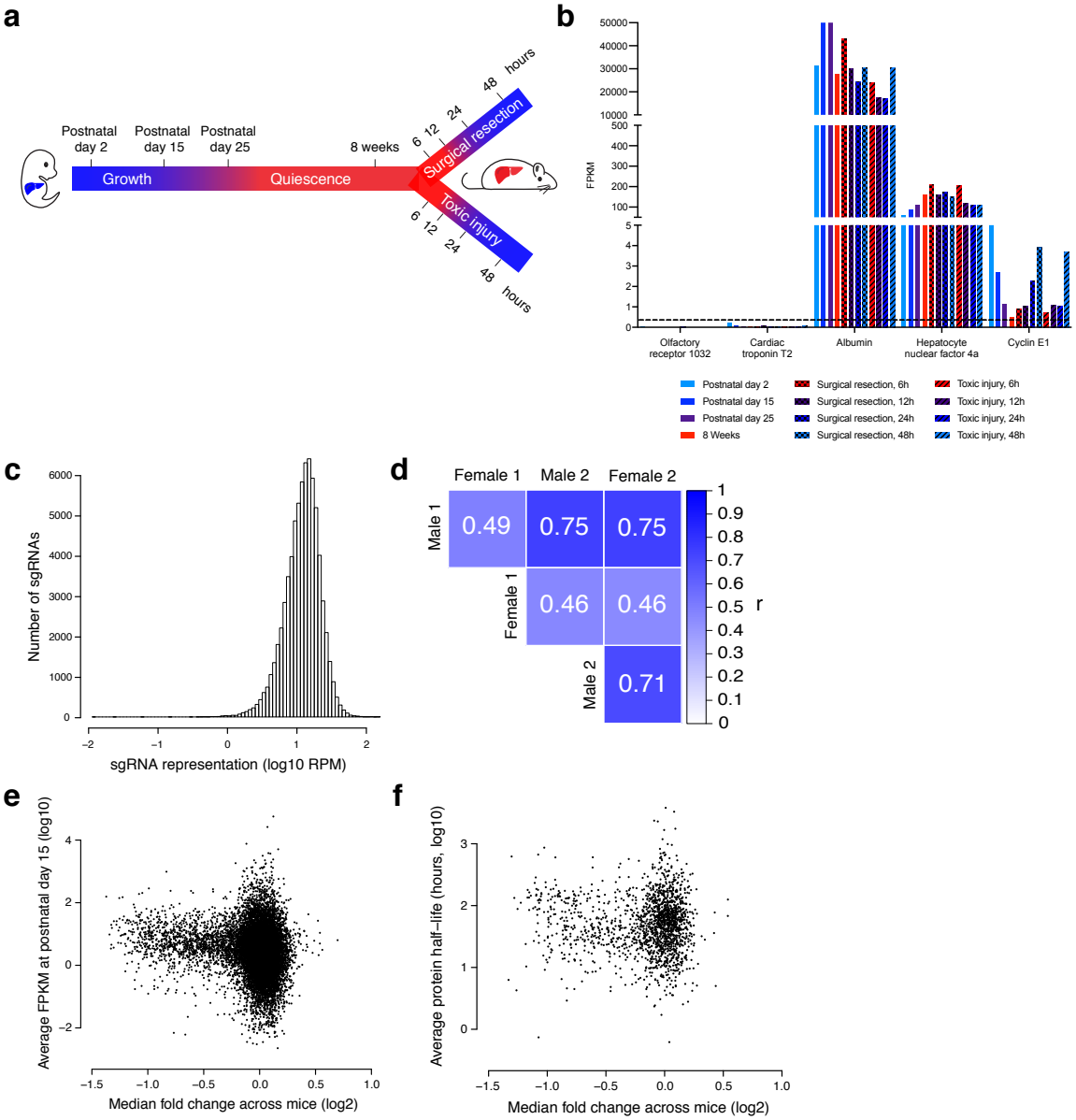
Litter	Sex	Volume (μL)	Mass (mg)	Proportion Hepatocytes	Average Hepatocyte Volume (μL)	Number of Hepatocytes
1	F	50	41.5	0.777041117	0.000002937	13,228,483.44
1	M	75	56	0.71753026	0.000003311	16,253,328.14
2	F	50	50.1	0.695385104	0.000002877	12,085,246.85
2	M	50	36.8	0.727514847	0.000002543	14,304,263.60



1000 **Extended Data Fig. 1 | Efficient delivery of sgRNAs and depletion of long-lived**
1001 **proteins in the liver. a,** Table of values used to estimate number of hepatocytes in
1002 postnatal day one livers. Liver volume was measured by volume displacement and
1003 percent hepatocytes and hepatocyte volume were measured by immunostaining and
1004 microscopy. **b,** Images of endogenous mCherry and mTurq2 fluorescence in livers from
1005 mice four days or 14 weeks after injection of 2.5 x 10^7 TU of an equal mixture of
1006 sgAAVS1-mCherry and sgAAVS1-mTurq2 lentiviruses. Scale bars, 100 μm. **c,** Percent
1007 mCherry-, mTurq2-, and double-positive hepatocytes in livers from mice four days or 14
1008 weeks after injection of 2.5 x 10^7 TU of an equal mixture of sgAAVS1-mCherry and
1009 sgAAVS1-mTurq2 lentiviruses. n = 3 mice per time point and 200 hepatocytes per
1010 mouse. Error bars indicate standard deviation. **d,** Images of livers from LSL-Cas9 mice
1011 injected with 0 or 2 x 10^{11} GC of AAV-Cre on postnatal day five and harvested four days
1012 thereafter immunostained for ASGR1 (hepatocyte marker, magenta) and Cas9 (green)
1013 and counterstained with Hoechst (blue). Scale bars, 15 μm. **e,** Percent Cas9-positive
1014 hepatocytes in livers from LSL-Cas9 mice injected with varying doses of AAV-Cre on
1015 postnatal day five and harvested four days thereafter as determined by Cas9
1016 immunostaining. n = 1 male and 1 female mouse per dose and 200 hepatocytes per
1017 mouse. Error bars indicate standard deviation. **f,** Images of livers from LSL-Cas9 mice
1018 injected with sgLmnb2-mCherry followed by PBS or AAV-Cre immunostained for
1019 mCherry (magenta) and lamin B2 (green) and counterstained with Hoechst (blue). Scale
1020 bars, 45 μm. **g,** Nuclear lamin B2 intensity per μm in mCherry-positive and mCherry-
1021 negative hepatocytes from LSL-Cas9 mice injected with sgLmnb2-mCherry followed by
1022 PBS or AAV-Cre. n = 1 male and 1 female mouse per condition and 25 cells per mouse.

1023 Closed and open circles represent values from male and female mouse, respectively. **h,**
1024 Images of livers from unperturbed postnatal day 12 LSL-Cas9 mice or postnatal day 12
1025 LSL-Cas9 mice injected with 1.25 x 10^7 TU of sgMaob-mCherry or sgLmnb2-mCherry
1026 lentivirus on postnatal day 1 followed by AAV-Cre on postnatal day 5 immunostained for
1027 CD45 (green) and mCherry (magenta) and counterstained for Hoechst (blue). Scale
1028 bars, 45 μm. **i,** Number of CD45-positive cells per 40X field in unperturbed postnatal
1029 day 12 LSL-Cas9 mice or postnatal day 12 LSL-Cas9 mice injected with 1.25 x 10^7 TU
1030 of sgMaob-mCherry or sgLmnb2-mCherry lentivirus on postnatal day 1 followed by
1031 AAV-Cre on postnatal day 5 as determined by CD45 immunostaining. n = 2 male and 2
1032 female mice per condition and five fields per mouse. Closed and open circles represent
1033 values from male and female mice, respectively. **j,** Images of livers from unperturbed
1034 postnatal day 12 LSL-Cas9 mice or postnatal day 12 LSL-Cas9 mice injected with 1.25
1035 x 10^7 TU of sgMaob-mCherry or sgLmnb2-mCherry lentivirus on postnatal day 1
1036 followed by AAV-Cre on postnatal day 5 immunostained for Ki67 (white), mCherry
1037 (magenta) and actin (green) and counterstained for Hoechst (blue). Scale bars, 45 μm.
1038 **k,** Percent Ki67-positive hepatocytes in livers from unperturbed postnatal day 12 LSL-
1039 Cas9 mice or postnatal day 12 LSL-Cas9 mice injected with 1.25 x 10^7 TU of sgMaob-
1040 mCherry or sgLmnb2-mCherry lentivirus on postnatal day 1 followed by AAV-Cre on
1041 postnatal day 5 as determined by Ki67 immunostaining. n = 2 male and 2 female mice
1042 per condition and 50 cells per mouse. Closed and open circles represent values from
1043 male and female mice, respectively.

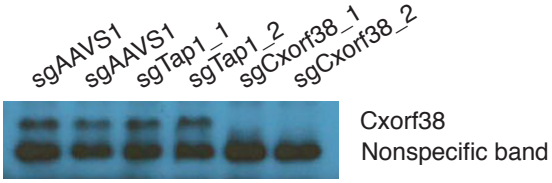
Extended Data Figure 2



Extended Data Fig. 2 | A genome-wide screen for hepatocyte fitness in the neonatal mouse liver. a, Time points during liver growth, quiescence, and regeneration from which livers were harvested for RNA sequencing. Partial hepatectomy and carbon tetrachloride were used as surgical resection and toxic injury models, respectively. **b,** Average RNA expression (FPKM) of representative protein-coding genes at different

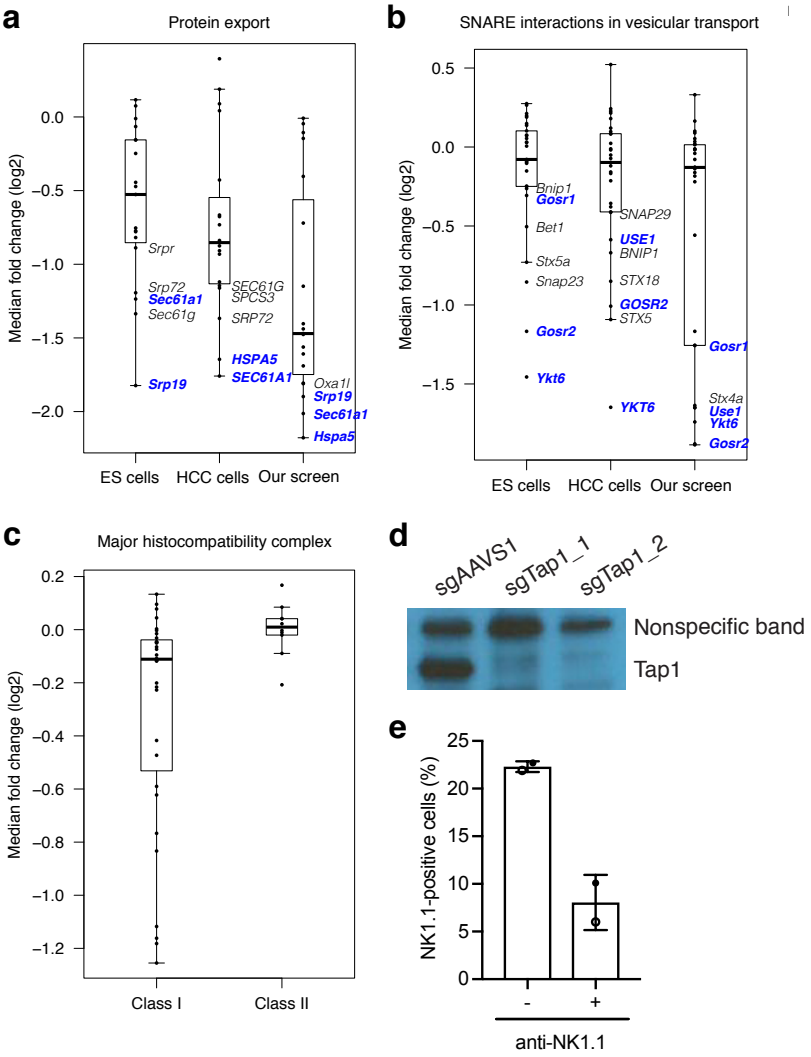
1050 time points of liver growth, quiescence, and regeneration. Dashed line indicates FPKM
1051 cutoff of 0.3. n = 3 male mice per time point **c,** Number of sgRNAs with a given
1052 representation (log10 RPM) for all sgRNAs in the library. **d,** Pearson correlation (r) for
1053 each plot in Fig. 2d. **e,** Average RNA expression at postnatal day 15 (log10 FPKM)
1054 relative to median fold change across mice (log2) for each gene. Spearman ρ = -0.15. **f,**
1055 Average protein half-life (log10 hours) relative to median fold change across mice (log2)
1056 for each gene. Spearman ρ = 0.01.

Extended Data Figure 3



Extended Data Fig. 3 | An *in vivo* screen uncovers sex-dependent fitness effects and canonical tumor suppressor genes. Western blot showing Cxorf38 protein levels in Cas9-AML12 cells stably transduced with sgAAVS1, sgTap1, or sgCxorf38. A nonspecific band is shown as a loading control.

Extended Data Figure 4



Extended Data Fig. 4 | Class I MHC is essential for hepatocyte fitness. a, Quantile-normalized median fold change (log2) of genes in the KEGG gene set for protein export in ES cell screens, HCC cell line screens, and our screen. Genes depleted in more than one screen are highlighted in blue. The bounds of the box indicate the first and third quartiles, and the whiskers extend to the furthest data point that is within 1.5 times the interquartile range. **b,** Quantile-normalized median fold change (log2) for genes in the KEGG gene set for SNARE interactions in vesicular transport in ES cell screens, HCC

1070 cell line screens, and our screen. Genes depleted in more than one screen are
1071 highlighted in blue. The bounds of the box indicate the first and third quartiles, and the
1072 whiskers extend to the furthest data point that is within 1.5 times the interquartile range.
1073 **c,** Median fold change (log2) across mice for genes associated with class I MHC or
1074 class II MHC within the KEGG gene set for antigen processing and presentation in our
1075 screen. The bounds of the box indicate the first and third quartiles, and the whiskers
1076 extend to the furthest data point that is within 1.5 times the interquartile range. **d,**
1077 Western blot showing Tap1 protein levels in Cas9-AML12 cells stably transduced with
1078 sgAAVS1 or sgTap1. A nonspecific band is shown as a loading control. **e**, Percent of
1079 NK1.1-positive NK cells among CD3-negative non-parenchymal cells in livers of mice
1080 without (-) or with (+) intraperitoneal injections of anti-NK1.1 every three days from
1081 postnatal day 5 to postnatal day 26.