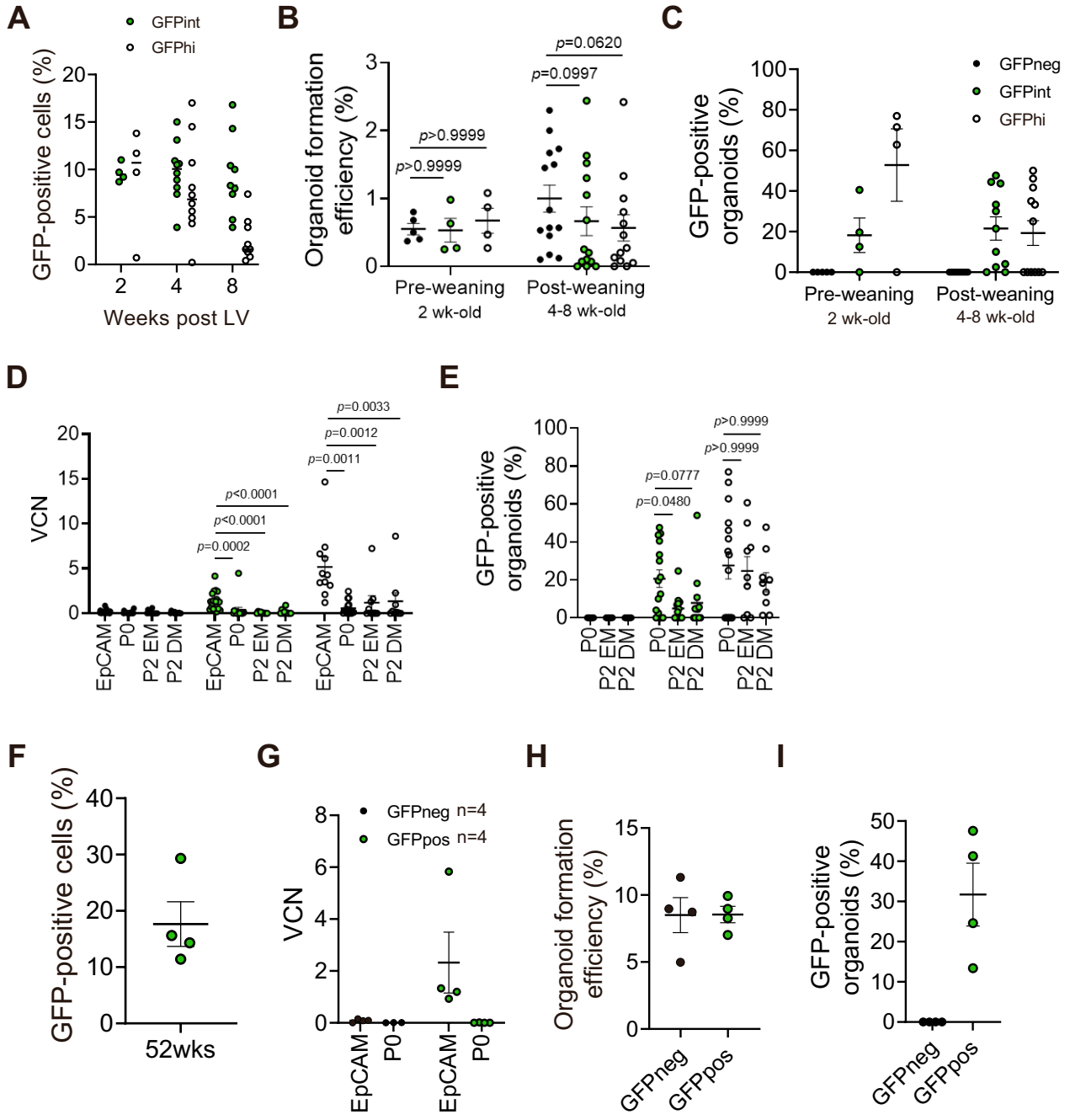
Cell Reports, Volume *44*

Supplemental information

Identification of hepatocyte-primed cholangiocytes in the homeostatic liver by *in vivo* lentiviral gene transfer to mice and non-human primates

Michela Milani, Francesco Starinieri, Anna Fabiano, Stefano Beretta, Tiziana Plati, Cesare Canepari, Mauro Biffi, Fabio Russo, Valeria Berno, Rossana Norata, Francesca Sanvito, Ivan Merelli, Luigi Aloia, Meritxell Huch, Luigi Naldini, and Alessio Cantore

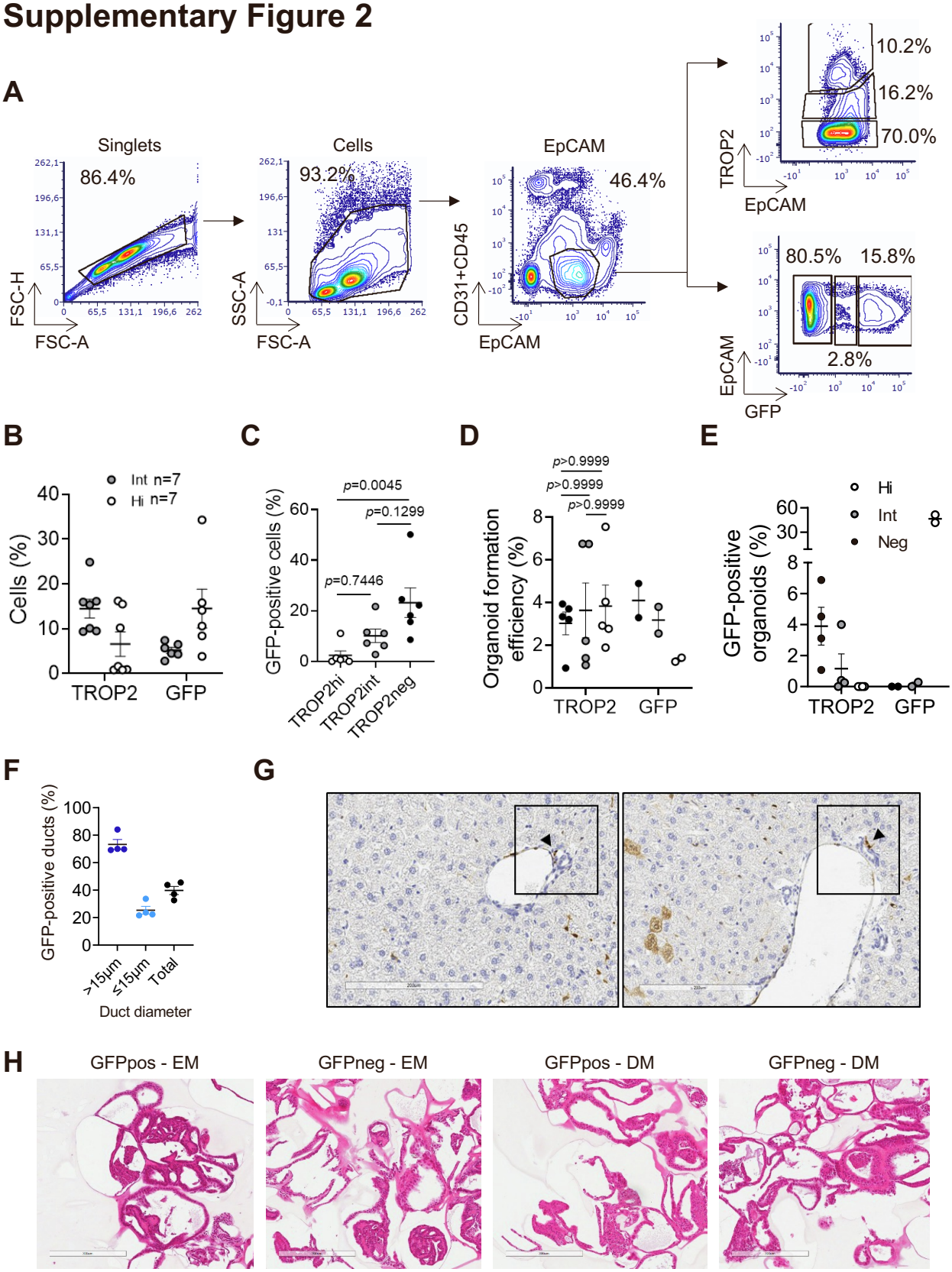
Supplementary Figure 1



Supplementary Figure 1: Mouse BEC express GFP under the control of a hepatocyte-specific promoter and GFP-positive BEC are maintained *in vivo* life-long. (A) Single values and mean with SEM of GFP-intermediate (GFPint) or GFP-high (GFPhi) EpCAM-positive cells at the indicated time-points after LV *in vivo* administration from mice treated as newborns with a LV encoding GFP under the control of the hepatocyte-specific enhanced transthyretin (ET) promoter ($5x10^{10}$ TU/Kg). 2wks n=4, 4wks n=11, 8wks n=9. Pool of 3 independent experiments. **(B, C)** Single values and mean with SEM of the organoid formation efficiency (B) or the percentage of GFP-positive organoids (C)

of GFP-negative (GFPneg), GFPint or GFPhi EpCAM-positive cells according to the age of mice at the moment of liver collection. Pre-weaning: GFPneg n=5, GFPint n=4, GFPhi n=4; post-weaning: GFPneg n=14, GFPint n=14, GFPhi n=13. Kruskal-Wallis test with Dunn's multiple comparisons test. **(D)** Single values and mean with SEM of VCN in GFPneg (n=19), GFPint (n=15) or GFPhi (n=15) EpCAM-positive sorted cells, organoids (P0) or organoids kept in culture in expansion medium (EM) or differentiation medium (DM) after two passages (P2). Kruskal-Wallis test with Dunn's multiple comparisons test. **(E)** Single values and mean with SEM of GFP-positive organoids shown in (D). Kruskal-Wallis test with Dunn's multiple comparisons test. **(F)** Single values and mean with SEM of GFP-positive EpCAM-positive cells 1 year after LV *in vivo* administration to newborn mice ($5x10^{10}$ TU/Kg, n=4). **(G)** Single values and mean with SEM of VCN in GFPneg or GFP-positive (GFPpos) sorted EpCAM-positive cells or organoids (P0). **(H, I)** Single values and mean with SEM of the organoid formation efficiency (H) or the percentage of GFP-positive organoids (I) of GFPneg or GFPpos EpCAM-positive cells.

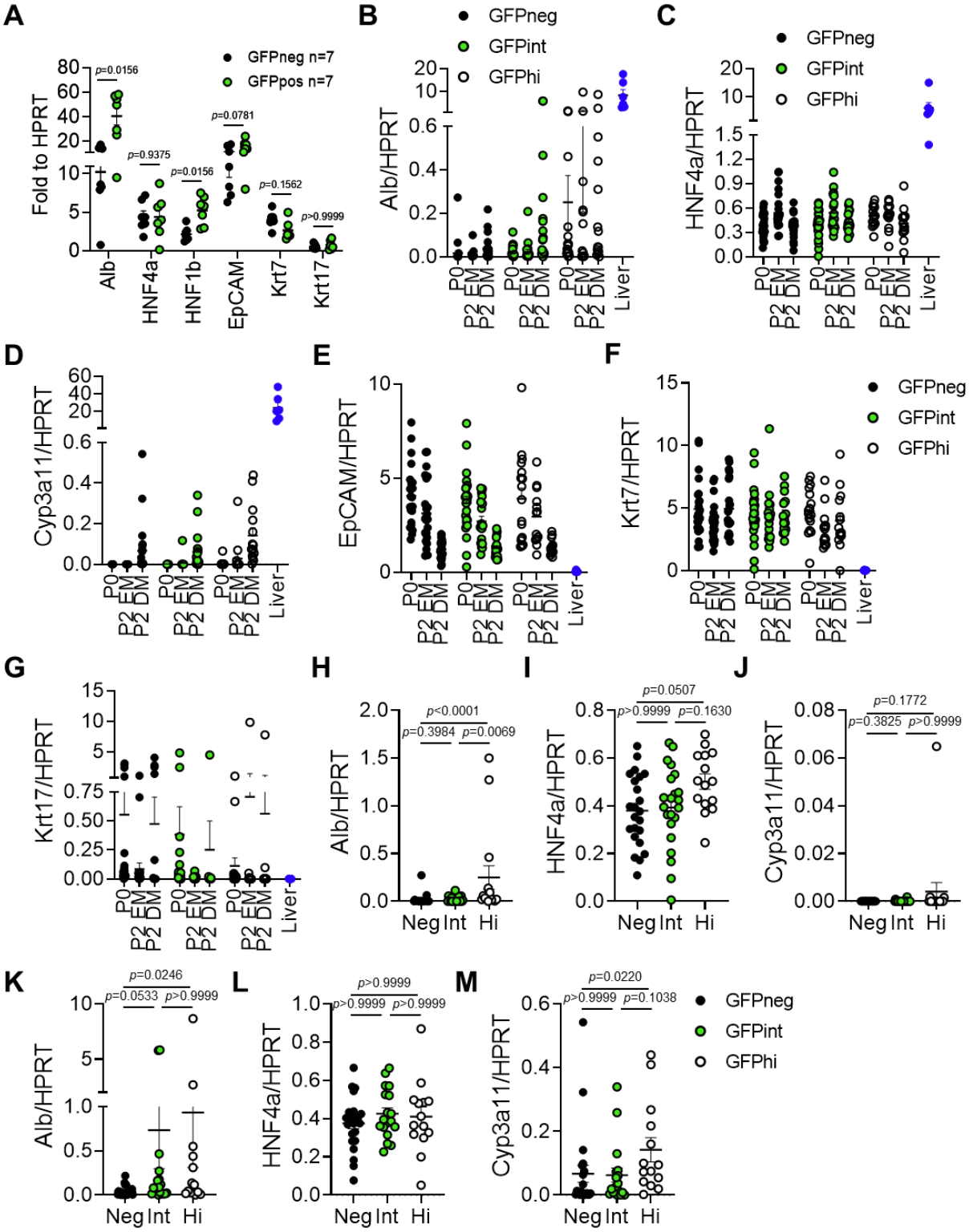
Supplementary Figure 2



Supplementary Figure 2: Heterogeneity assessment of mouse cholangiocytes. (A) Gating strategy for sorting of cholangiocytes (EpCAM-positive, CD31- and CD45-negative) according to level of expression of TROP2 (n=5) or GFP (n=2). **(B)** Single values and mean with SEM of TROP2-positive

or GFP-positive EpCAM-positive cells 8wks after LV *in vivo* administration to newborn mice ($5x10^{10}$ TU/Kg, n=7). Int: intermediate; hi: high. **(C)** Single values and mean with SEM of GFP-positive EpCAM-positive cells according to TROP2 expression levels. Friedman test with Dunn's multiple comparisons test. **(D, E)** Single values and mean with SEM of the organoid formation efficiency (D) or the percentage of GFP-positive organoids (E) of TROP2neg, TROP2int or TROP2hi (n=5) EpCAM positive cells, or GFPneg, GFPint or GFPhi (n=2) EpCAM-positive cells. **(F)** Single values and mean with SEM of the percentage of bile ducts containing at least one GFP-positive cell (GFP-positive ducts) according to duct diameter (n=4). **(G)** Representative images of GFP-positive ducts (black arrow). Note that not all cholangiocytes in a GFP-positive duct are GFP-positive. Scale bar: 200 µm. **(H)** Representative images of hematoxylin/eosin stained sections of liver organoids cultured in EM or DM derived from GFP-positive (GFPpos) or GFP-negative (GFPneg) BEC, as indicated. Scale bar: 300 µm.

Supplementary Figure 3

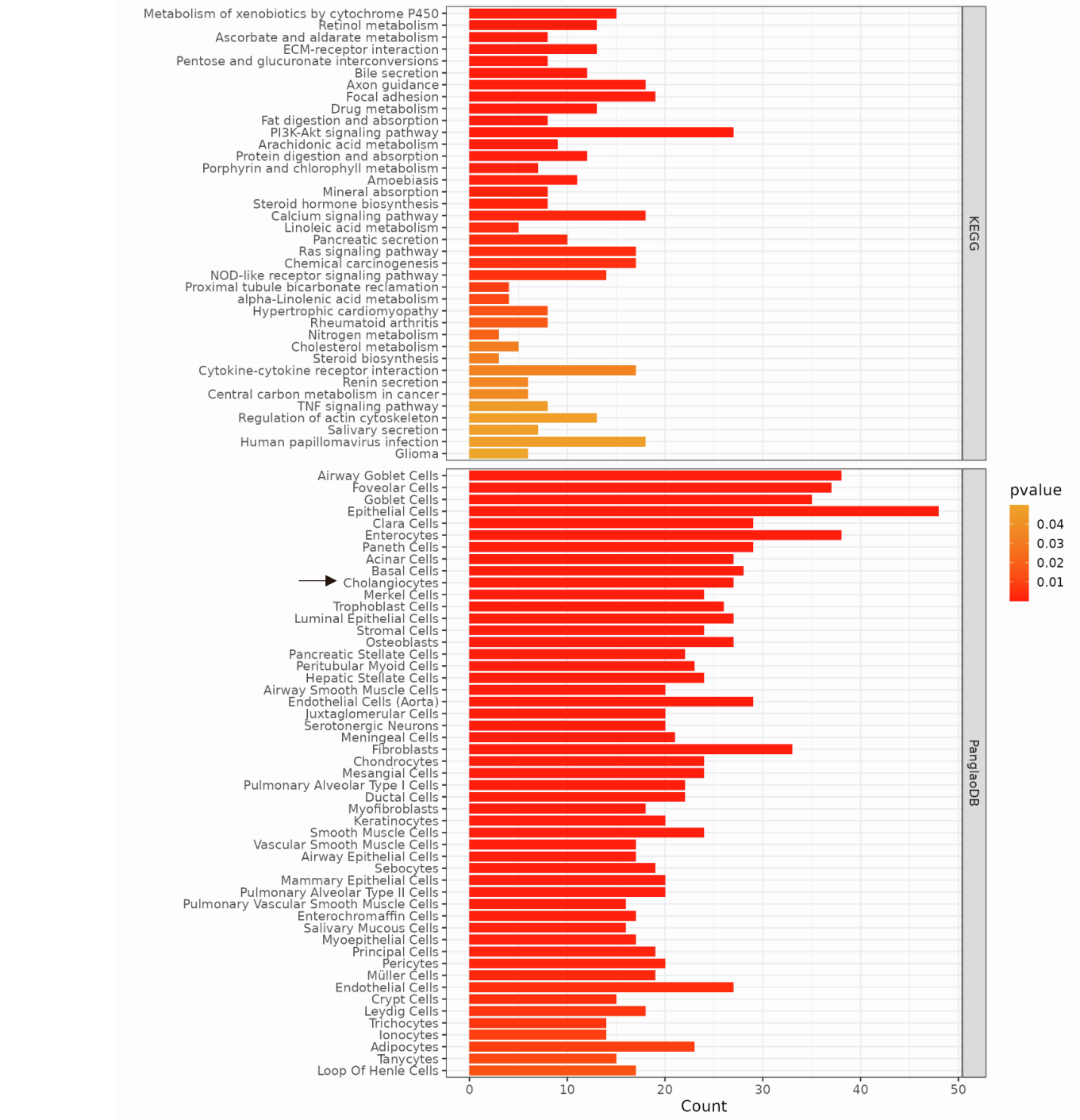


Supplementary Figure 3: Gene expression analyses reveal increased hepatocyte-specific genes in organoids derived from GFP-positive cholangiocytes. (A) Single values and mean with SEM of the expression of the indicated genes in sorted GFP-positive or –negative EpCAM-positive BEC isolated from 8wk-old mice (n=7) treated as newborns with i.v. administration of LV expressing GFP

under the control of the ubiquitous PGK promoter. **(B-G)** Single values and mean with SEM of the expression of albumin (B), HNF4a (C), Cyp3a11 (D), EpCAM (E), Krt17 (F) or Krt7 (G) normalized on HPRT in organoids cultured in EM or DM at the indicated time-points derived from GFPneg (n=24), GFPint (n=22) or GFPhi (n=16) EpCAM-positive cells, as indicated, isolated from 4-8wk-old mice treated as newborns with i.v. administration of LV expressing GFP under the control of the ubiquitous PGK promoter or the hepatocyte-specific ET promoter. Expression of these genes in murine livers is reported for comparison (Liver, n=6). **(H-M)** Single values and mean with SEM of the expression of albumin (H, K), HNF4a (I, L) or Cyp3a11 (J, M) of organoids shown in (B-G) at P0 (H-J) or after culture in DM (K-M) reported here to show statistical analyses. Kruskal-Wallis test with Dunn's multiple comparisons test.

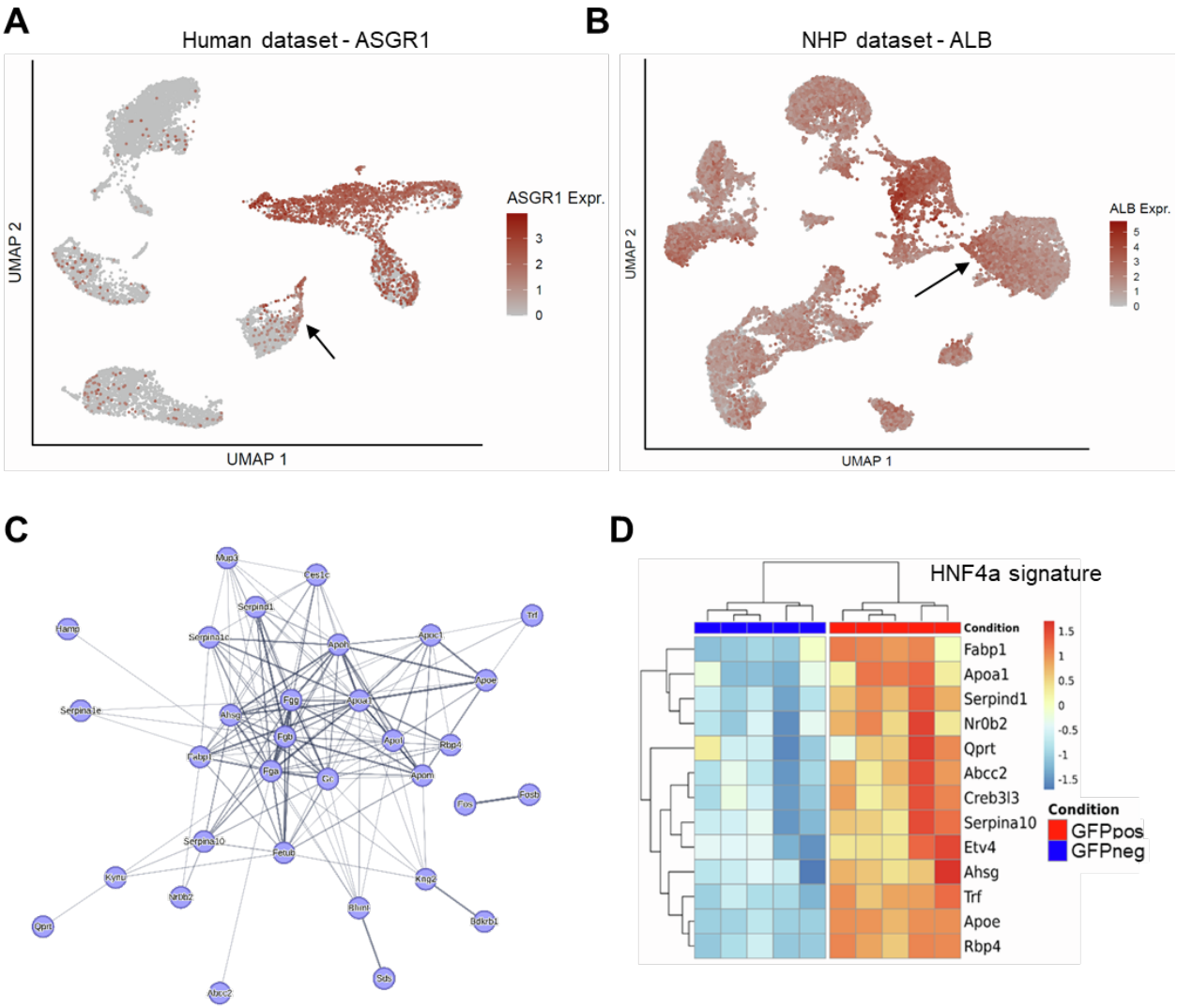
Supplementary Figure 4

A



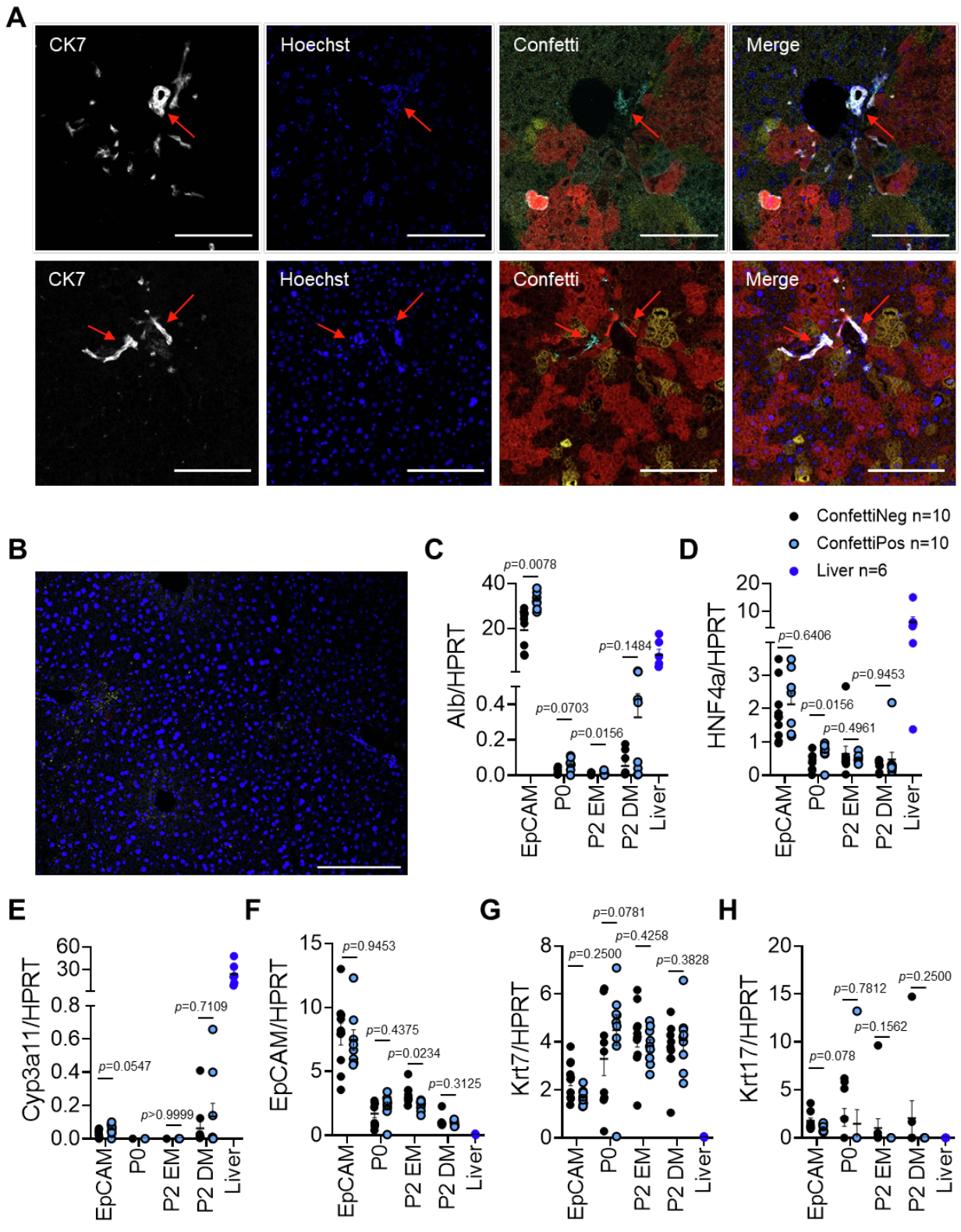
Supplementary Figure 4: Transcriptomic analysis reveal a subpopulation of cholangiocytes primed toward hepatocyte lineage in mice. (**A**) Barplot of enriched categories (adjusted p<0.05) on DEGs downregulated in GFPpos compared to GFPneg cells. Counts on the X axis represent the number of enriched genes in the corresponding term, while bar colors represent the statistical significance (p-value).

Supplementary Figure 5



Supplementary Figure 5: Transcriptomic analyses reveal a subpopulation of cholangiocytes primed toward hepatocyte lineage. (A) UMAP plot showing expression of the hepatocyte marker asialoglycoprotein Receptor 1 (ASGR1) in the human dataset shown in Figure 3D. **(B)** UMAP plots showing the expression of *ALB* in the NHP scRNAseq dataset are shown in Figure 3F. **(C)** Gene network obtained from the StringDB analysis on upregulated DEGs in GFPpos compared to GFPneg cells. **(D)** Heatmap showing the expression of DEGs in the core enrichment of the HNF4a category from the ChEA transcription factors database.

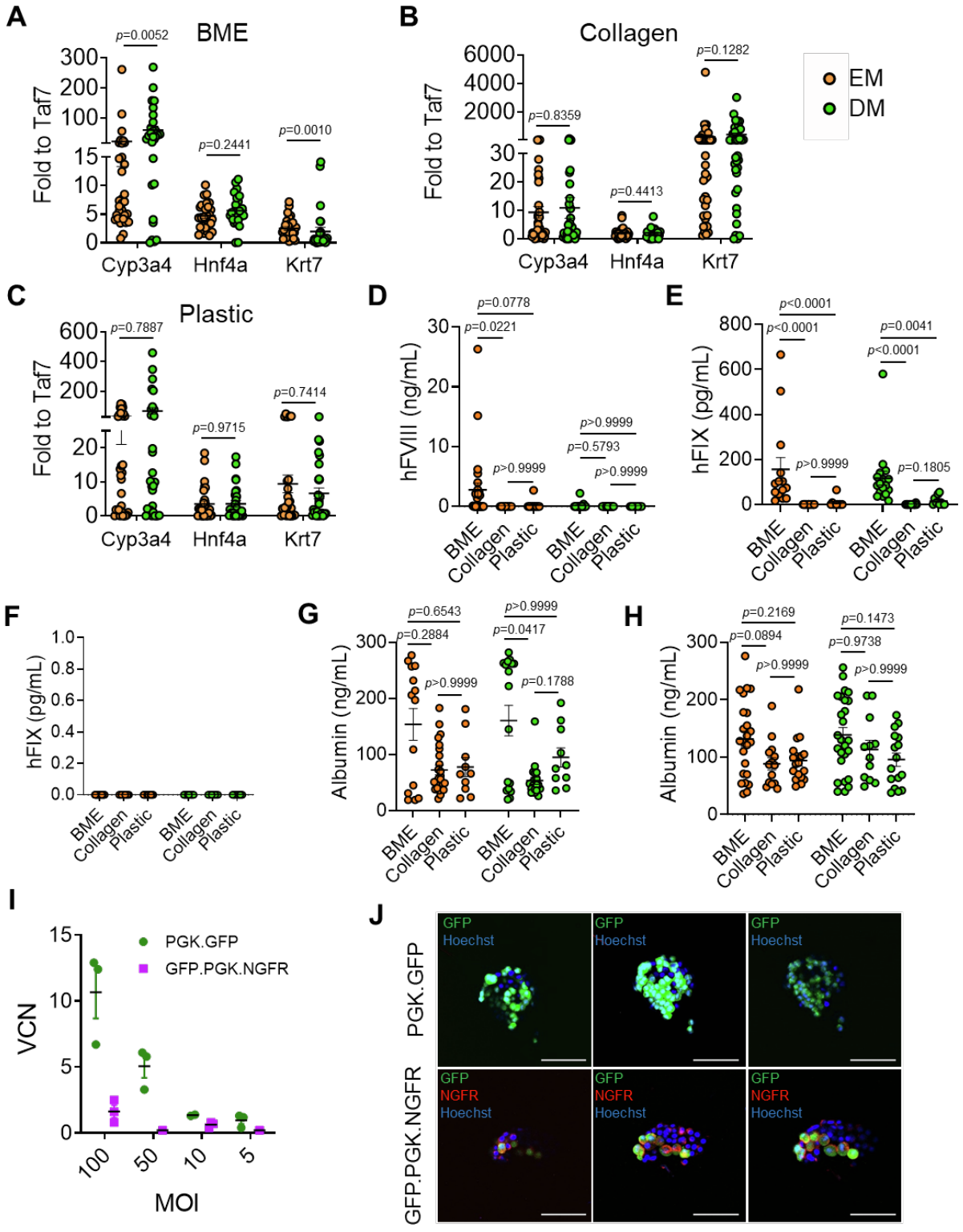
Supplementary Figure 6



Supplementary Figure 6: Gene expression analyses of organoids derived from confetti-positive or -negative cholangiocytes. (A) Representative images of confetti-positive (cyan-positive or red-positive, red arrows) cholangiocytes (CK7-positive, in white) in the liver of 2 different 3wk-old Alb-CreERT2/Confetti mice in which Cre-mediated recombination of the Confetti construct has been

induced with subcutaneous administration of tamoxifen when newborn (see Figure 4A). Hoechst: nuclei. Scale bar: 100 μm. **(B)** Representative image of the liver of an Alb.CreERT2/R26-Confetti mouse not treated with Tamoxifen. Blue: Hoechst. Scale bar: 200 μm. **(C-H)** Single values and mean with SEM of the expression of albumin (C), HNF4a (D), Cyp3a11 (E), EpCAM (F), Krt17 (G) or Krt7 (H) normalized on HPRT in EpCAM-positive cells or organoids cultured in EM or DM at the indicated time-points derived from ConfettiNeg (n=10) or ConfettiPos (n=10) EpCAM-positive cells, as indicated, isolated from 8wk-old Alb-CreERT2/Confetti mice treated as newborns with subcutanous administration of tamoxifen to activate the inducible CreERT2 (see Figure 3A). Expression of these genes in murine livers is reported for comparison (Liver, n=6, same data shown in Figure S3B-G). Wilcoxon matched-pairs signed rank test.

Supplementary Figure 7



Supplementary Figure 7: *Ex vivo* LV transduction and gene expression of NHP single organoids. (A-C) Single values and mean with SEM of the relative expression of the indicated genes in single liver organoids grown on basal membrane matrix extract (BME-, A), plastic (B) or collagen- (C) coated plates. **(D-H)** Single values and mean with SEM of the concentration of hFVIII (D), hFIX

(E, F) or NHP-albumin (G, H) measured in the supernatant of single organoids cultured on matrix with different stiffness, as indicated. Single organoids were derived from untreated NHP (F, H), LV.FVIII or LV.FVIII-XTEN treated NHP (D) or LV.FIX treated NHP (E, G). **(I)** Single values and mean with SEM of VCN measured in NHP organoids after LV transduction with the indicated LV at the indicated multiplicity of infection (MOI). **(J)** Representative immunofluorescence images of organoids transduced with the indicated LV. Scale bar: 100 μm.