TFEB regulates murine liver cell fate during development and regeneration

TFEB influences cell fate in developing and adult liver

Nunzia Pastore[1,2*], Tuong Huynh[1,2], Niculin J. Herz[1,2], Alessia Calcagni`[1,2], Tiemo J. Klisch[1,2], Lorenzo Brunetti[3,4], Kangho Kim[5], Marco De Giorgi[6], Ayrea Hurley[6], Annamaria Carissimo[7], Margherita Mutarelli[7], Niya Aleksieva[8], Luca D'Orsi[7], William R. Lagor[6], David D. Moore[5], Carmine Settembre[7,10], Milton J. Finegold[9], Stuart J. Forbes[8] and Andrea Ballabio[1,2,7,10*]

[1]*Jan and Dan Duncan Neurological Research Institute, Texas Children Hospital, Houston, Texas, 77030, USA*

[2]*Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, 77030, USA*

[3]*Stem Cells and Regenerative Medicine Center, Baylor College of Medicine, Houston, TX 77030, USA*

[4]*Center for Cell and Gene Therapy, Baylor College of Medicine, Houston, TX 77030, USA*

[5]*Department of Molecular and Cell Biology, Baylor College of Medicine, Houston, Texas, 77030, USA*

[6]*Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, Texas, 77030, USA*

[7]*Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli (NA), 80078, Italy*

[8]*MRC Centre for Regenerative Medicine University of Edinburgh, Edinburgh EH16 4UU, UK*

[9]*Department of Pathology, Baylor College of Medicine, Houston, TX, 77030, USA*

[10]*Department of Translational Medicine, Medical Genetics, Federico II University, Naples, 80131, Italy*

**Correspondence should be addressed to:*
Nunzia Pastore (pastore@bcm.edu) and Andrea Ballabio (ballabio@tigem.it)

Transcription Factor EB (TFEB), a master regulator of lysosomal biogenesis and autophagy, regulates liver glucose and lipid metabolism in response to nutrients. Here we show that TFEB also plays an important role in cell lineage specification and proliferation both during development and upon liver regeneration. *In vitro* and *in vivo* gain- and loss-of function studies revealed that TFEB drives the differentiation status of liver progenitor cells (LPCs) into the progenitor/cholangiocyte lineage while inhibiting differentiation into the hepatocyte lineage. Accordingly, TFEB liver-specific overexpressing mice showed cells with hybrid features of hepatocytes and LPCs, ectopic bile duct formation, cholestasis and bile duct neoplasms. On the contrary, TFEB conditional KO mice in the liver exhibited an aberrant response to hepatic injury characterized by depressed LPC expansion. Genetic interaction studies showed that *Sox9,* marker of precursor and biliary cells, is a direct TFEB target and a primary mediator of its effects on liver cell fate. In summary, our findings reveal a previously unrecognized role for TFEB in controlling cell lineage commitment and identify a novel

unexplored pathway that is involved in liver homeostasis and may play a role in biliary cancer.

Introduction

The adult liver is the largest internal organ and provides many essential metabolic, exocrine, and endocrine functions[1]. Being a complex organ with several cell types, liver development involves multiple cell fate decisions. For instance, during development hepatic endoderm cells, known as hepatoblasts (HBs), differentiate into hepatocytes or cholangiocytes depending on their localization with respect to the portal vein: HBs exposed to ligands from portal venous endothelial cells differentiate into primitive ductal plate cells and then form bile ducts, whereas HBs located further away from the portal vein differentiate into hepatocytes[2,3]. Hepatocytes in the adult liver rarely divide, however, under conditions in which alterations of liver mass occurs, such as surgical removal or cell loss caused by drugs or viruses, quiescent hepatocytes become proliferative and replicate to restore full liver functional capacity[4]. In certain injury models, hepatocytes can transdifferentiate into ductal biliary epithelial cells (BECs) to ensure tissue regeneration[5-9]. However, when hepatocyte proliferation is compromised, BECs become active and subsequently differentiate into hepatocytes[10,11]. Liver stem/progenitor cells (LPCs) may appear in chronic liver damage when hepatocyte proliferation is compromised and differentiate in both hepatocytes and bile ducts[12]. In humans, LPCs are evident in pathological ductular reactions often observed in a variety of liver diseases, including fatty liver diseases[2], chronic viral hepatitis[13], cirrhosis[14] and acute hepatic injury[15]. Thus, LPCs are considered potential targets for liver cell transplantation and repopulation[16,17]. However, although LPCs have capacity to differentiate into hepatocytes and biliary cells *in vitro* and to form hepatocytes buds repopulating the liver *in vivo*[18], their ability to participate to liver regeneration in human clinical setting is still unclear[19,20]. A specialized type of cells with both LPCs and mature hepatocytes features, called hybrid hepatocytes (HybHPs), express both SRY (sex determining region Y)-box (SOX9), a marker of LPCs, and the hepatocyte marker hepatocyte nuclear factor 4α (HNF4α) [21]. HybHPs are located at the periportal region of normal liver and are efficient in liver repair when non-centrilobular hepatocytes are damaged. Thus, it appears that liver injury triggers several regenerative responses depending on the size and the proliferative capacity of the remaining liver tissue[22].

The mechanisms regulating liver cell proliferation and differentiation are highly controlled to achieve accurate tissue growth and development, and deregulation of the signaling pathways involved in liver cell differentiation can impair regeneration and trigger the development of tumors with hepato-cellular and cholangio-cellular differentiation features[23,24]. Thus, a better understanding of the processes that control liver cell differentiation in physiological and

pathological conditions may contribute to the identification of novel and druggable targets and represent a potential therapeutic approach for the treatment of liver diseases.

Transcription factor EB (TFEB) belongs to the MiT-TFE family of transcription factors, which also includes MITF, TFE3, and TFEC[25]. In the last decade, several studies have explored the role of TFEB in physiological settings and in response to environmental cues. TFEB was first described as essential for placental vascularization[26]. Subsequently, TFEB has been implicated in the expression of lysosomal and autophagic genes[27,28]. TFEB is phosphorylated by the mTOR kinase and participates in a lysosomal signaling mechanism that is essential for nutrient sensing and maintenance of cellular homeostasis and energy metabolism[29-31]. In addition, TFEB has an important role in the regulation of body metabolism in liver and muscle and in the adaptation to environmental cues (e.g. fasting, high-fat diet, exercise)[32,33]. Furthermore, overexpression of MiT-TFE genes has been implicated in the pathogenesis of a variety of tumors, including renal cell carcinoma, melanoma, and pancreatic cancer[34-37], suggesting their involvement in cell differentiation and proliferation. However, a role of TFEB in cell fate determination and liver cancer has not been investigated so far.

Here, we show that TFEB plays a critical role in controlling cell fate and proliferation in the mammalian liver during embryogenesis and repair. Indeed, we found that TFEB is highly expressed in LPCs and BECs and that genetic manipulation of TFEB expression in the liver alters cell lineage specification in both developing and injured adult liver. Our data identify *Sox9* as one important downstream target of TFEB in hepatic cell differentiation, highlighting the importance of this pathway in liver development, regeneration and cancer.

Results

TFEB is highly expressed in the biliary compartment

To investigate whether TFEB plays a role in liver development, we examined its expression during liver specification. We used a mouse line in which the endogenous *Tcfeb* allele is disrupted by homologous recombination at exons 4 and 5 through the insertion of the β-Galactosidase coding sequence[29] (**Supplementary Fig. 1a**). Since *Tcfeb*^{KO} mice are embryonic lethal[26,29], heterozygous embryos (*Tcfeb*$^{LacZ/+}$) were collected at several embryonic stages to analyze *Tcfeb* promoter activity. *Tcfeb* expression was detected in the liver starting at embryonic stage E14.5 and increases overtime, just prior to the developmental differentiation of HBs into hepatocytes or biliary epithelial cells (BECs)[1], while no expression was detected at E12.5 (**Supplementary Fig. 1b**). Consistent with these results, the expression of the endogenous *Tcfeb* mRNA and protein expression levels in fetal and neonatal livers showed a gradual increase during liver growth (**Supplementary Fig. 1c and**

1d). Immunostaining analysis revealed a dynamic expression pattern of TFEB during liver development. At E15.5 TFEB was diffuse in the entire parenchyma (**Fig. 1a**). At postnatal stage P0 and adult stage, a stronger expression was detected in the portal vein endothelium (**Fig. 1a and 1b**), while pericentral hepatocytes (adjacent to the central vein) showed low and diffuse TFEB expression (**Fig. 1b**). Indeed, TFEB was mainly detected in HNF4α^{-}/SOX9^{+} ductal plate cells and mature ducts (**Fig. 1b and 1c**). X-Galactosidase staining in *Tcfeb*$^{LacZ/+}$ mice showed the same pattern of expression (**Supplementary Fig. 1e**). Together, our results show that TFEB expression is highly enriched in progenitor/ductal cells, while mature hepatocytes display lower TFEB levels, suggesting a possible role for TFEB in liver cell fate specification.

TFEB influences liver cell differentiation *in vitro*

To investigate whether TFEB plays a role in liver cell commitment, we generated CRISPR/Cas9 *Tcfeb*^{KO} (TFEB^{KO})(**Supplementary Fig. 2a**) and TFEB overexpressing (TFEB^{OE}) mouse HBs. To generate TFEB^{OE} HBs, we isolated HBs from *Tcfeb* conditional overexpressing mice that carry *Tcfeb*-3xFlag$^{fs/fs}$ under the control of a strong CMV early enhancer/chicken beta-actin (CAG) promoter (**Supplementary Fig. 2b**)[28,35]. To induce *Tcfeb* overexpression, we infected HBs with an HDAd-BOS-CRE virus. HBs are bi-potent cells that can differentiate into hepatocytes or cholangiocytes in culture when plated on uncoated or matrigel coated plates, respectively[38]. Consistently, control HBs differentiated into hepatocytes forming hepatocyte clusters (**Fig. 2a**) and showed induction of the expression of hepatocyte-specific genes such as *Alb*, *Hnf4α*, *AldoB* and *Otc*, with concomitant reduction of the expression of the precursor markers *Sox9*, *Sox4*, *Afp* and *Cd24* (**Fig. 2b**). Notably, HBs lacking TFEB showed significantly increased expression of the hepatocyte-specific markers and reduced *Sox9* levels compared to controls, with no differences in the expression of the precursor markers *Afp*, *Sox4 and Cd24* (**Fig. 2b**), suggesting that TFEB loss-of-function preferentially induces the hepatocyte differentiation program. On the contrary, TFEB overexpressing HBs did not completely differentiate into hepatocytes, as demonstrated by the smaller size of the hepatocyte-like aggregates (**Fig. 2a**), lower expression levels of hepatocyte-specific genes and higher levels of the precursor-specific markers (**Fig. 2b**), suggesting that TFEB overexpression prevents the hepatic differentiation of HBs, while maintaining precursor features. These results were confirmed by immunoblot and immunofluorescence analysis on HBs 3 days after hepatocytic differentiation (**Fig. 2c and 2d**). To confirm the defective differentiation program, we examined the effect of TFEB overexpression and depletion on the transcriptome of HBs 3 days after differentiation toward the hepatocytic lineage. Transcriptional profiles of TFEB^{KO} and TFEB^{OE} HBs confirmed the altered expression of hepatocyte- and progenitor/cholangiocyte- specific genes (**Fig. 2e**).

KEGG analysis on the differentially expressed genes showed upregulation of hepatocyte-specific pathways, such as drug metabolism- cytochrome P450, fat digestion and absorption and cholesterol metabolism in $TFEB^{KO}$ HBs that were instead downregulated in $TFEB^{OE}$ HBs (**Supplementary Tables 1 and 2**). Moreover, Gene set enrichment analysis (GSEA) demonstrated that HNF4α targets (a master transcriptional regulator of hepatocyte differentiation) are enriched among up-regulated genes in $TFEB^{KO}$ HBs and show a reduced expression, despite not significant, in $TFEB^{OE}$ HBs (**Supplementary Fig. 2c**). We also examined the distribution of HBs in the three phases of the cell cycle (G1 *vs* S *vs* G2/M). $TFEB^{KO}$ HBs showed significant increase in the percentage of S phase cells compared to controls, suggesting that TFEB depletion induces S phase arrest in HBs (**Fig. 2f**). Interestingly, TFEB overexpression in HBs resulted in increased percentage of cells in G2/M phase compared with control cells with a concomitant reduction in S phase (**Fig. 2f**), suggesting increased proliferation.

In contrast to the complete inhibition of hepatocyte specification, neither TFEB depletion nor overexpression impaired biliary differentiation of HBs. Indeed, $TFEB^{KO}$ and $TFEB^{OE}$ cells supported tubule formation (**Supplementary Fig. 2d**) and expressed biliary genes (i.e. *Hif1α, Hnf6*, *Ggt1*) after cholangiocytic differentiation despite the reduced or increased expression levels of both *Sox9* and *Afp*, respectively (**Supplementary Fig. 2e**).

Together these results indicate that TFEB plays a role in the lineage commitment of liver precursor cells.

TFEB overexpression induces a progenitor/biliary phenotype *in vivo*

To test whether TFEB overexpression alters cell differentiation from a progenitor state and impairs homeostasis of mature hepatocytes *in vivo*, we crossed the *Tcfeb*-3xFlag$^{fs/fs}$ mouse line (**Supplementary Fig. 2b**) with a transgenic line carrying *Alb-Cre* recombinase to obtain *Tcfeb*-3xFlag$^{fs/fs;}$*Alb-Cre* mice (hereafter referred to as Tg). *Albumin* is expressed by the bipotential HB progenitors, thus enabling us to investigate the role of TFEB during liver cell specification. tdTomato expression and *in situ* hybridization analysis confirmed TFEB overexpression in hepatocytes and BECs at E18.5, P0 and P9 (**Supplementary Fig. 3a and 3b**). Hepatoblast-specific expression of TFEB was assessed by qPCR analysis on liver extracts of Tg mice showing an approximately 10-fold increase of *Tcfeb* mRNA levels in E18.5, P0 and P9 livers and up to 65-fold in 3-month-old mice (**Supplementary Fig. 3c**), consistent with the progressive increase in *Alb* mRNA during liver specification[39].

To examine the effects of TFEB overexpression *in vivo*, we carried out microarray analysis at an early stage (P9) showing a total of 8,400 differentially expressed genes. KEGG analysis revealed that several up-regulated genes are involved in oxidative phosphorylation,

proteasome, cell cycle, Hippo signaling pathway and endocytosis, among others, while down-regulated genes are mostly involved in hepatocyte-specific pathways, such as amino acid, cholesterol, and lipid metabolism, drug metabolism (P450), PPAR signaling pathway (**Fig. 3a and Supplementary Table 3**). Consistent with *in vitro* data, gene expression profile of Tg livers demonstrated a reduction in the expression of hepatocyte-specific genes and an increase in progenitor/cholangiocyte related genes compared with control livers (**Fig. 3b and 3c and Supplementary Table 4**). Immunoblotting analysis confirmed the reduction of HNF4α and the increase of SOX9 protein levels in liver extracts from Tg mice at P9 (**Fig. 3d**). Similar results were obtained in primary hepatocytes isolated from 21-day-old Tg and control mice (**Supplementary Fig. 4a and 4b**). Interestingly, while no $CK19^+$ cells were detected in control hepatocytes, we found $HNF4\alpha^+/CK19^+$ Tg cells, suggesting a hybrid phenotype of Tg hepatocytes (**Supplementary Fig. 4c**). Indeed, immunostaining for the cholangiocyte marker CK19 in Tg livers from P0 mice showed $CK19^+$ cells with hepatocyte-like morphology positive for the hepatocyte marker HNF4α mainly in the periportal area (**Fig. 3e and Supplementary Fig. 5a**). The concomitant positive signals for both a cholangiocyte and a hepatocyte marker suggest defective differentiation. At P9, Tg mice exhibited dilated bile ducts and ectopic tubules throughout the lobule (**Fig. 3e and Supplementary Fig. 5a**). Immunofluorescence analysis showed that most of the ductal $CK19^+$ cells were positive for SOX9 (**Supplementary Fig. 5b**), as expected, but other cells scattered in the entire parenchyma were $HNF4\alpha^+/SOX9^+/CK19^-$ in P0 and P9 livers (**Fig 3e and Supplementary Fig 6a**), suggesting that they were hepatocytes with bi-phenotypic features. *In situ* hybridization analysis confirmed that SOX9 was expressed in the entire liver in Tg mice compared to controls (**Supplementary Fig. 6b**).

Together, these observations suggested that TFEB overexpression directs cell fate of HBs towards a progenitor/cholangiocyte lineage.

Sox9 is a direct target of TFEB

Our transcriptomic analysis in $TFEB^{OE}$ HBs and livers showed enrichment in several progenitor-specific genes, including *Sox9*. *Sox9* overexpression is known to be sufficient to induce biliary genes and suppress differentiation into the hepatocyte lineage[40]. Therefore, we investigated whether *Sox9* is a direct transcriptional target of TFEB and a mediator of TFEB effects on liver cell differentiation. We analyzed the promoter region of *Sox9* gene and identified five putative TFEB target sites (i.e. the CLEAR sites)[27] that were validated by chromatin immuno-precipitation (ChIP) analysis. Compared with a control sequence, the sequences closest to the *Sox9* transcriptional start site (TSS) were significantly enriched in liver samples from Tg mice (**Fig. 4a**). The region containing CLEAR sites closest to the TSS was cloned into a luciferase reporter plasmid to evaluate its responsiveness to TFEB in HBs.

Overexpression of TFEB increased luciferase activity in HBs, whereas deletion of the CLEAR sites failed to induce transactivation (**Fig. 4b**). Taken together these data indicate that TFEB binds to the *Sox9* promoter, identifying *Sox9* as a direct transcriptional target of TFEB.
To further validate our findings, we performed genetic interaction studies by crossing *Tcfeb*-3xFlag$^{fs/fs}$;*Alb-Cre* with *Sox9*$^{f/f}$ mice to generate double transgenic *Tcfeb*-3xFlag$^{fs/fs}$;*Alb-Cre*;*Sox9*$^{f/+}$ mice (hereafter Tg;*Sox9*$^{f/+}$). The expression of *Sox9* was significantly reduced in livers from Tg;*Sox9*$^{f/+}$compared with Tg mice (**Fig. 4c and 4d**). qPCR analysis on livers from Tg;*Sox9*$^{f/+}$ at P9 showed higher expression levels of hepatocyte-specific markers (e.g. *Alb*, *AldoB* and *Otc*) and a reduction of cholangiocyte-markers (e.g. *Krt7* and *Krt19*) compared with Tg mice (**Fig. 4c**). Interestingly, immunostaining analysis showed that *Sox9* knockdown resulted in reduced number of CK19$^+$ cells compared to Tg liver (**Fig. 4e**). These data suggest that SOX9 mediates, at least partially, the effects of TFEB on the determination of liver cell fate.

TFEB depletion impairs LPC expansion upon liver injury and regeneration

We then investigated the effects of TFEB deletion on liver cell differentiation in a previously described TFEB liver-specific conditional KO mouse line *Alb-Cre;Tcfeb*$^{f/f}$ (*Tcfeb*^{LiKO})[29]. While immunofluorescence analysis did not reveal any significant cell differentiation defect in TFEB depleted livers compared to age-matched controls at P0 and 1 month (**Supplementary Fig. 7**), DDC-induced liver damage (**Fig. 5a**) in *Tcfeb*^{LiKO} mice resulted in a reduction in the expression of *Sox9* and *Krt19* during the recovery phase (**Fig. 5b**), which was confirmed by immunoblot analysis and immunostaining for SOX9 at 14 days after recovery (**Fig. 5c and 5d**). These mice also showed a reduced ductular reaction as measured by CK19 immunostaining (**Fig. 5d**). Interestingly, while control livers showed hybrid hepatocytes positive for both HNF4α and SOX9 around the portal vein as consequence of liver damage, as previously shown[41], *Tcfeb*^{LiKO} mice showed SOX9$^+$ cells only in the portal area (**Fig. 5d and 5e**) and a strong reduction of proliferating SOX9$^+$ and CK19$^+$ cells, indicating impaired LPC activation (**Fig. 5f and 5g**). No main differences were observed in weight recovery after injury in the two genotypes (**Supplementary Fig. 8a**), while an increase in serum markers of liver damage (e.g. bilirubin and AST) was detected in *Tcfeb*^{LiKO} mice compared to control mice (**Supplementary Fig. 8b**). These data suggest that liver cell differentiation after injury requires TFEB induction. Consistently, we found that TFEB translocates from the cytoplasm to the nucleus 7 days after recovery from DDC-induced liver injury (**Supplementary Fig. 8c**), a time point during which we detected a reduction in mTOR activity, as measured by detecting the phosphorylation of the mTOR substrate 4EBP1 (**Supplementary Fig. 8d**).

TFEB overexpression in regenerating adult liver subverts cell identity

The effects of TFEB overexpression in the progenitor/cholangiocyte compartment were investigated by crossing the *Tcfeb*-3xFlag$^{fs/fs}$;*R26R^{LSL}tdTomato* mice with an inducible *Krt19*^{CreERT} mouse line that labels the biliary epithelium with a 40% recombination efficiency (**Fig. 6a and 6b**)[11]. Four-weeks after tamoxifen injection, tdTom expression was observed only in biliary ducts and in small periportal ductules (**Fig. 6a**). Based on the results obtained in *Alb-Cre* mice, we hypothesized that TFEB overexpression would expand the CK19$^+$ population in the periportal area by generating a ductal reaction. Indeed, we observed an increase in CK19$^+$/tdTom$^+$ cells in Tg mice compared to controls (**Fig. 6b**), hyperplasia of the bile ducts and a population of cells growing within the ductal epithelium and forming multilayered structures (**Fig. 6a**). In addition, some of the CK19$^+$/tdTom$^+$ cells appeared with a rounded morphology, compared to the cuboidal biliary epithelium observed in control mice, with a migrating phenotype that mimics a ductular reaction (**Fig. 6a**). This phenotype strongly resembles the one observed in YAP-overexpressing biliary cells[42].

Subsequently, to stimulate a progenitor/biliary-derived liver regenerative response, we induced chronic liver damage by feeding mice with a 0.1% 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC)-containing diet for 3 weeks and then switch to a normal diet for 14 days, a protocol causing transient LPC activation (**Fig. 6c**) [6]. Interestingly, while control mice showed tdTom$^+$ cholangiocytes in the portal area and the expected ductular reaction as shown by tdTom, CK19 and SOX9 immunostaining (**Fig. 6d**), most of the CK19$^+$ and SOX9$^+$ cholangiocytes were tdTom$^+$ in Tg mice, suggesting increased proliferation of the tdTom$^+$ cells upon TFEB overexpression, as also confirmed by BrdU analysis (**Fig. 6d**). Notably, these mice recapitulate the same phenotype observed in the *Alb-Cre*;*Tcfeb*-3xFlag$^{fs/fs}$ mice.

To overexpress TFEB specifically in the hepatocytes, we injected *Tcfeb*-3xFlag$^{fs/fs}$;*R26R^{LSL}tdTomato* mice with a CRE-expressing adeno-associated virus (AAV8-TBG-CRE). As expected, AAV8 injection induced CRE expression in 99% of the hepatocytes and no expression in the biliary cells as demonstrated by tdTom staining (**Fig 7a**)[6]. These mice showed 20-fold increase of *Tcfeb* expression at 4-weeks after injection (**Fig 7b**), but no alteration in cell identity compared to controls (**Fig 7c**). These results suggest that TFEB overexpression in differentiated adult hepatocytes is not sufficient to induce their conversion to cholangiocytes.

To evaluate whether TFEB overexpression in hepatocytes alters cell identity upon liver injury, we fed *Tcfeb*-3xFlag$^{fs/fs}$;*R26R^{LSL}tdTomato* mice injected with AAV8-TBG-CTRL or AAV8-TBG-CRE with the DDC-containing diet. We first used a protocol of acute injury by feeding mice a DDC-containing diet for 5 days and switching to a normal diet for 14 days (**Fig 7d**).

Interestingly, CRE-injected mice showed tdTom^{+} hepatocytes positive for SOX9 and CK19, strongly suggesting a conversion of TFEB^{OE} cells in progenitor/cholangiocyte-like cells upon liver damage (**Fig 7e**). Notably, the chronic exposure to DDC-containing food for 3 weeks following by 14 days of recovery resulted in liver tumors in CRE-injected mice (**Fig 7f and 7g**). Histological analysis showed that most of the CK19^{+}/SOX9^{+} cells were tdTom^{+}, and that some tdTom+ hepatocytes also expressed SOX9 (**Fig 7h**), confirming that TFEB overexpression in hepatocytes leads to trans-differentiation in progenitor/cholangiocytes in conditions of liver damage when cells are more prone to proliferation and trans-differentiation.
Together our data strongly confirm that TFEB overexpression in progenitors/cholangiocytes is able to influence proliferation and differentiation.

Liver cysts, cholestasis and cholangiocarcinoma-like phenotype in TFEB^{OE} mice

We then analyzed the phenotypic consequence of the altered cell differentiation associated with TFEB overexpression in the liver. Tg mice exhibited significant hepatocellular derangement as shown by an increase of liver damage markers, as well as a significant increase in both total bilirubin and direct bilirubin at 15 days and 3 months (**Supplementary Table 5**) consistent with cholestasis. Serum bile acid and cholesterol levels were also increased in Tg mice compared to controls (**Supplementary Table 5**). Analysis of three-dimensional structure of the biliary system in adult liver, performed by retrograde injection of ink in the biliary tree in 2-month-old mice, confirmed a defect in bile duct development (**Fig. 8a**). Indeed, we observed reduced density of branches arising from major branches and abnormal major branches in appearance consistent with partial obstruction of bile flow, correlating with increased bilirubin and ALT levels. Furthermore, Tg mice showed a progressive increase in liver mass relative to total body mass with a 2, 4 and 10-fold increase at 2 weeks, 2 months, and 5 months of age, respectively (**Fig. 8b and 8c**). At 3 months, TFEB overexpression was detected in the entire parenchyma and in particular in the dilated ducts as demonstrated by immunohistochemistry and *in situ* hybridization analysis (**Fig. 8d**). Livers from Tg mice displayed variable degrees of hyperplasia of bile ducts, which were increased in size and number, and an altered morphology of epithelial cells. In addition, they showed multifocal biliary cystic hyperplasia with cysts lined by flattened epithelium at 3 months of age (**Fig. 8e**). Densely packed cholangiocytes, as revealed by CK19 immunostaining, associated with increased fibrosis revealed by Sirius red staining, were also detected suggesting the presence of a neoplasms of the bile ducts. All of these phenotypic features are consistent with the presence of a cholangiocarcinoma (CCA)-like phenotype. Livers from 6-month-old Tg mice showed multiple cysts replacing most of the hepatic parenchyma. At this age mice began to die or needed to be euthanized due to poor health status. Increased liver mass in Tg mice was associated with a higher proliferation index compared to control mice, as shown by

increased Ki67 staining (**Fig. 8e**). A similar, albeit milder and with later onset, phenotype was observed in a different transgenic mouse line in which *Tcfeb* is overexpressed at lower levels (2.5-fold increase) suggesting a dose- and time-dependent effect of TFEB overexpression (**Supplementary Fig. 9a-9d**).

These data are consistent with a pro-cholangiocytic function of TFEB and suggest that TFEB induction may be involved in the pathogenesis of cholangiocarcinoma.

Discussion

Recent studies have identified an important role for MiT-TFE transcription factors (TFs) in the regulation of basic cellular processes, such as lysosome, autophagosome and melanosome biogenesis[27,28,43,44]. These TFs control organismal adaptation to environmental cues, such as nutrient availability, physical exercise and infections[32,33,37,45-49]. While some studies reported the involvement of MiT-TFE genes in cell differentiation, the precise role that these TFs play in cell specification and in embryonic development has remained elusive. TFEB full KO mice die at E10 due to a placental vascularization defect, indicating that embryonic development requires TFEB mediated transcriptional control[26]. In addition, previous studies showed that MITF and TFEB are involved in the differentiation processes of melanocytes and osteoclasts, respectively[50]. A recent study also reported a role for TFEB in the activation and response to growth factor in quiescent neuronal stem cells (qNSCs)[51]. Furthermore, TFE3, another member of the MiT family, was shown to be involved in stem-cell commitment by enabling ESCs to withstand differentiation condition[52]. In the present study, we reveal a novel role for TFEB in liver cell differentiation during development and in the regenerative response to injury. We found that TFEB expression levels increase over time during liver specification. Moreover, we observed differences in the levels of TFEB expression between pericentral vs periportal area suggesting that different levels of TFEB could determine different progenitor cell fates. Indeed, our data show that TFEB mainly specify the cholangiocyte lineage. This is also supported by the finding that TFEB overexpression in hepatoblasts dictates a progenitor/cholangiocytic fate resulting in a rapid replacement of the entire liver by biliary structures. In this regard, it will be interesting to determine the mechanisms that allow the differential expression of TFEB in specific subsets of cells.

Hepatocyte identity is defined by the expression of a core group of transcription factors primarily driven by the CCAAT/enhancer binding protein alpha (C/EBPα)[53], a key hepatic transcription factor that also controls the expression of genes involved in ammonia detoxification and glucose and lipid homeostasis[54]. Another key regulator of cell identity in the liver is SOX9, a transcription factor that drives bile duct morphogenesis and is recognized as

a specific marker of precursors and biliary cells in the developing liver[55]. Remarkably, C/EBPα and SOX9 form a mutually antagonistic system controlling the hepatocyte versus biliary fate during normal liver homeostasis and regeneration[40]. Our *in vitro* and *in vivo* studies indicate that TFEB directly controls the expression of *Sox9* during progenitor/ductal specification and after liver injury. Indeed, our epistatic analysis using *Sox9*-deficient mice provides evidence that TFEB-mediated regulation of SOX9 is required for proper differentiation of bi-potential progenitors, although additional mechanisms downstream of TFEB are likely to play a role.

Consistent with a role of TFEB in the determination of liver cell fate, we found that mice overexpressing TFEB in the liver showed increased proliferation of progenitor cells, expansion of immature BECs and inhibition of the hepatocyte lineage. An opposite phenotype was observed in *Tcfeb*-depleted HBs and liver, which appeared more prone to differentiation into the hepatocyte lineage (**Supplementary Fig 10**).

Although liver parenchymal cells turn over slowly, the liver displays high regenerative capacity, capable of restoring 70% tissue loss within a few weeks[56]. This ability is vital for the liver to maintain constant mass. The population of liver cell responding to injury may differ depending on the site, type and the duration of injury. HybHPs are located in the periportal area and express progenitor markers, such as *Sox9*, and hepatocyte markers, such as *Hnf4α.* This population of specialized hepatocytes are able to reconstitute the liver mass after various chronic injuries[21]. *Sox9*$^{+}$/*Hnf4α*$^{+}$ bi-phenotypic hepatocytes able to differentiate in either hepatocytes or cholangiocytes have been also identified in DDC-injured livers[41]. These cells represent the intermediate status of lineage conversion during liver regeneration. We found that conditional deletion of TFEB in the liver results in impaired LPCs activation. Moreover, we demonstrated that TFEB depleted hepatocytes fail to express *Sox9* after liver damage, thus suggesting that TFEB-dependent *Sox9* induction plays an important role in liver homeostasis in response to injury. This observation is in line with the known role of TFEB in the cellular adaptation to environmental cues such as various types of stress conditions[33,47].

It is known that overexpression of endogenous MiT-TFE transcription factors, as a result of chromosomal abnormalities such as translocations or gene amplifications, can drive tumorigenesis in renal cell carcinoma and melanoma[34,35,37,57] and to support cancer growth in pancreatic cancer[35-37]. However, a specific role for TFEB in cell differentiation, proliferation and neoplastic transformation in the liver has never been demonstrated. Here, we show that *Albumin*-CRE-driven TFEB overexpression in the liver of transgenic mice results in a severe phenotype characterized by a progressive increase in liver mass, hyperplasia of bile ducts, altered biliary tree structure, multifocal biliary cystic hyperplasia and fibrosis. Cholestasis, bile duct proliferation and cystic alterations rapidly progress ultimately leading to biliary tumor with

features resembling cholangiocarcinoma (CCA). Previous studies showed that SOX9 is highly up-regulated in many premalignant lesions and in tumor tissues and plays a role in tumor development[58-60]. Our results suggest that TFEB-mediated up-regulation of *Sox9* expression plays a role in liver cancer development by inducing progenitor cell proliferation.

In conclusion, our study reveals an important role for TFEB in liver development and cell fate under physiological and pathological conditions. Further studies of this mechanism may lead to the identification of novel therapeutic targets for the modulation of liver regeneration after injury.

Methods

Mouse experiments. All mice used in experiments were males and were maintained on a C57BL/6 strain background. All experiments were approved by the Baylor College of Medicine Institutional Animal Care and Use Committee (IACUC) and conform to the legal mandates and federal guidelines for the care and maintenance of laboratory animals.

Tcfeb$^{LacZ/LacZ}$, *Tcfeb*$^{flox/flox}$ and T*cfeb*-3xFlag$^{fs/fs}$ transgenic mouse line generation has been previously described[28,29]. Wild-type, *Sox9*$^{flox/flox}$, *R26R*^{LSL}*tdTomato, Albumin*-Cre and *Krt19*^{CreERT} mouse lines were obtained from the Jackson laboratory (Bar, Harbor, ME).

The *Krt19*^{CreERT} was induced by three individual intraperitoneal injections of tamoxifen at a dose of 4 mg. *Krt19*^{CreERT} mice received 2 weeks of normal diet after the last tamoxifen injection before starting an injury diet regime.

To induce liver injury, mice were given 0.1% DDC food (Custom Animal Diet, Bangor, PA) as indicated in the text. After injury, mice were given normal chow and drinking water.

Cell culture. Hepatoblasts (HBs) were prepared from control and T*cfeb*-3xFlag$^{fs/fs}$ mice at embryonic day 13.5 and immortalized by plating at clonal density[61]. HBs were maintained in HB media (RPMI) (Thermo Fisher Scientific, Wilmington, DE) containing 10% FBS, 1% penicillin-streptomycin, 50 ng/ml epidermal growth factor, 30 ng/ml insulin-like growth factor II (PeproTech, Rocky Hill, NJ), 10 μg/ml insulin (Roche Diagnostics, Indianapolis, IN) on plates coated with rat tail collagen (BD Bioscience, San Jose, CA) in a humidified atmosphere with 5% CO_2 at 37°C. Cells were infected with HDAd-BOS-CRE or HDAd-CMV-LacZ vectors (previously described[62,63]) and then cultured for the specific assay. For hepatocyte differentiation, HBs were cultured on uncoated tissue culture dishes in HB medium for up to 5 days. For bile duct differentiation, 6 cm tissue culture dishes were coated with 0.5 ml Basement Membrane Matrix (BD Bioscience, San Jose, CA) and allowed to set for 1 h. HBs in HB cell media supplemented with 100 ng/ml recombinant mouse hepatocyte growth factor (HGF) (R&D System, Minneapolis, MN) were then added to the plate. Tubule formation was

monitored at 24 hours, 3 days and 10 days.

Primary mouse hepatocytes were isolated using a two-step perfusion technique as previously described[64]. Briefly, WT or transgenic mouse liver was perfused with collagenase (C5138, Sigma-Aldrich, Saint Louis, MO) and parenchymal cells were squeezed out from the liver. The hepatocyte suspension was further purified using a 40% of Percoll gradient (P4937, Sigma-Aldrich, Saint Louis, MO), washed with Hepatocyte Wash Medium (17004-024; Thermo Fisher Scientific) and plated at an appropriate cell density. Cells were harvested for analysis after 24 hours.

Generation of *Tcfeb*^{KO} hepatoblasts through CRISPR-Cas9. Two sgRNAs targeting adjacent regions of *Tcfeb* exon 3 were designed and synthetized as previously described[65,66] with minimal modification. Briefly, forward primers containing the T7 promoter sequence, the proto-spacer sequence and the sgRNA scaffold overlap sequence were designed through CRISPRscan[67]. Full-lenght sgRNA scaffold was obtained through an overlap PCR with the universal scaffold reverse primer. Tfeb sgRNA 1 Forward: taatacgactcactataGGGTATCTGTCTGAGACCTAgttttagagctagaaatagc
Tfeb sgRNA 2 Forward: taatacgactcactataGGCAGGCTTCGGGGAACCTTgttttagagctagaaatagc
Universal scaffold Reverse: gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgct
PCR products were used for *in vitro* transcription (IVT). IVT was performed using the HiScribe T7 High Yield RNA Synthesis kit (NEB) following manufacturer instructions. Purified sgRNAs were pre-complexed with *sp*Cas9 protein (IDT). Finally, sgRNA-Cas9 ribonucleoproteins (RNP) containing the two sgRNAs were simultaneously transfected into primary mouse Hepatoblasts using the NEON Transfection System (ThermoFisher) and the following electroporation conditions: Buffer R, 1400 V, 10 ms, 3 pulses. Deletion efficiency was assessed through PCR using the following primers: *Tcfeb* Forward: GTCCACTTCCAGTCGCCC; *Tcfeb* Reverse: AGGCTAGAGGCCCATAAAGAA.

Flow cytometry and cell cycle analysis. HBs were cultured on collagen-coated plates, harvested with trypsin-EDTA (Thermo Fisher Scientific, Wilmington, DE), and washed in PBS. For cell cycle analysis, cells were fixed in 1 ml cold 70% ethanol for 30 minutes on ice and then centrifuged for 5 minutes at high speed. The pellet was washed twice in PBS and treated with RNAseA (Life Technologies, Carlsbad, CA). Propidium Iodide (PI) solution was added to the cells and the mixture incubated for 10 minutes at room temperature. Cell debris and dead cells were excluded from FSC-A/SSC-A dot-plots (**Supplementary Figure 11a**). DNA content and relative cell cycle phases was determined based on PI relative fluorescence levels on

singlets (**Supplementary Figure 11b**). FACS experiments were performed using an LSRII cytometer (BD Bioscience, San Jose, CA Bioscience, San Jose, CA). Data were analyzed using FACSDiva (BD Bioscience, San Jose, CA) and Flow-Jo (Flow-Jo, LLC, Ashland, OR) software. Cell cycle phases were quantified using the Flow-Jo Cell Cycle platform (univariate). In all experiments at least 10000 events were acquired.

RNA extraction and quantitative RT-PCR. Total RNA was extracted in TRIzol reagent (Life Technologies, Carlsbad, CA) using RNeasy kit (Qiagen, Hilden, Germany). RNA was reverse transcribed using a first strand complementary deoxyribonucleic acid kit with random primers according to the manufacturer's protocol (Applied Biosystems). The RT-PCR reactions were performed using the CFX96 Real Time System (Bio-Rad, Hercules, CA). The PCR reaction was performed using iTaq SYBR Green Supermix (Bio-Rad, Hercules, CA) with the following thermocycler conditions: pre-heating, 5 min at 95°C; cycling, 40 cycles of 15 s at 95°C, 15 s at 60°C and 25 s at 72°C. Results were expressed in terms of cycle threshold(C_t). The C_t values were averaged for each duplicate. The *ß$_2$-microglobulin*, *Ribosomal protein S16* or *Cyclophilin* genes were used as endogenous controls (reference markers). Differences between the mean C_t values of the tested genes and those of the reference gene were calculated as $\Delta C_{tgene}=C_{tgene}-C_{treference}$. Relative fold increase in expression level was determined as $2^{-\Delta\Delta Ct}$. Primers used for RT-PCR are listed in **Supplementary Table 6**

Western blotting. Liver and cells samples were homogenized in RIPA buffer (50mM Tris–HCl pH 7.4, 150mM NaCl, 1% Triton X-100, 1mM EDTA pH 8.0, 0.1% SDS) containing complete protease inhibitor cocktail (Roche Diagnostics, Indianapolis, IN). Samples were incubated for 20 minutes at 4°C and centrifuged at 16,000 g for 10 minutes. The pellet was discarded, and cell lysates were used for Western blot analysis. Ten to twenty micrograms of liver protein were run on a 4-20% SDS-PAGE polyacrylamide gel by electrophoresis. After transfer to a PVDF membrane, the blots were blocked in TBS-Tween 20 containing 5% non-fat milk for 1 hour at room temperature, then exposed to primary antibody was applied overnight at 4°C. Anti-rabbit IgG or anti-mouse IgG conjugated with horseradish peroxidase (GE Healthcare, Little Chalfont, UK) and ECL (Pierce, Thermo Fisher Scientific, Wilmington, DE) was used for detection. Antibodies used for immuno-blots are listed in **Supplementary Table 7**.

Cellular fractionation. Enriched nuclear and cytosolic cellular subfractions were isolated by differential centrifugation, as previously described[68]. Briefly, the liver was minced on ice and homogenized using a Teflon pestle and mortar and suspended in mitochondrial isolation buffer (MIB; 250 mM sucrose, 20 mM HEPES, 10 mM KCl, 1.5 mM MgCl2, 1 mM EDTA, 1 mM

EGTA) supplemented with protease and phosphatase inhibitor cocktails (Complete and PhosSTOP Roche, Roche Diagnostics, Basel, Switzerland). The homogenates were then centrifuged at 1,000 g for 10 min at 4°C to pellet the nuclei while mitochondrial and cytosolic fractions were contained within the supernatant. The supernatant fraction was re-centrifuged twice at 16,000 g for 20 min at 4°C to pellet the mitochondria and supernatant containing cytosolic subfraction was collected. Pellets containing nuclei were re-suspended in nuclear lysis buffer (1.5 mM MgCl2, 0.2 mM EDTA, 20 mM HEPES, 0.5 M NaCl, 20% glycerol, 1% Triton X-100), incubated on ice for 30 min, and then sonicated 3 Å~ 10 s followed by a final centrifugation step of 15 min at 16,000 g. The supernatant was collected to obtain the enriched nuclear fraction. Fraction purity was determined by Western blot analysis.

In situ **hybridization**. Livers from P9 and 3-month-old WT and Tg mice were collected and frozen without prior fixation in OCT cryoprotection media. Tissue was cut into 25-μm sections, mounted on slides with ProlongGold with DAPI (Life Technologies, Carlsbad, CA) and used for RNA *in situ hybridization* (ISH). We utilized the previously described method[69] but modified the development of the signal by using a Cy3-labeled tyramide instead of the biotin-labeled tyramide. *Tcfeb*-probe-F1 5'-GCGGCAGAAGAAAGACAATC-3' and *Tcfeb*-probe-R1 5'-AGGTGATGGAACGGAGACTG-3' were used to amplify 1300 bp from mouse liver cDNA and cloned into pGEM_T Easy vector (Promega, Madison, MI). This template was used to generate a DIG-labeled mRNA probe using in vitro transcription reagents from Roche (Roche Diagnostics, Indianapolis, IN).

Liver staining. Livers were dissected, fixed with buffered 10% formalin overnight at 4°C and stored in 70% ethanol, embedded into paraffin blocks and cut into 6-μm sections. For Sirius red staining the sections were rehydrated and stained for 1 hour in picro-sirius red solution (0.1% Sirius red in a saturated aqueous solution of picric acid) (Sigma-Aldrich, Saint Louis, MO). After two changes of acidified water (5ml glacial acetic acid in 1 L of water), the sections were dehydrated in three washes of 100% ethanol, cleared in xylene, and mounted on a resinous medium. X-Gal and H&E stainings were performed following the IHC World protocol. For immuno-staining, the sections were rehydrated to PBS pH 7.4 and permeabilized with 0.2% Triton in PBS. Heat Induced Epitope Retrieval using the citrate buffer method (pH 6.0) was performed to retrieve the antigen sites. The sections were then incubated for 1 hour at room temperature with blocking solution (3% BSA, 5% donkey serum, 20mM MgCl2, 0.3% Tween 20 in PBS pH 7.4). For co-stainings, primary antibodies were incubated overnight at 4°C. Secondary antibodies made in donkey were: AlexaFluor-488 anti-rabbit and AlexaFluor-594 anti-mouse (Life Technologies, Carlsbad, CA). Slides were mounted on ProlongGold with

DAPI (Life Technologies, Carlsbad, CA). For IHC, the Vectastain ABC kit and DAB (Vector Laboratories, Burlingame, CA) were used following the manufacturer's instructions. Sections were counterstained using Mayer's hematoxylin (Electron Microscopy Sciences, Hatfield, PA). For BrdU staining, 2 mg of BrdU (BD Bioscience) diluted in PBS 1x was injected by intraperitoneal injection 4 hours before tissue collection. Livers were dissected, fixed with buffered 10% formalin overnight at 4°C and embedded into OCT blocks and cut into 10-μm sections. BrdU signal was revealed by using a rat anti-BrdU antibody. Stained liver sections were examined under a Zeiss Axiocam MR microscope. For staining quantification, Image J Software (NIH) was used to calculate the percent area positively stained in five random low power views.

Gene expression analysis. Total RNA from HBs 3 days after hepatocytic differentiation was quantified using the Qubit 2.0 fluorimetric Assay (Thermo Fisher Scientific). Libraries were prepared from 100ng of total RNA using the QuantSeq 3' mRNA-Seq Library Prep Kit FWD for Illumina (Lexogen GmbH). Quality of libraries was assessed by using screen tape High sensitivity DNA D1000 (Agilent Technologies). Libraries were sequenced on a NovaSeq 6000 sequencing system using an S1, 100 cycles flow cell (Illumina Inc.). Sequence reads were trimmed using bbduk software (bbmap suite 37.31) to remove adapter sequences, poly-A tails and low-quality end bases (regions with average quality below 6). Alignment was performed with STAR 2.6.0a[70] on mm10 reference assembly. The expression levels of genes were determined with htseq-count 0.9.1[71] by using mm10 Ensembl assembly (release 90). We have filtered out all genes having < 1 cpm in less than 2 samples. Differential expression analysis was performed using edgeR[72].

Microarray analysis was performed on liver from P9 mice of each of the two genotypes (n=3 per group). All samples were processed on Affymetrix Mouse 430A 2.0 arrays using GeneChip 3'-IVT Plus and Hybridization Wash and Stain kits by means of Affymetrix standard protocols. Raw intensity values of the 6 arrays were processed and normalized by Robust Multi-Array Average Method[73] using the Bioconductor R package Affy[74]. The data have been deposited in NCBIs Gene Expression Omnibus[75] (GEO) and are accessible through GEO Series accession number GSE35015.

Enrichment analyses of KEGG pathways on differentially expressed genes were performed using the Bioconductor R package clusterProfiler[76] (FDR <= 0.05). Gene set enrichment analysis (GSEA) was run on HNF4α targets downloaded from the Molecular Signatures Database (MSigDB)[77].

Promoter analysis. The analysis of TFEB binding sites was performed using the CLEAR matrix[78] and the matchPWM algorithm[79] implemented in the Biostrings package. Putative binding sites were filtered with threshold of significance set at 0.8.

ChIP. ChIP was performed using livers of 2-month-old Alb-Cre;*Tcfeb*-3xFlag and control mice as previously described[33]. Primers used for RT-PCR are listed in **Supplementary Table 8.**

Luciferase assay. The *Sox9* promoter was amplified from mouse genomic DNA and cloned into the pGL4 plasmid (Promega, Madison, MI). A quick-change site directed mutagenesis kit (Agilent, Santa Clara, CA) was used for the mutagenesis of CLEAR sites in the *Sox9* promoter. Primers used for mutagenesis are listed in **Supplementary Table 9**. Control HBs and HBs overexpressing *Tcfeb* were transfected with the *Sox9* promoter-reporter luciferase construct and pRL-CMV (Promega, Madison, MI). Cells were harvested 24 hours after transfection and subjected to luminescence detection by dual luciferase assay using the Dual-Luciferase Reporter Assay system (Promega, Madison, MI). A Turner Designs Luminometer (DLReady) (Promega, Madison, MI) was used to measure luminescence normalized against *Renilla* luciferase activity.

Biliary tree casting. Two-month-old mice were euthanized with isoflurane (Vedco Saint Joseph MO). Ink (Hyatt's) was injected into the common bile duct using a 30-gauge needle. The entire liver was removed, formalin-fixed and clarified in a 1:2 solution of benzyl alcohol and benzyl benzoate.

Blood chemistry analysis. Blood was collected from the retro-orbital plexus under isoflurane (Vedco Saint Joseph MO) anaesthesia. Serum was frozen at -20°C or used immediately after collection. Total Bile Acids (TBA) Enzymatic Cycling Assay Kit was obtained from BQKITS (San Diego, CA).

Vector production and injections. pAAV-TBG-GFP-pA (GFP) was obtained from Thomas Vallim (University of California, Los Angeles). pAAV-TBG-PI-Cre-rBG (Cre Recombinase) and plasmids required for AAV packaging, adenoviral helper plasmid pAdDeltaF6 (PL-F-PVADF6) and AAV8 packaging plasmid pAAV2/8 (PL-T-PV0007) were obtained from the University of Pennsylvania Vector Core. AAV were generated as previously described with some modifications[80,81]. Each AAV transgene construct was co-transfected with the packaging constructs into 293T cells (ATCC, CRL-3216) using polyethylenimine (PEI). Cell pellets were harvested and purified using a single cesium chloride density gradient centrifugation. Fractions containing AAV vector genomes were pooled and then dialyzed against PBS using

a 100kD Spectra-Por® Float-A-Lyzer® G2 dialysis device (Spectrum Labs, G235059) to remove the cesium chloride. Purified AAV were concentrated using a Sartorius™ Vivaspin™ Turbo 4 Ultrafiltration Unit (VS04T42) and stored at -80°C until use. AAV titers were calculated after DNase digestion using qPCR standard method. Primers used for titer are included in **Supplementary table 10**. Viruses were administered by retro-orbital injection at the dose of $2x10^{11}$ genome copies/mouse.

Statistical analyses. Data are expressed as averages ± standard error. Statistical significance was computed using Student's 2-tail t-test. A *p*-value<0.05 is considered statistically significant.

Author contribution

N.P. performed the experiments and interpreted the results. T.H., N.J.H., A.C. and L.D. provided technical support to N.P. T.J.K. performed ChIP experiments. L.B. helped with the generation of CRISPR/CAS9 clones and with FACS analysis. K.H.K. isolated primary hepatocytes. M.D.G., H.A. and W.L. provided AAV8 viruses. A.C. and M.M. analyzed the gene expression data. N.A., D.D.M., C.S, M.F. and S.J.F. contributed to interpretation of the results. A.B. and N.P designed the overall study, supervised the work, and wrote the manuscript.

Acknowledgements

We thank Marcus Grompe, Huda Zoghbi, Nicola Brunetti-Pierri, Roberto Zoncu, Rushika Perera and Graciana Diez Roux for critical reading of the manuscript. We are grateful to Diego Carrella and Rossella De Cegli (TIGEM) for promoter analysis and ChIP primer design. We thank Dr Thomas Vallim (University of California, Los Angeles) for the pAAV-TBG-GFP plasmid. We also thank the NGS Core at TIGEM for help with gene expression analysis. This work was supported by a grant from the US National Institutes of Health (R01-NS078072), AIRC (Italian Association for Cancer Research) (IG 2015-17639), Foundation Louis-Jeantet and Telethon Foundation to AB. The project was supported in part by the RNA In Situ Hybridization Core facility at Baylor College of Medicine which is supported by a Shared Instrumentation grant from the NIH (1S10OD016167) and the NIH IDDRC Grant (1U54HD083092-02) from the Eunice Kennedy Shriver National Institute of Child Health & Human Development. This project was also supported in part by the NIDDK Digestive Disease Center Core with funding from the NIH (NIH P30DK58338) and in part by the Pathology and histology Core at Baylor College of Medicine with funding from the NIH (NCI-CA125123).

Conflict of Interest

A. Ballabio is a co-founder of CASMA Therapeutics, Boston, MA, USA.

References

1 Zorn, A. M. in *StemBook* (2008).

2 Roskams, T. A., Libbrecht, L. & Desmet, V. J. Progenitor cells in diseased human liver. *Seminars in liver disease* **23**, 385-396, doi:10.1055/s-2004-815564 (2003).

3 Boulter, L., Lu, W. Y. & Forbes, S. J. Differentiation of progenitors in the liver: a matter of local choice. *The Journal of clinical investigation* **123**, 1867-1873, doi:10.1172/JCI66026 (2013).

4 Sell, S. Heterogeneity and plasticity of hepatocyte lineage cells. *Hepatology* **33**, 738-750, doi:10.1053/jhep.2001.21900 (2001).

5 Tarlow, B. D. *et al.* Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes. *Cell stem cell* **15**, 605-618, doi:10.1016/j.stem.2014.09.008 (2014).

6 Yanger, K. *et al.* Robust cellular reprogramming occurs spontaneously during liver regeneration. *Genes & development* **27**, 719-724, doi:10.1101/gad.207803.112 (2013).

7 Sekiya, S. & Suzuki, A. Hepatocytes, rather than cholangiocytes, can be the major source of primitive ductules in the chronically injured mouse liver. *The American journal of pathology* **184**, 1468-1478, doi:10.1016/j.ajpath.2014.01.005 (2014).

8 Yanger, K. *et al.* Adult hepatocytes are generated by self-duplication rather than stem cell differentiation. *Cell Stem Cell* **15**, 340-349, doi:10.1016/j.stem.2014.06.003 (2014).

9 Schaub, J. R. *et al.* De novo formation of the biliary system by TGFbeta-mediated hepatocyte transdifferentiation. *Nature* **557**, 247-251, doi:10.1038/s41586-018-0075-5 (2018).

10 Deng, X. *et al.* Chronic Liver Injury Induces Conversion of Biliary Epithelial Cells into Hepatocytes. *Cell Stem Cell*, doi:10.1016/j.stem.2018.05.022 (2018).

11 Raven, A. *et al.* Cholangiocytes act as facultative liver stem cells during impaired hepatocyte regeneration. *Nature* **547**, 350-354, doi:10.1038/nature23015 (2017).

12 Yovchev, M. I. *et al.* Identification of adult hepatic progenitor cells capable of repopulating injured rat liver. *Hepatology* **47**, 636-647, doi:10.1002/hep.22047 (2008).

13 Lowes, K. N., Croager, E. J., Abraham, L. J., Olynyk, J. K. & Yeoh, G. C. Upregulation of lymphotoxin beta expression in liver progenitor (oval) cells in chronic hepatitis C. *Gut* **52**, 1327-1332 (2003).

14 Wang, X. *et al.* Osteopontin induces ductular reaction contributing to liver fibrosis. *Gut* **63**, 1805-1818, doi:10.1136/gutjnl-2013-306373 (2014).

15 Khuu, D. N., Nyabi, O., Maerckx, C., Sokal, E. & Najimi, M. Adult human liver mesenchymal stem/progenitor cells participate in mouse liver regeneration after hepatectomy. *Cell transplantation* **22**, 1369-1380, doi:10.3727/096368912X659853 (2013).

16 Li, Z. *et al.* Human hepatocyte growth factor (hHGF)-modified hepatic oval cells improve liver transplant survival. *PLoS One* **7**, e44805, doi:10.1371/journal.pone.0044805 (2012).

17 Wu, C. X., Zou, Q., Zhu, Z. Y., Gao, Y. T. & Wang, Y. J. Intrahepatic transplantation of hepatic oval cells for fulminant hepatic failure in rats. *World J Gastroenterol* **15**, 1506-1511 (2009).

18 Stueck, A. E. & Wanless, I. R. Hepatocyte buds derived from progenitor cells repopulate regions of parenchymal extinction in human cirrhosis. *Hepatology* **61**, 1696-1707, doi:10.1002/hep.27706 (2015).

19 Malato, Y. *et al.* Fate tracing of mature hepatocytes in mouse liver homeostasis and regeneration. *J Clin Invest* **121**, 4850-4860, doi:10.1172/JCI59261 (2011).

20 Kopp, J. L., Grompe, M. & Sander, M. Stem cells versus plasticity in liver and pancreas regeneration. *Nature cell biology* **18**, 238-245, doi:10.1038/ncb3309 (2016).

21 Font-Burgada, J. *et al.* Hybrid Periportal Hepatocytes Regenerate the Injured Liver without Giving Rise to Cancer. *Cell* **162**, 766-779, doi:10.1016/j.cell.2015.07.026 (2015).

22 Gilgenkrantz, H. & Collin de l'Hortet, A. Understanding Liver Regeneration: From Mechanisms to Regenerative Medicine. *Am J Pathol* **188**, 1316-1327, doi:10.1016/j.ajpath.2018.03.008 (2018).

23 Roskams, T. Liver stem cells and their implication in hepatocellular and cholangiocarcinoma. *Oncogene* **25**, 3818-3822, doi:10.1038/sj.onc.1209558 (2006).

24 Yamashita, T. & Wang, X. W. Cancer stem cells in the development of liver cancer. *The Journal of clinical investigation* **123**, 1911-1918, doi:10.1172/JCI66024 (2013).

25 Hemesath, T. J. *et al.* microphthalmia, a critical factor in melanocyte development, defines a discrete transcription factor family. *Genes & development* **8**, 2770-2780 (1994).

26 Steingrimsson, E., Tessarollo, L., Reid, S. W., Jenkins, N. A. & Copeland, N. G. The bHLH-Zip transcription factor Tfeb is essential for placental vascularization. *Development* **125**, 4607-4616 (1998).

27 Sardiello, M. *et al.* A gene network regulating lysosomal biogenesis and function. *Science* **325**, 473-477, doi:10.1126/science.1174447 (2009).

28 Settembre, C. *et al.* TFEB links autophagy to lysosomal biogenesis. *Science* **332**, 1429-1433, doi:10.1126/science.1204592 (2011).

29 Settembre, C. *et al.* A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *The EMBO journal* **31**, 1095-1108, doi:10.1038/emboj.2012.32 (2012).

30 Roczniak-Ferguson, A. *et al.* The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. *Sci Signal* **5**, ra42, doi:10.1126/scisignal.2002790 (2012).

31 Martina, J. A., Chen, Y., Gucek, M. & Puertollano, R. MTORC1 functions as a transcriptional regulator of autophagy by preventing nuclear transport of TFEB. *Autophagy* **8**, 903-914, doi:10.4161/auto.19653 (2012).

32 Mansueto, G. *et al.* Transcription Factor EB Controls Metabolic Flexibility during Exercise. *Cell Metab* **25**, 182-196, doi:10.1016/j.cmet.2016.11.003 (2017).

33 Settembre, C. *et al.* TFEB controls cellular lipid metabolism through a starvation-induced autoregulatory loop. *Nature cell biology* **15**, 647-658, doi:10.1038/ncb2718 (2013).

34 Haq, R. & Fisher, D. E. Biology and clinical relevance of the micropthalmia family of transcription factors in human cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **29**, 3474-3482, doi:10.1200/JCO.2010.32.6223 (2011).

35 Calcagni, A. *et al.* Modelling TFE renal cell carcinoma in mice reveals a critical role of WNT signaling. *eLife* **5**, doi:10.7554/eLife.17047 (2016).

36 Perera, R. M. *et al.* Transcriptional control of autophagy-lysosome function drives pancreatic cancer metabolism. *Nature*, doi:10.1038/nature14587 (2015).

37 Di Malta, C. *et al.* Transcriptional activation of RagD GTPase controls mTORC1 and promotes cancer growth. *Science* **356**, 1188-1192, doi:10.1126/science.aag2553 (2017).

38 Lemaigre, F. P. Development of the biliary tract. *Mechanisms of development* **120**, 81-87 (2003).

39 Weisend, C. M., Kundert, J. A., Suvorova, E. S., Prigge, J. R. & Schmidt, E. E. Cre activity in fetal albCre mouse hepatocytes: Utility for developmental studies. *Genesis* **47**, 789-792, doi:10.1002/dvg.20568 (2009).

40 O'Neill, K. E. *et al.* Hepatocyte-ductal transdifferentiation is mediated by reciprocal repression of SOX9 and C/EBPalpha. *Cellular reprogramming* **16**, 314-323, doi:10.1089/cell.2014.0032 (2014).

41 Tanimizu, N., Nishikawa, Y., Ichinohe, N., Akiyama, H. & Mitaka, T. Sry HMG box protein 9-positive (Sox9+) epithelial cell adhesion molecule-negative (EpCAM-) biphenotypic cells derived from hepatocytes are involved in mouse liver regeneration. *J Biol Chem* **289**, 7589-7598, doi:10.1074/jbc.M113.517243 (2014).

42 Yimlamai, D. *et al.* Hippo pathway activity influences liver cell fate. *Cell* **157**, 1324-1338, doi:10.1016/j.cell.2014.03.060 (2014).

43 Martina, J. A. *et al.* The nutrient-responsive transcription factor TFE3 promotes autophagy, lysosomal biogenesis, and clearance of cellular debris. *Science signaling* **7**, ra9, doi:10.1126/scisignal.2004754 (2014).

44 Levy, C., Khaled, M. & Fisher, D. E. MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol Med* **12**, 406-414, doi:10.1016/j.molmed.2006.07.008 (2006).

45 Pastore, N. *et al.* TFE3 regulates whole-body energy metabolism in cooperation with TFEB. *EMBO molecular medicine* **9**, 605-621, doi:10.15252/emmm.201607204 (2017).

46 Najibi, M., Labed, S. A., Visvikis, O. & Irazoqui, J. E. An Evolutionarily Conserved PLC-PKD-TFEB Pathway for Host Defense. *Cell reports* **15**, 1728-1742, doi:10.1016/j.celrep.2016.04.052 (2016).

47 Pastore, N. *et al.* TFEB and TFE3 cooperate in the regulation of the innate immune response in activated macrophages. *Autophagy* **12**, 1240-1258, doi:10.1080/15548627.2016.1179405 (2016).

48 Visvikis, O. *et al.* Innate host defense requires TFEB-mediated transcription of cytoprotective and antimicrobial genes. *Immunity* **40**, 896-909, doi:10.1016/j.immuni.2014.05.002 (2014).

49 Pastore, N. *et al.* Nutrient-sensitive transcription factors TFEB and TFE3 couple autophagy and metabolism to the peripheral clock. *EMBO J*, doi:10.15252/embj.2018101347 (2019).

50 Steingrimsson, E., Copeland, N. G. & Jenkins, N. A. Melanocytes and the microphthalmia transcription factor network. *Annual review of genetics* **38**, 365-411, doi:10.1146/annurev.genet.38.072902.092717 (2004).

51 Leeman, D. S. *et al.* Lysosome activation clears aggregates and enhances quiescent neural stem cell activation during aging. *Science* **359**, 1277-1283, doi:10.1126/science.aag3048 (2018).

52 Betschinger, J. *et al.* Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153**, 335-347, doi:10.1016/j.cell.2013.03.012 (2013).

53 Kyrmizi, I. *et al.* Plasticity and expanding complexity of the hepatic transcription factor network during liver development. *Genes & development* **20**, 2293-2305, doi:10.1101/gad.390906 (2006).

54 Inoue, Y., Inoue, J., Lambert, G., Yim, S. H. & Gonzalez, F. J. Disruption of hepatic C/EBPalpha results in impaired glucose tolerance and age-dependent hepatosteatosis. *The Journal of biological chemistry* **279**, 44740-44748, doi:10.1074/jbc.M405177200 (2004).

55 Antoniou, A. *et al.* Intrahepatic bile ducts develop according to a new mode of tubulogenesis regulated by the transcription factor SOX9. *Gastroenterology* **136**, 2325-2333, doi:10.1053/j.gastro.2009.02.051 (2009).

56 Michalopoulos, G. K. Principles of liver regeneration and growth homeostasis. *Compr Physiol* **3**, 485-513, doi:10.1002/cphy.c120014 (2013).

57 Kauffman, E. C. *et al.* Molecular genetics and cellular features of TFE3 and TFEB fusion kidney cancers. *Nature reviews. Urology* **11**, 465-475, doi:10.1038/nrurol.2014.162 (2014).

58 Thomsen, M. K. *et al.* SOX9 elevation in the prostate promotes proliferation and cooperates with PTEN loss to drive tumor formation. *Cancer research* **70**, 979-987, doi:10.1158/0008-5472.CAN-09-2370 (2010).

59 Ling, S. *et al.* An EGFR-ERK-SOX9 signaling cascade links urothelial development and regeneration to cancer. *Cancer research* **71**, 3812-3821, doi:10.1158/0008-5472.CAN-10-3072 (2011).

60 Matheu, A. *et al.* Oncogenicity of the developmental transcription factor Sox9. *Cancer research* **72**, 1301-1315, doi:10.1158/0008-5472.CAN-11-3660 (2012).

61 Strick-Marchand, H. & Weiss, M. C. Inducible differentiation and morphogenesis of bipotential liver cell lines from wild-type mouse embryos. *Hepatology* **36**, 794-804, doi:10.1053/jhep.2002.36123 (2002).

62 Palmer, D. & Ng, P. Improved system for helper-dependent adenoviral vector production. *Molecular therapy : the journal of the American Society of Gene Therapy* **8**, 846-852 (2003).

63 Terashima, T. *et al.* DRG-targeted helper-dependent adenoviruses mediate selective gene delivery for therapeutic rescue of sensory neuronopathies in mice. *The Journal of clinical investigation* **119**, 2100-2112 (2009).

64 Seglen, P. O. Preparation of isolated rat liver cells. *Methods in cell biology* **13**, 29-83 (1976).

65 Gundry, M. C. *et al.* Highly Efficient Genome Editing of Murine and Human Hematopoietic Progenitor Cells by CRISPR/Cas9. *Cell reports* **17**, 1453-1461, doi:10.1016/j.celrep.2016.09.092 (2016).

66 Brunetti, L., Gundry, M. C., Kitano, A., Nakada, D. & Goodell, M. A. Highly Efficient Gene Disruption of Murine and Human Hematopoietic Progenitor Cells by CRISPR/Cas9. *Journal of visualized experiments : JoVE* **134**, doi:doi:10.3791/57278 (2018).

67 Moreno-Mateos, M. A. *et al.* CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nature methods* **12**, 982-988, doi:10.1038/nmeth.3543 (2015).

68 Pastore, N. *et al.* TFE3 regulates whole-body energy metabolism in cooperation with TFEB. *EMBO molecular medicine*, doi:10.15252/emmm.201607204 (2017).

69 Yaylaoglu, M. B. *et al.* Comprehensive expression atlas of fibroblast growth factors and their receptors generated by a novel robotic in situ hybridization platform. *Developmental dynamics : an official publication of the American Association of Anatomists* **234**, 371-386, doi:10.1002/dvdy.20441 (2005).

70 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

71 Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).

72 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

73 Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264, doi:10.1093/biostatistics/4.2.249 (2003).

74 Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307-315, doi:10.1093/bioinformatics/btg405 (2004).

75 Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids research* **30**, 207-210 (2002).

76 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).

77 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

78 Palmieri, M. *et al.* Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways. *Hum Mol Genet* **20**, 3852-3866, doi:10.1093/hmg/ddr306 (2011).

79 Wasserman, W. W. & Sandelin, A. Applied bioinformatics for the identification of regulatory elements. *Nature reviews. Genetics* **5**, 276-287, doi:10.1038/nrg1315 (2004).

80 Jarrett, K. E. *et al.* Somatic Editing of Ldlr With Adeno-Associated Viral-CRISPR Is an Efficient Tool for Atherosclerosis Research. *Arterioscler Thromb Vasc Biol* **38**, 1997-2006, doi:10.1161/ATVBAHA.118.311221 (2018).

81 Lagor, W. R., Johnston, J. C., Lock, M., Vandenberghe, L. H. & Rader, D. J. Adeno-associated viruses as liver-directed gene delivery vehicles: focus on lipoprotein metabolism. *Methods Mol Biol* **1027**, 273-307, doi:10.1007/978-1-60327-369-5_13 (2013).

Figure legends

Figure 1: TFEB expression is enriched in ductal/progenitor cells.

(**a**) Immunohistochemistry analysis of TFEB in wild-type (WT) liver at the indicated stages showing prominent signaling in bile ductules. Arrows indicate ductal cells. PV=portal vein. Scale bar 20 μm. (**b,c**) Representative immunofluorescence stains for TFEB/HNF4α (**b**) and TFEB/SOX9 (**c**) showing TFEB levels in the central (CV) and in the portal (PV) vein area. Scale bar 20 μm.

Figure 2: TFEB influences hepatoblast differentiation *in vitro*.

(**a**) Hepatocyte sphere formation of HBs of the indicated genotypes 3 days after differentiation. (**b**) mRNA levels of the indicated genes were quantified by quantitative RT-PCR of total RNA isolated from control (CTRL), TFEB overexpressing ($TFEB^{OE}$) and TFEB depleted ($TFEB^{KO}$) HBs undifferentiated (collagen coated plates) or after 5 days of hepatocyte differentiation (uncoated plates). Values are indicated as mean ± SEM of n=3 and expressed as fold difference compared with CTRL undifferentiated HBs. (**c**) Immunoblotting analysis and relative quantification of the precursor/cholangiocyte marker SOX9 and the hepatocyte marker HNF4α in $TFEB^{KO}$ and $TFEB^{OE}$ HBs after 5 days of hepatocytic differentiation. (**d**) Immunostaining for the indicated markers on $TFEB^{KO}$ and $TFEB^{OE}$ HBs 5 days after hepatocytic differentiation. Scale bar 5 μm. (**e**) Gene expression profiling of hepatocyte and biliary genes of differentiated HBs of the indicated genotypes. (**f**) Flow cytometry analysis of the cell cycle distribution of undifferentiated CTRL, $TFEB^{KO}$ and $TFEB^{OE}$ HBs and relative quantification (n=3). Data are represented as mean ± SEM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ two-tailed Student's t-test.

Figure 3: TFEB overexpression inhibits hepatocyte differentiation *in vivo*.

(**a**) KEGG pathways enriched in upregulated and downregulated genes of P9 liver overexpressing TFEB compared to CTRL. (**b**) Heat map showing the relative fold change of hepatocyte and biliary genes. (**c**) RT-PCR analysis of hepatocyte- and cholangiocyte- specific markers in liver isolated from P0 and P9 mice (n=3). The dashed line represents gene expression levels in CTRL liver. Data are represented as mean ± SEM. *$p \leq 0.05$, **$p \leq 0.01$ two-tailed Student's t-test. (**d**) Immunoblotting analysis of TFEB, HNF4α and SOX9 in liver extracts from mice at P9. Actin was used as loading control. (**e**) Left panel: Representative images of liver sections from P0 and P9 mice stained for CK19. Scale bar 50 μm. Right panels: triple immunostaining for HNF4α (red), SOX9 (green) and CK19 (white) in liver sections from P0 and P9 mice of the indicated genotype showing the hybrid characteristics of hepatocytes and cholangiocytes in Tg mice. Note cells expressing SOX9 and CK19 (box1), SOX9 and HNF4α (box2) or SOX9/HNF4α/CK19 (box3). Scale bar 20 μm.

Figure 4: TFEB directly controls *Sox9* expression.

(**a**) ChIP analysis of TFEB binding to the *Sox9* promoter in Tg livers. CLEAR sites in the promoter region of *Sox9* are indicated by red squares. Bar graphs show the amount of immuno-precipitated DNA as detected by RT-PCR analysis (*n*=2 independent experiments). Values were normalized to the input and plotted as relative enrichment over the control sequence. (**b**) Luciferase expression upon TFEB overexpression was measured in HBs infected with HDAd-LacZ (CTRL) or HDAd-CRE expressing viruses ($TFEB^{OE}$) for 24 hours and transfected with wild type or mutated *Sox9*-promoter luciferase reporter plasmids (*n*=3). Deletion of the CLEAR sites resulted in reduced luciferase transactivation. (**c**) RT-PCR expression analysis of hepatocyte- and cholangiocyte- specific markers in livers isolated from P9 mice for the indicated genotypes (*n*=3). (**d**) Representative immunoblot for SOX9 in liver of the indicated genotypes and relative quantification. (**e**) CK19 immunostaining of liver sections from P9 mice of the indicated genotypes. Scale bar 50 μm. Data are represented as mean ± SEM. **p* ≤ 0.05, ***p* ≤ 0.01, ****p* ≤ 0.001, *****p* ≤ 0.0001 two-tailed Student's t-test.

Figure 5: TFEB depletion impairs ductular reaction and cell proliferation after liver injury.

(**a**) Experimental strategy for liver injury protocol for $TFEB^{LiKO}$ mice. (**b,c**) Expression levels of *Sox9* and *Krt19* mRNA (*n*=5) at different time points (**b**) and immunoblotting with relative quantification (*n*=4) (**c**) of livers isolated from $TFEB^{f/f}$ and $TFEB^{LiKO}$ mice 14 days after discontinuation of DDC food. Values are expressed as fold difference relative to $TFEB^{f/f}$ mice. (**d**) SOX9 and CK19 immuno-fluorescence with relative quantifications (*n*=3) of liver isolated from $TFEB^{f/f}$ and $TFEB^{LiKO}$ mice 14 days after recovery and relative quantifications. Scale bar 50 μm. (**e**) HNF4α/SOX9 dual staining of liver sections from $TFEB^{f/f}$ and $TFEB^{LiKO}$ mice. Scale bar 5 μm. (**f,g**) Quantification of $BrdU^{+}$ cells (**f**) and representative images (**g**). Scale bar 50 μm. Data are represented as mean ± SEM. **p* ≤ 0.05, ***p* ≤ 0.01 two-tailed Student's t-test.

Figure 6: TFEB overexpression in progenitor/cholangiocytes gives rise to ectopic ductal structures.

(**a**) Representative images of CK19/tdTom dual immunofluorescence of CTRL and $TFEB^{OE}$ liver four-weeks after tamoxifen injection. Scale bar 50 μm. (**b**) Quantification of $CK19^{+}$ biliary epithelial cells that were $tdTom^{+}$ (*n*=3 mice analyzed, 5 sections per mouse). Data are represented as mean ± SEM. ***p* ≤ 0.01 two-tailed Student's t-test. (**c**) Experimental strategy to lineage trace biliary epithelial cells on a background of liver injury. (**d**) tdTom/HNF4α, tdTom/SOX9 and tdTom/CK19 Immunofluorescent images at 14 days after DDC diet. Scale bar 50 μm.

Figure 7: TFEB-overexpressing hepatocytes dedifferentiate in progenitor/cholangiocyte-like cells upon liver injury.

(a) tdTom/CK19 and tdTom/HNF4α immunofluorescence in AAV8-TBG-CRE-treated livers 4 weeks after injection. CRE-treated HNF4α$^+$ hepatocytes are tdTom$^+$ in contrast to CK19$^+$ ductal cells located in the portal tract (PV). Insets: images of AAV-CTRL injected liver indicated as negative control (NC). Scale bar 50 μm. (**b**) Transcript levels of *Tcfeb* in livers isolated from AAV8-TBG-CRE treated mice 4 weeks after injection. *Tcfeb* mRNA levels were normalized to the CTRL. Data are represented as mean ± SEM. ****$p \leq 0.0001$ two-tailed Student's t-test. (**c**) CK19/HNF4α immunofluorescence analysis in AAV8-TBG-CRE-treated and CTRL liver 4 weeks after injection showing no alteration in cell identity. Scale bar 50 μm. (**d**) Experimental design of the AAV8 injection followed by acute liver injury protocol. (**e**) 14-days post DDC acute injury hepatocytes dedifferentiate in progenitor/biliary cells. (**f**) Experimental design of the AAV8 injection followed by chronic liver injury protocol (**g**) Gross appearance of AAV-injected livers from day 14 of recovery after DDC prolonged exposure. (**h**) Immunofluorescence analysis of AAV8-treated livers at day 14 of recovery after DDC diet. Scale bar 50 μm.

Figure 8: Bile duct neoplasm in TFEB overexpressing liver.

(**a**) Retrograde injection of ink revealed abnormal development of the bile duct in Tg liver compared to control. (**b**) Gross appearance of livers in Tg mice at several ages. (**c**) Quantification of liver-to-body weight ratio (LW/BW) in Tg mice at the indicated age (n=3-15). Values were normalized to age-matched control liver as indicated by the dashed line. Data are represented as mean ± SEM. ***$p \leq 0.001$, ****$p \leq 0.0001$ two-tailed Student's t-test. (**d**) Immunohistochemistry and *in-situ* hybridization analysis for TFEB in CTRL and Tg liver at 3 months of age. Scale bar 50 μm. (**e**) Histological characterization of Tg liver phenotype at 3 months of age. Lower panel: Ki67 staining and relative quantification showing increased cell proliferation in Tg liver compared to CTRL. Scale bar 100 μm.