# Lineage tracing : efficient tools to determine the fate of hepatic cells in health and disease.

Prof. Frédéric Lemaigre, MD, PhD (*)

(*) Corresponding author

Université catholique de Louvain, de Duve Institute, Avenue Hippocrate 75, B1-7503

1200 Brussels, Belgium

email: frederic.lemaigre@uclouvain.be

Phone: +32 2 764 75 83  Fax: +32 2 764 75 73

**Keywords**: differentiation, recombinase, single-cell labeling, cell fate determination

**Abstract**

Hepatic cells undergo profound phenotypic changes during development, chronic injury, regeneration and tumor formation. Several lineage tracing methods have been developed to determine the fate of hepatic and other cell types in normal and disease conditions. They are based on prospective chemical, radioactive or genetic labeling of a precursor population, followed by the detection of the label in the progeny. The advent of single-cell analysis adds a new dimension in our capacity to trace the fate of liver cells, including for the retrospective identification of lineages. Here we describe the main lineage tracing methods that were applied to liver biology and discuss novel methods that may help solving open questions. We also examine the strengths and limitations of the techniques in order to provide the reader with the information required to select the most appropriate approach in future studies and critically assess published data.

Hepatic cells undergo extensive changes in form and function during development, injury, regeneration and tumorigenesis. Such changes occur in normal development by differentiation of progenitor cells to descendants that progressively mature to fully functional adult liver cells. Following acute and chronic injury or during tumor formation, adult cells may dedifferentiate, *i.e.* lose their mature phenotype and revert to a progenitor-like state. Adult cells can also transdifferentiate by switching their phenotype to that of another cell type, and this might include a transient dedifferentiated state [1]. Therefore, when confronted with the phenotype of cells at a specific stage of development or disease, questions are raised about the cell type of origin and the future destiny of the studied cells. A number of lineage tracing approaches have been designed to address this issue, each with their strengths and limitations, and risks of misinterpretation [2, 3]. Here we discuss the technical and conceptual approaches for lineage tracing cells in liver development, homeostasis and disease. The aim of the chapter is to provide the readers with background information required to select the optimal strategy in future studies and to critically assess the data summarized in other sections of the textbook.

**Hepatic cell fate determination by chemical labeling of cells**

In the label retaining assay (Fig. 1a), cells are pulse-labeled using a radioactive or chemical tracer such as tritiated thymidine or 5-bromo-2'-deoxyuridine (BrdU) which becomes incorporated in DNA during cell division, in order to follow the fate of dividing cells that are marked by the tracer. This approach was used to locate putative stem cells: asymmetric stem cell division generates two daughter cells, namely a new stem cell and a transit amplifying cell. The label is diluted when the

amplifying cell proliferates, but remains detectable in the stem cell which resides in its niche. This technique suffers from a number of limitations. It critically lacks cell specificity, and the decrease in immunodetectability of BrdU as a result of proliferation-induced dilution of the label does not allow to faithfully identify the stem cell's progeny beyond the third division. Moreover, BrdU and tritiated thymidine can affect cell metabolism, and either promote or inhibit proliferation, or trigger apoptosis. BrdU can even affect differentiation [4]. For these reasons, the label retaining assay is usually replaced by potentially less toxic and more specific methods.

In a different experimental setting, vital dyes were used to label cell patches in the embryo to determine their positional fate during liver development. Chloromethybenzamido-1-1-dioctadecyl-3,3,3V,-3V-tetramethylindocarbocyanine (Dil) can be dropped on cells and fluoresces once incorporated into membranes. Accurately targeting narrow areas of mouse embryos at embryonic day 8 with Dil revealed that three distinct endodermal regions coalesce to constitute the liver bud around E9-E10 [5] (Fig. 1b). This approach, which implements the concept of fate mapping by direct observation in the study of liver development, is useful to provide dynamical information about the location and migration of cells. However, the gained information is limited in time due to the dilution of the membrane-associated dye following cell division and it does not consider the differentiation status of the labeled cells.

**Genetic marks to determine cell fate**

Given the limitation associated with dilution of chemical or radioactive labels, it appeared more efficient to trace cells harboring mutations that are transmitted and

permanently detectable in the cells' progeny. Tracing cells for mosaic expression of a transgene or mutated gene, or for random inactivation of an X-linked gene, enabled to identified patches of mutated cells surrounded by non-mutated cells within developing tissues or liver regenerative nodules [6-8]. While the results were highly suggestive of clonality of the mutated cells, the origin of these cells cannot be identified with certainty, and cells which are more motile and do not remain clustered within the tissue cannot reliably be assigned to a single clone or lineage.

The use of liver chimeras is a variant of genetic marking. In a classical study, dipeptidylpeptidase IV (DPP-IV)-positive hepatocytes were administered to DPP-IV-negative rats subjected to partial hepatectomy and retrorsine-induced inhibition of hepatocyte proliferation. DPP-IV-positive hepatocytes integrated in the liver, thereby generating chimeric livers which were subsequently subjected to bile duct ligation to stimulate ductular proliferation. DPP-IV-expressing cells were eventually detected in the biliary epithelium of the recipient livers [9]. Along the same lines, oval cells purified from fumarylacetoacetate hydrolase (Fah)-positive livers were able to repopulate Fah-negative livers and to convert to hepatocytes [10]. These two studies provided insight into the transdifferentiation capacity of hepatocytes and oval cells, but their conclusions were debated, as the possibility was raised that the normal differentiation potential of the cells could have been altered by the purification and reimplantation conditions. An additional caveat is that fusion between the implanted cells and cells of the recipient liver may confound the interpretation of the data. Together, the limitations of the genetic tracing techniques illustrated here prompted researchers to further resort to *in situ* genetic marking of cells.

**Labeling of cells by genetic recombination**

The use of bacteriophage cyclization recombinase (Cre) to genetically mark cells and their progeny has rapidly become the gold standard in lineage tracing experiments. In this setting the genetically marked cell population contains two transgenes, the first coding for Cre or a variant of Cre, and the second coding for a reporter protein that is expressed only after Cre-mediated excision of a DNA fragment containing a transcriptional stop cassette fllanked by 34-base pair-long locus of X-over P1 (*loxP*) sequences that are recognized by Cre. Thus the expression of the reporter is induced in the targeted cells and persists in their descendants, whether or not they express Cre, since the Cre-mediated genetic elimination of the stop cassette is irreversible and heritable. Therefore, the expression of the reporter serves as a marker of the targeted cells and their progeny, and is expected to permanently label the lineage. Yet, one must keep in mind that tumor formation may be associated with chromosomal rearrangements that affect the expression of the marker gene, thereby breaking the rule of permanent labeling of the lineage. This elegant approach permitted to uncover key aspects related to the source and phenotypic evolution of liver cells during development and disease. However, its implementation turned out to be more complex than anticipated and has led to a number of misinterpretations due to unexpected difficulties associated both with the production of Cre in the desired precursor cell population and with the expression of the reporter protein.

The ideal strategy for expression of Cre implies that Cre is exclusively expressed and active in the targeted precursor cells. Since the latter differentiate to their progeny, the duration of Cre expression and/or activity must be strictly controlled to avoid that excision of the reporter's stop cassette also occurs in the descendant cells on alleles

that escaped recombination at the precursor stage. In other terms, Cre activity must be tightly controlled in space and time.

The control of Cre expression in space is usually achieved by constructing transgenes that drive Cre transcription under control of transcriptional regulatory regions that are specifically active in the precursor cell population. A list of Cre lines available for lineage tracing studies in liver has been published [2]. They bear transgenes which drive Cre expression under control of regulatory regions active in hepatocytes (*Albumin*, *Cyp1A1* (*Ah*), *Mx1*, *Transthyretin*, *$\alpha$-fetoprotein*), cholangiocytes (*Ck19*, *Hnf1$\beta$*, *Osteopontin*, *Sox9*, *Prominin1*), stellate cells (*Gfap*, *Col$\alpha$2*, *Lrat*, *Pdgfr$\beta$, $\alpha$Sma*), macrophages (*Fsp1*), mesothelium (*Wt1*, *MesP1*) or undifferentiated progenitors (*Lgr5*, *FoxL1*). Importantly, most of the listed regulatory regions are not as cell type-specific as initially hoped for and display leakiness. For instance, the *Albumin-Cre* and *Ah-Cre* transgenes drive expression of Cre in hepatocytes, as expected, but also in a subset of cholangiocytes; *mGFAP-Cre* produces Cre in stellate cells, but also in cholangiocytes. In addition, the tissue-specificity of the Cre-driving regulatory regions may vary with time. For instance the *Sox9-CreER* transgene is specifically transcribed in developing cholangiocytes of embryonic liver, but after birth it is functional in cholangiocytes and in a subset of periportal hepatocytes [11]. Therefore, when selecting a Cre transgene supposedly expressed in a well-defined cell population, one is often confronted with insufficient cell specificity, leading to labeling of unintended cell types and confound interpretation of lineage tracing data.

An elegant dual recombinase system has been designed to more accurately label the desired cell population in liver while inhibiting the labeling of an undesired cell type

[12, 13]. In this system the mice harbor three transgenes, namely a reporter construct bearing coding sequences for two fluorescent reporters (tdTomato and ZsGreen) and two transgenes coding for Dre recombinase or Cre recombinase (Fig. 1c). Similar to Cre recombinase, Dre eliminates DNA regions flanked by *rox* sequences. Labeling of an undesired cell population by CreER recombinase (a version of Cre whose activity is controlled by tamoxifen; see also below) is prevented if Dre recombinase has previously excised the reporter construct in the undesired cell type and if the Dre-mediated excision concomitantly renders the reporter structurally unresponsive to CreER. This approach has for instance been implemented to demonstrate that hepatocytes can convert to biliary-like cells following liver injury, and that the traced cell population was not contaminated by inappropriate labeling of cholangiocytes which might have been the source of the new biliary-like cells formed during the disease process. The success of this elegant approach requires high efficiency of both the Dre and Cre recombinases, and implies that the undesired cell type is identified prior to the lineage tracing of the desired cell type.

Four procedures can be followed to achieve temporal control of Cre activity. The first and most classical one resorts to the use of tamoxifen-inducible CreER which consists in a fusion of Cre with a mutated hormone binding domain of the estrogen receptor (ER). The latter is insensitive to estradiol but has high affinity for tamoxifen which promotes nuclear translocation and functional activation of CreER. Variants of CreER have been designed and CreER$^{T2}$, which contains a triple mutation of human ER hormone binding domain, is the most popular and most sensitive to tamoxifen [14]. CrePR is a less often selected alternative; it was made by fusing Cre recombinase with a mutated progesterone receptor (PR) ligand binding domain, which is responsive to the synthetic steroid RU486 [15]. A second approach consists

in inducing the activity of the promoter driving Cre transcription, which in liver was performed by administering β-naphtoflavone or poly(I:C) to induce transcription of the *Ah*- or *Mx1*-driven Cre transgenes, respectively. Third, tight temporal control of Cre expression can also be achieved by means a doxycyclin-regulated system. Here, Cre expression is driven by a promoter activated by tTA or rtTA, which is produced from a transgene and which simulates Cre transcription in the absence or presence of doxycyclin, respectively [16]. In a fourth setting, the temporally-controlled onset of Cre expression can be achieved by vector-mediated injection of transgenes in adult animals. In this case cell-type specificity is ensured by two control mechanisms: the transgene bears a cell type-specific promoter driving Cre expression and is inserted in a vector targeting a specific liver cell type. Good examples are the use of Adenoviruses or of the Adeno-associated virus 2/8-TTR-Cre vector which specifically infects hepatocytes and induces expression of Cre under control of the transthyretin promoter [17]. Alternatively, hydrodynamic injection of plasmid DNA coding for Cre into the tail vein will preferentially target hepatocytes but no other liver cell type [18, 19].

The temporal control of CreER activity by tamoxifen is very appealing, yet its use requires much care. Too low doses of tamoxifen are associated with insufficient activation of CreER and labeling of only a small fraction of the CreER-expressing cells. Still, one can take advantage of low dose-related efficiency to perform single-cell labeling and clonal analysis. For instance, Kamimoto and coworkers administered low doses of tamoxifen to *Prom1-CreER$^{T2}$;R26R-tdTomato* mice and labeled 0.2% of cholangiocytes. The mice were then subjected to a hepatotoxin to induce a ductular reaction and to investigate the three dimensional architecture of single-cell-derived clonal progenies [20]. Higher doses of tamoxifen are required to

efficiently activate CreER and homogeneously label a specified population, but tamoxifen being a mixed agonist/antagonist of the estrogen receptor, it is toxic during development and its administration in mice beyond E14 often leads to abortion. Moreover, tamoxifen administration is potentially noxious in adults as well since it increases serum transaminase levels [21] and induces ectopic *Sox9* expression in hepatocytes [11]. It also came as a surprise that tamoxifen appeared to be long-lived in the organism and able to induce detectable CreER activation up to four weeks after injection, albeit at low efficiency [22]. This is a potential source of misinterpretation of lineage tracing data, in particular when investigating the fate of cells in experimentally-induced disease conditions that modify the cell-specificity of CreER expression: in case CreER is expressed as anticipated in the specified cells, but subsequently also induced in an unintended lineage as a result of injury, CreER will become activated in the specified cells but also in the intended lineage if tamoxifen is persisting in the organism. Therefore, a one-month wash-out period is required between tamoxifen activation of CreER in the specified population and the induction of disease. Studies that have not applied the precautionary principle of a wash-out period should be evaluated with care.

The structure of the transgene bearing Cre recombinase can also lead to misleading results. While most constructs used for liver studies and expressing Cre or CreER were introduced in the animal models as conventional transgenes or as bacterial artificial chromosomes (BAC), knockin strategies have been selected in a number of studies. Mice that express CreER from *IRES-CreER* coding sequences knocked into the 3' UTR of the *Sox9* locus (*Sox9-IRES-CreER*) generated results different from those that produce CreER from a transgenic BAC construct. In the latter, mice harbor exogenous *Sox9-CreER* sequences in the BAC, leaving the two endogenous *Sox9*

alleles of the recipient genome unaltered. Several groups found that cholangiocytes do not participate to hepatocyte replacement in normal homeostatic liver, supporting the findings obtained with lineage tracing experiments performed with the *Sox9-CreER* BAC construct. Instead, continuous supply of hepatocytes by cholangiocytes was monitored in homeostatic livers using *Sox9-IRES-CreER* as a tool to lineage trace the cholangiocytes. The discrepancy might result from the fact that inserting *IRES-CreER* in the *Sox9* locus reduces the expression of endogenous *Sox9*, potentially causing abnormal behavior of the cholangiocytes [11, 23, 24]. In this context, and again in contrast to the proposal that cholangiocytes supply homeostatic liver with hepatocytes, recent lineage tracing experiments using Axin-CreER$^{T2}$ or Tert-CreER$^{T2}$ provided evidence that homeostatic hepatocyte renewal originates from pericentral hepatocytes expressing Axin2 and from widely distributed hepatocytes displaying high levels of telomerase activity [25, 26].

Finally, it is important to mention that viral vectors delivering Cre recombinase to the liver can cause adverse effects. Thus adenoviral vectors can stimulate inflammatory responses in the liver, thereby altering the fate choice of cells. Also, other complications might arise when viral vectors integrate into gene promoters or transcriptional units, thereby increasing the risk of deregulation of normal cell fate decisions.

**Reporter transgenes for genetic recombination**

The most common reporter transgenes for lineage tracing code for a fluorescent protein under control of the ubiquitously transcribed ROSA locus and are preceded by a *loxP*-flanked stop cassette. Sequence variants for *loxP* have been developed

with mutations that prevent recombination with the canonical *loxP* sites but allow recombination with *lox* site variants harboring identical mutations [27, 28]. *rox* and *frt* sites are alternatives that are recombined by Dre and FLP recombinase, respectively; their availability makes it possible to design experiments with dual recombination for accurate cell-specific labeling as described above (Fig. 1c) [12], or to sequentially recombine *loxP*- and *frt*-flanked genes as in the work by Schaub and coworkers [29]. This group knocked out *Rbpj* and *Hnf6* in developing livers using *Alb-Cre;Rbpj*$^{loxP/loxP}$*;Hnf6*$^{loxP/loxP}$*;R26ZG* mice; this impairs bile duct development, leading to bile duct paucity in newborns as seen in Alagille syndrome. Subsequently the mice were injected at the adult stage with a AAV8-Ttr-FLP virus to induce *frt*-flanked green fluorescent protein from the *R26ZG* allele specifically in hepatocytes; compensatory bile duct development in adults resulted from hepatocyte-to-cholangiocyte transdifferentiation as demonstrated by the fluorescent green labeling of the newly formed bile ducts.

Several variants of fluorescent proteins are available [30, 31]. Fluorescent reporters are preferred over β-galactosidase, since the latter cannot be detected by flow cytometry. However, not all fluorescent proteins are appropriate for flow cytometry: tdTomato is the brightest and is easily visualized, but its epifluorescence can bleed in other detection channels and interfere with co-detection of other markers. A number of lineage tracing experiments resorted to clever combinations of fluorescent reporters. The dual recombinase system mentioned above is again a good example [12]. Along the same lines, Iverson and coworkers measured the hepatocyte birthrate using a two color fluorescent marker gene that switches from expressing membrane-targeted red fluorescent protein tdTomato to expressing membrane-targeted green fluorescent protein, following Cre-mediated excision of the *loxP*-flanked mTomato

sequence [32] (Fig. 1d). Perhaps the most sophisticated combination of fluorescent markers was achieved by generating multicolor reporters in which Cre recombinase excises as stop cassette and stochastically places one of four fluorescent reporters under control of a promoter [33, 34] (Fig. 1e). Using such strategy, clonal progeny can be visualized in the liver: Tarlow and coworkers used *Sox9-CreER$^{T2}$*;*ROSA26-confetti* to track clonality of oval cells in liver injury models [22]. Similarly, Schaub and coworkers performed clonal analysis to demonstrate that bile duct reconstitution in their model of Alagille syndrome resulted from transdifferentiation of a large number of hepatocytes and not from transdifferentiation and proliferation of a small subset of hepatocytes [29].

**Single-cell analysis**

Single-cell methods have considerably expanded our potential to investigate gene expression during cell fate changes [35]. Single-cell RNA sequencing (scRNAseq) followed by classification of the cells based on their transcriptomic profile can identify subgroups of cells within a heterogeneous population. Since cell differentiation is a continuous process transcriptome analysis of individual cells within a differentiating population is expected to reveal a continuous spectrum of states that represents the sequential steps in cell fate progression. Algorithms referred to as pseudotime methods were designed to order the cells according to their transcriptomic similarities which reflect their progressing differentiation status and recapitulate their developmental trajectory [36]. Within developmental trajectories, the detection of initiation states and end-points, and of transient states and branching decisions allows to propose lineage trees that are based on individual cells' expression profiles

[37, 38]. Moreover, investigation of gene expression fluctuations during development of a lineage can identify candidate regulators of differentiation as well as gene regulatory networks driving cell fate determination [39].

Strictly speaking, pseudotime methods do not provide definitive proof that a progeny derives from the proposed precursor. They are based on statistical analyses of whole populations and lack the functional dynamics associated with labeling of a specific cell type, followed by visual detection and spatial location of the progeny. Yet, in its current state, pseudotime analysis implemented in the study of liver development confirmed and extended the conclusions from tracing experiments performed by standard cell labeling procedures. Indeed, genetic labeling of liver progenitors has provided evidence that hepatocytes and cholangiocytes derive from common progenitors [40], and several teams have shown that hepatoblasts give rise to cholangiocytes when exposed to cholangiocyte-inducing signaling [41]. Single-cell transcriptomic analyses were performed on mouse fetal liver cells that were sorted on the basis of their expression of Delta-like 1 (Dlk1), which marks hepatoblasts throughout gestation, and Epithelial Cell Adhesion Molecule (EpCAM), a marker of early hepatoblasts and developing cholangiocytes [42, 43]. The results confirmed the linear and continuous differentiation of hepatoblasts to hepatocytes, and the generation of cholangiocytes through a branched differentiation pathway emerging from the hepatoblast population around embryonic day 14.5.

The limitations of pseudotime analysis will likely be overcome by combining spatial transcriptome data with scRNA-seq, or by scRNA-seq of cells whose precursors have been genetically modified by introducing barcodes in their genome. The barcoded genome can be detected by genomic sequencing, or scRNA-seq in case the barcode is transcribed. An example of the latter is provided by CRISPR-Cas9-

induced mutational barcoding of cells at an early stage of Zebrafish development, followed by single cell transcriptomic analysis at a later stage; this identified lineages of cells, including liver cells, expressing the barcode [44, 45]. The use of barcoding technology for tracing liver cells is still in its infancy, but has been successfully implemented in other organs where the strengths and limitations of such approaches are currently emerging [45, 46].

Finally, the lineage tracing approaches described so far prospectively introduced marks in a precursor cell population, followed by detection of the mark in the progeny. Retrospective lineage tracing, *i.e.* inferring past developmental relationships on the basis of the analysis of progeny, is particularly important when investigating diseased tissues. Single-cell detection of lineage marks such as somatic mutations, copy-number variants, single-nucleotide variants and microsatellites can now be used for reconstructing the genealogy of cell populations [46]. Understanding liver disease is likely to rely heavily on the implementation of such approaches in the near future.

**Acknowledgements**

Work by the author is supported by the D.G. Higher Education and Scientific Research of the French Community of Belgium (ARC 15/20-065), the F.R.S.-FNRS (Belgium; grants T.007214 and J.0037.17), the Belgian Foundation Against Cancer (grant 2014-125), the Fonds Spéciaux de Recherche (Université catholique de Louvain), the Fonds pour la Formation à la Recherche dans l'Industrie et l'Agriculture (Belgium; grant 1.E071.18), the Fondation Maisin (Belgium), and Innovation for Liver

Tissue Engineering (iLITE; Recherche Hospitalo-Universitaire-Assistance Publique-Hopitaux de Paris, France).

## References


1. Merrell, A.J. and Stanger, B.Z. (2016) Adult cell plasticity in vivo: de-differentiation and transdifferentiation are back in style. *Nat Rev Mol Cell Biol,* **17**, 413-25.
2. Lemaigre, F.P. (2015) Determining the fate of hepatic cells by lineage tracing: facts and pitfalls. *Hepatology,* **61**, 2100-3.
3. Kretzschmar, K. and Watt, F.M. (2012) Lineage tracing. *Cell,* **148**, 33-45.
4. Evarts, R.P., Hu, Z., Omori, N. *et al.* (1996) Precursor-product relationship between oval cells and hepatocytes: comparison between tritiated thymidine and bromodeoxyuridine as tracers. *Carcinogenesis,* **17**, 2143-51.
5. Tremblay, K.D. and Zaret, K.S. (2005) Distinct populations of endoderm cells converge to generate the embryonic liver bud and ventral foregut tissues. *Dev Biol,* **280**, 87-99.
6. Roth, K.A., Hermiston, M.L. and Gordon, J.I. (1991) Use of transgenic mice to infer the biological properties of small intestinal stem cells and to examine the lineage relationships of their descendants. *Proc Natl Acad Sci USA,* **88**, 9407-11.
7. Fellous, T.G., Islam, S., Tadrous, P.J. *et al.* (2009) Locating the stem cell niche and tracing hepatocyte lineages in human liver. *Hepatology,* **49**, 1655-63.
8. Paradis, V., Dargere, D., Bonvoust, F. *et al.* (2000) Clonal analysis of micronodules in virus C-induced liver cirrhosis using laser capture microdissection (LCM) and HUMARA assay. *Lab Invest,* **80**, 1553-9.
9. Michalopoulos, G.K., Barua, L. and Bowen, W.C. (2005) Transdifferentiation of rat hepatocytes into biliary cells after bile duct ligation and toxic biliary injury. *Hepatology,* **41**, 535-44.
10. Wang, X., Foster, M., Al-Dhalimy, M. *et al.* (2003) The origin and liver repopulating capacity of murine oval cells. *Proc Natl Acad Sci U S A,* **100 Suppl 1**, 11881-8.
11. Carpentier, R., Suner, R.E., Van Hul, N. *et al.* (2011) Embryonic ductal plate cells give rise to cholangiocytes, periportal hepatocytes, and adult liver progenitor cells. *Gastroenterology,* **141**, 1432-8.
12. He, L., Li, Y., Li, Y. *et al.* (2017) Enhancing the precision of genetic lineage tracing using dual recombinases. *Nat Med,* **23**, 1488-98.
13. Wang, Y., Huang, X., He, L. *et al.* (2017) Genetic tracing of hepatocytes in liver homeostasis, injury, and regeneration. *J Biol Chem,* **292**, 8594-604.

14. Feil, R., Wagner, J., Metzger, D. *et al.* (1997) Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochem Biophys Res Commun,* **237**, 752-7.

15. Kyrkanides, S., Miller, J., Bowers, W.J. *et al.* (2003) Transcriptional and posttranslational regulation of Cre recombinase by RU486 as the basis for an enhanced inducible expression system. *Mol Ther,* **8**, 790-5.

16. Schonig, K., Schwenk, F., Rajewsky, K. *et al.* (2002) Stringent doxycycline dependent control of CRE recombinase in vivo. *Nucleic Acids Res,* **30**, e134.

17. Malato, Y., Naqvi, S., Schurmann, N. *et al.* (2011) Fate tracing of mature hepatocytes in mouse liver homeostasis and regeneration. *J Clin Invest,* **121**, 4850-60.

18. Liu, F., Song, Y. and Liu, D. (1999) Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA. *Gene Ther,* **6**, 1258-66.

19. Zhang, G., Budker, V. and Wolff, J.A. (1999) High levels of foreign gene expression in hepatocytes after tail vein injections of naked plasmid DNA. *Hum Gene Ther,* **10**, 1735-7.

20. Kamimoto, K., Kaneko, K., Kok, C.Y. *et al.* (2016) Heterogeneity and stochastic growth regulation of biliary epithelial cells dictate dynamic epithelial tissue remodeling. *Elife*, **5**, e15034.

21. Rodrigo-Torres, D., Affo, S., Coll, M. *et al.* (2014) The biliary epithelium gives rise to liver progenitor cells. *Hepatology*, **60**, 1367-77.

22. Tarlow, B.D., Finegold, M.J. and Grompe, M. (2014) Clonal tracing of Sox9(+) liver progenitors in mouse oval cell injury. *Hepatology,* **60**, 278-89.

23. Furuyama, K., Kawaguchi, Y., Akiyama, H. *et al.* (2011) Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. *Nat Genet,* **43**, 34-41.

24. Kawaguchi, Y. (2013) Sox9 and programming of liver and pancreatic progenitors. *J Clin Invest,* **123**, 1881-6.

25. Wang, B., Zhao, L., Fish, M. *et al.* (2015) Self-renewing diploid Axin2(+) cells fuel homeostatic renewal of the liver. *Nature,* **524**, 180-5.

26. Lin, S., Nascimento, E.M., Gajera, C.R. *et al.* (2018) Distributed hepatocytes expressing telomerase repopulate the liver in homeostasis and injury. *Nature,* **556**, 244-8.

27. Lee, G. and Saito, I. (1998) Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. *Gene,* **216**, 55-65.

28. Missirlis, P.I., Smailus, D.E. and Holt, R.A. (2006) A high-throughput screen identifying sequence and promiscuity characteristics of the loxP spacer region in Cre-mediated recombination. *BMC Genomics,* **7**, 73.

29. Schaub, J.R., Huppert, K.A., Kurial, S.N.T. *et al.* (2018) De novo formation of the biliary system by TGFbeta-mediated hepatocyte transdifferentiation. *Nature,* May 2. [Epub ahead of print].

30. Shaner, N.C., Steinbach, P.A. and Tsien, R.Y. (2005) A guide to choosing fluorescent proteins. *Nat Methods,* **2**, 905-9.

31. Cranfill, P.J., Sell, B.R., Baird, M.A. *et al.* (2016) Quantitative assessment of fluorescent proteins. *Nat Methods,* **13**, 557-62.

32. Iverson, S.V., Comstock, K.M., Kundert, J.A. *et al.* (2011) Contributions of new hepatocyte lineages to liver growth, maintenance, and regeneration in mice. *Hepatology,* **54**, 655-63.

33. Livet, J., Weissman, T.A., Kang, H. *et al.* (2007) Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature,* **450**, 56-62.

34. Snippert, H.J., van der Flier, L.G., Sato, T. *et al.* (2010) Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell,* **143**, 134-44.

35. Griffiths, J.A., Scialdone, A. and Marioni, J.C. (2018) Using single-cell genomics to understand developmental processes and cell fate decisions. *Mol Syst Biol,* **14**, e8046.

36. Trapnell, C., Cacchiarelli, D., Grimsby, J. *et al.* (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol,* **32**, 381-6.

37. Setty, M., Tadmor, M.D., Reich-Zeliger, S. *et al.* (2016) Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat Biotechnol,* **34**, 637-45.

38. Haghverdi, L., Buttner, M., Wolf, F.A. *et al.* (2016) Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods,* **13**, 845-8.

39. Moignard, V., Woodhouse, S., Haghverdi, L. *et al.* (2015) Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nat Biotechnol,* **33**, 269-76.

40. Coffinier, C., Gresh, L., Fiette, L. *et al.* (2002) Bile system morphogenesis defects and liver dysfunction upon targeted deletion of HNF1beta. *Development,* **129**, 1829-38.

41. Lemaigre, F.P. (2009) Mechanisms of liver development: concepts for understanding liver disorders and design of novel therapies. *Gastroenterology,* **137**, 62-79.

42. Yang, L., Wang, W.H., Qiu, W.L. *et al.* (2017) A single-cell transcriptomic analysis reveals precise pathways and regulatory mechanisms underlying hepatoblast differentiation. *Hepatology,* **66**, 1387-401.

43. Su, X., Shi, Y., Zou, X. *et al.* (2017) Single-cell RNA-Seq analysis reveals dynamic trajectories during mouse liver development. *BMC Genomics,* **18**, 946.

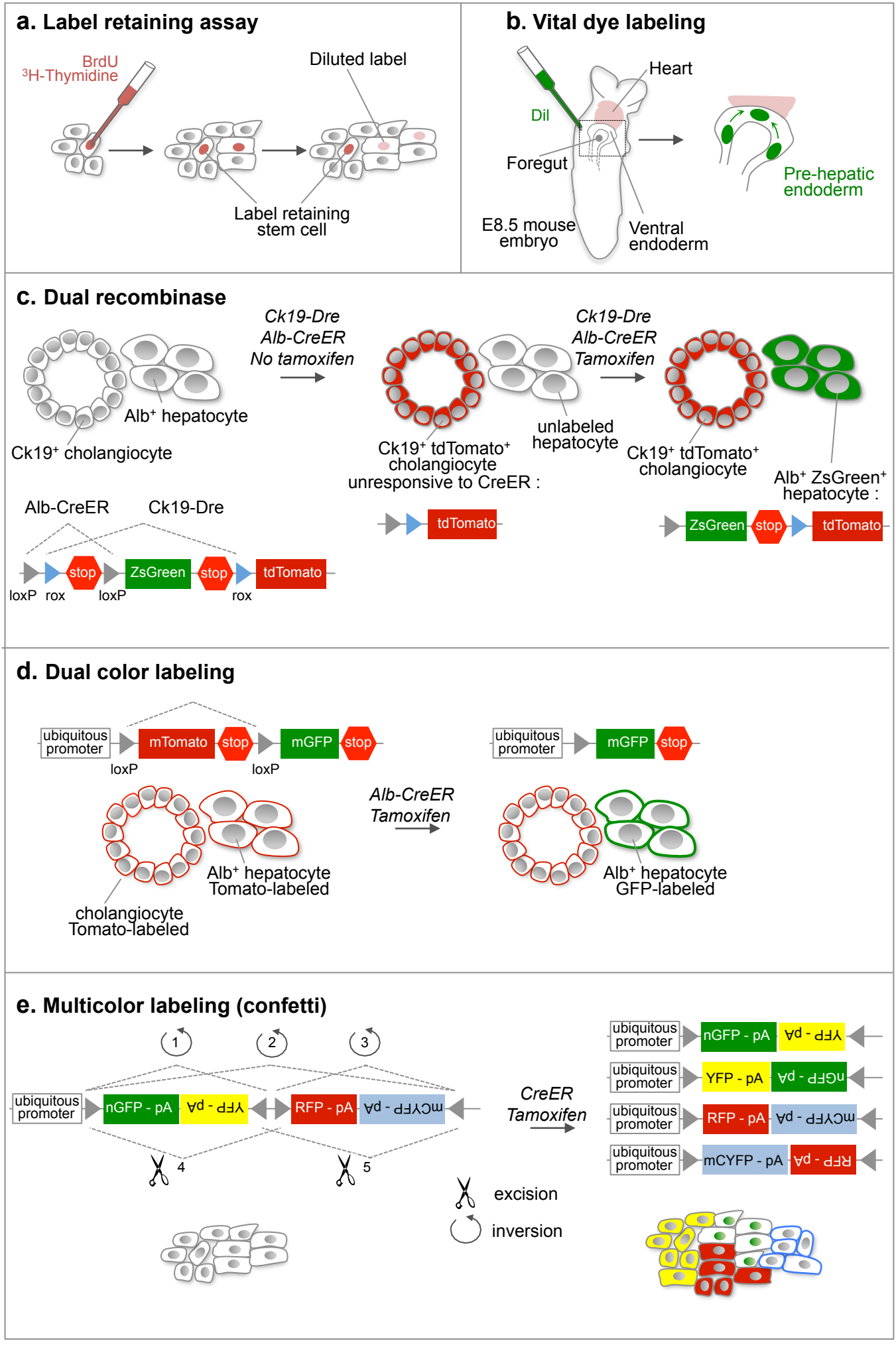
44. Spanjaard, B., Hu, B., Mitic, N. *et al.* (2018) Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat Biotechnol,* **36**, 469-73.

45. Yuan, G.C., Cai, L., Elowitz, M. *et al.* (2017) Challenges and emerging directions in single-cell analysis. *Genome Biol,* **18**, 84.

46. Woodworth, M.B., Girskis, K.M. and Walsh, C.A. (2017) Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nat Rev Genet,* **18**, 230-44.

## Figure legend

**Figure 1.** Lineage tracing methods implemented in the study of liver biology.


a. Label retaining assay
BrdU
3H-Thymidine
Diluted label
Label retaining
stem cell
b. Vital dye labeling
Heart
DiI
Foregut
E8.5 mouse
embryo
Ventral
endoderm
Pre-hepatic
endoderm
c. Dual recombinase
Ck19-Dre
Alb-CreER
No tamoxifen
Alb+ hepatocyte
Ck19+ cholangiocyte
Ck19-Dre
Alb-CreER
Tamoxifen
Ck19+ tdTomato+
cholangiocyte
unresponsive to CreER :
unlabeled
hepatocyte
Ck19+ tdTomato+
cholangiocyte
Alb+ ZsGreen+
hepatocyte :
Alb-CreER
Ck19-Dre
loxP
rox
stop
ZsGreen
tdTomato
ZsGreen
stop
tdTomato
d. Dual color labeling
ubiquitous
promoter
mTomato
stop
mGFP
stop
loxP
Alb-CreER
Tamoxifen
Alb+ hepatocyte
Tomato-labeled
cholangiocyte
Tomato-labeled
Alb+ hepatocyte
GFP-labeled
e. Multicolor labeling (confetti)
1
2
3
nGFP - pA
YFP - pA
RFP - pA
mCYFP - pA
4
5
CreER
Tamoxifen
excision
inversion