# cmgh

ORIGINAL RESEARCH

The Dynamic Chromatin Architecture of the Regenerating Liver

Amber W. Wang,[1] Yue J. Wang,[1,2] Adam M. Zahm,[1] Ashleigh R. Morgan,[1] Kirk J. Wangensteen,[1,3] and Klaus H. Kaestner[1]

[1]Department of Genetics, [3]Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; [2]Department of Biomedical Sciences, Florida State University College of Medicine, Tallahassee, Florida
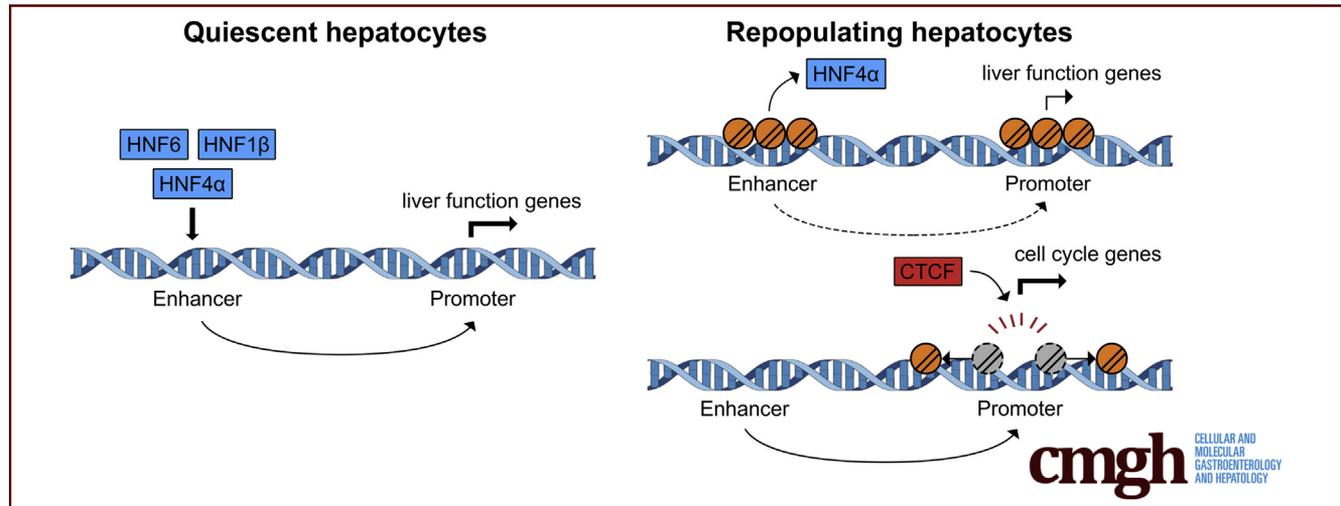
SUMMARY

We used the *Fah* null mouse repopulation model to determine how the chromatin structure is altered in hepatocytes during liver regeneration. We found that chromatin accessibility and transcription factor occupancy were highly dynamic in repopulating hepatocytes to enable the entry of hepatocytes into the cell cycle and the temporary cessation of certain biosynthetic functions.

BACKGROUND & AIMS: The adult liver is the main detoxification organ and routinely is exposed to environmental insults but retains the ability to restore its mass and function upon tissue damage. However, extensive injury can lead to liver failure, and chronic injury causes fibrosis, cirrhosis, and hepatocellular carcinoma. Currently, the transcriptional regulation of organ repair in the adult liver is incompletely understood.

METHODS: We isolated nuclei from quiescent as well as repopulating hepatocytes in a mouse model of hereditary tyrosinemia, which recapitulates the injury and repopulation seen in toxic liver injury in human beings. We then performed the assay for transposase accessible chromatin with high-throughput sequencing specifically in repopulating hepatocytes to identify differentially accessible chromatin regions and nucleosome positioning. In addition, we used motif analysis to predict differential transcription factor occupancy and validated the in silico results with chromatin immunoprecipitation followed by sequencing for hepatocyte nuclear factor 4α (HNF4α) and CCCTC-binding factor (CTCF).

RESULTS: Chromatin accessibility in repopulating hepatocytes was increased in the regulatory regions of genes promoting proliferation and decreased in the regulatory regions of genes involved in metabolism. The epigenetic changes at promoters and liver enhancers correspond with the regulation of gene expression, with enhancers of many liver function genes showing a less accessible state during the regenerative process. Moreover, increased CTCF occupancy at promoters and decreased HNF4α binding at enhancers implicate these factors as key drivers of the transcriptomic changes in replicating hepatocytes that enable liver repopulation.

CONCLUSIONS: Our analysis of hepatocyte-specific epigenomic changes during liver repopulation identified CTCF and HNF4α as key regulators of hepatocyte proliferation and regulation of metabolic programs. Thus, liver repopulation in the setting of toxic injury makes use of both general transcription factors (CTCF) for promoter activation, and reduced binding by a hepatocyte-enriched factor (HNF4α) to temporarily limit enhancer activity. All sequencing data in this study were deposited to the Gene Expression Omnibus database and can be downloaded with accession number GSE109466. *(Cell Mol Gastroenterol Hepatol 2020;9:121–143; https://doi.org/10.1016/j.jcmgh.2019.09.006)*

Keywords: Liver Regeneration; Hepatocyte; Chromatin Accessibility; ATAC-Seq; TRAP-Seq; RNA-Seq; ChIP-Seq; CTCF; HNF4α.

As the central metabolic organ in vertebrates, the liver regulates carbohydrate, protein, and lipid homeostasis; metabolizes nutrients, wastes, and xenobiotics; and synthesizes bile, amino acids, coagulation factors, and serum proteins.[1] To prevent acute liver failure from

exposure to harmful toxins, the liver has maintained an extraordinary ability to effectively restore its mass and function, in which the normally quiescent mature hepatocytes rapidly re-enter the cell cycle and divide.[2] Nonetheless, failure of regeneration can occur after exposure to harmful metabolites and environmental toxins, as often seen with the overconsumption of acetaminophen and alcohol.[3] Hence, understanding the genetic networks regulating the regenerative process can have an immense impact on the development of novel therapeutic strategies to treat acute liver failure.

The *Fah* null mouse model of human hereditary tyrosinemia type I provides a unique system to study the hepatocyte replication process after acute liver injury. Lack of the fumarylacetoacetate hydrolase (FAH) enzyme, which is essential for normal tyrosine catabolism, results in the accumulation of toxic intermediates followed by hepatocyte cell death.[4,5] *Fah^-/-* mice can be maintained in a healthy state by supplementation with the drug 2-(2-nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione (NTBC), which inhibits an upstream enzymatic step that prevents toxin production.[4] Alternatively, gene therapy that uses hydrodynamic tail-vein injection and the Sleeping Beauty transposon system to restore *Fah* expression in hepatocytes can rescue these mice.[6,7] When a small fraction (0.1%–1%) of hepatocytes express FAH after removal of NTBC, these hepatocytes competitively repopulate the liver in the context of injury through clonal expansion. Furthermore, this method allows lineage tracing of repopulating hepatocytes because only those with stable FAH expression can expand and repopulate the injured parenchyma.[7,8]

Eukaryotic DNA is highly organized and structured into compact chromatin to allow tight transcriptional control. Transcriptional regulation can be broadly categorized into 2 integrated layers: transcription factors and the transcriptional machinery, and chromatin structure and its regulatory proteins.[9] Expression of genes targeted by transcription factors depends on their binding to specific target DNA recognition sequences, combinatorial assembly with other cofactors, the concentration of the transcription factor, and post-translational modifications that affect protein localization.[10] The chromatin landscape is governed by DNA methylation, nucleosome position, histone modifications, and intrachromosomal and interchromosomal interactions.[10] Establishing the relationship of chromatin structure, transcriptional regulators, and the effects on gene expression is therefore vital to elucidating the transcriptional control governing the regenerative process. To date, most studies have relied on transcriptomic studies to document gene expression changes in the regenerating liver,[11–15] while 2 other studies focused on histone modifications.[16,17] However, these processes are downstream of chromatin reorganization and therefore do not capture the dynamic cross-talk of chromatin accessibility and transcriptional regulation. To identify transcriptomic changes specific to repopulating hepatocytes, we previously used the translating ribosome affinity purification (TRAP)[18] to isolate translating messenger RNAs only from repopulating hepatocytes.[15] To discern the dynamic chromatin patterns that underlie liver repopulation, we now implement the isolation of nuclei tagged in specific cell types (INTACT)[19] approach to isolate nuclei only from repopulating hepatocytes. This is achieved by expressing the green fluorescence protein (GFP)-tagged nuclear envelope protein Sad1 and UNC84 domain containing 1 (SUN1)-GFP together with FAH in *Fah^-/-* mice, followed by the sorting of GFP-positive nuclei from repopulating hepatocytes and assay for transposase accessible chromatin with high-throughput sequencing (ATAC-seq).[20] We identify promoter accessibility changes corresponding to up-regulation of cell-cycle genes and down-regulation of metabolic pathways, consistent with previous gene expression studies.[12,15] Integrative expression level and chromatin accessibility analysis has suggested that gene activation is associated primarily with increased promoter accessibility, while inactivation is correlated with the closure of selected promoters and enhancers. We propose a model in which a more accessible promoter allows increased transcription factor binding and gene activation, whereas decreased enhancer accessibility prevents binding of hepatocyte-enriched DNA binding proteins followed by inhibition of liver function genes so that the repopulating liver assumes a less differentiated state to promote cell growth and proliferation.

Results

Adaptation of INTACT to the Fah^-/- Model Allows for Isolation of Repopulating Hepatocyte Nuclei

Liver cells in human beings and mice rarely undergo division in homeostatic conditions.[2] However, with injury and repopulation, hepatocytes become facultative stem cells and divide to replenish liver mass and restore liver function.[2] We hypothesized that this change from quiescence to replication is accompanied by substantial and specific changes to chromatin accessibility. To analyze the chromatin specific to repopulating hepatocytes, we adapted the INTACT[19] method to the *Fah^-/-* model to label hepatocytes with the GFP-tagged nuclear envelope protein, SUN1-GFP, and performed fluorescence-activated cell sorting to isolate nuclei from whole liver at selected time points (Figure 1).
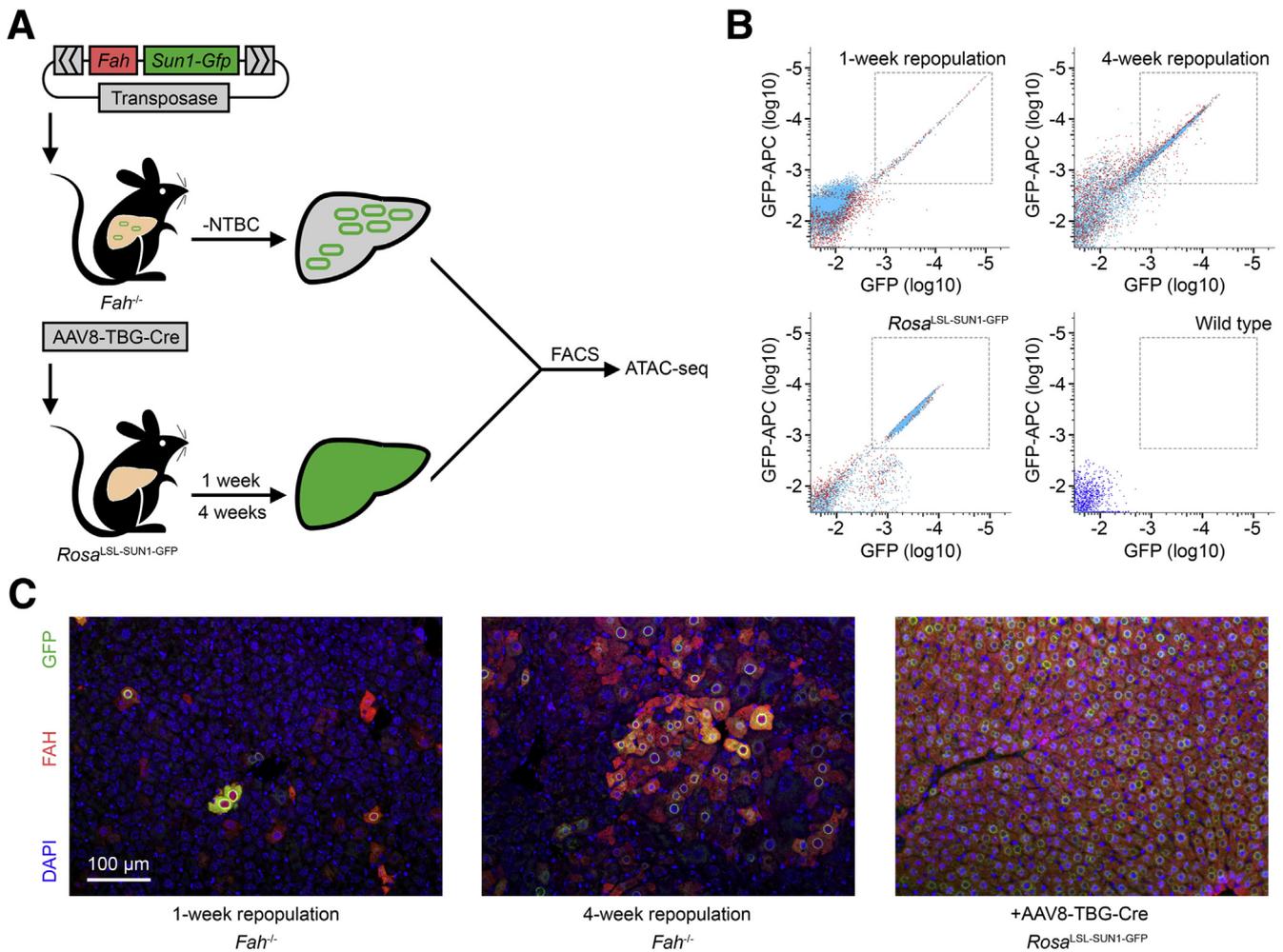
Figure 1. Implementation of the INTACT[19] method with the _Fah_[−/−] mouse model allows isolation of repopulating hepatocyte nuclei. (_A_) Schematic of coexpression of SUN1-GFP with FAH to label repopulating hepatocytes for FACS, followed by the ATAC-seq. (_B_) Representative images (n = 2) of repopulating hepatocyte nuclei show specific isolation with anti-GFP antibody labeling. _Gray boxes_ denote the sorting strategy to collect GFP[+] nuclei. (_C_) Representative images (n = 2) of immunofluorescent staining of GFP and FAH show coexpression of SUN1-GFP and FAH in repopulating hepatocytes of the _Fah_[−/−] mouse after 1 week (_left_) and 4 weeks (_middle_), and global expression of SUN1-GFP and FAH in all hepatocytes of the _Rosa_[LSL-SUN1-GFP] mouse 1 week after AAV8-TBG-Cre injection. FACS, fluorescence-activated cell sorting.

The SUN1-GFP fragment was subcloned into a FAH expression plasmid[7] so that all repopulating hepatocytes express GFP on the nuclear envelope. After hydrodynamic injection of the FAH-SUN1-GFP plasmid into _Fah_[−/−] mice, NTBC was removed and liver repopulation was allowed to proceed for 1 or 4 weeks (Figure 1A). As a control for healthy quiescent hepatocytes, _Rosa_[LSL-SUN1-GFP] transgenic mice[19] were injected with AAV8-TBG-Cre[21] to label all hepatocytes. Nuclei were isolated from repopulating hepatocytes exclusively at the selected time points by fluorescence-activated cell sorting with an anti-GFP antibody (Figure 1B). ATAC-seq[20] then was performed on the sorted nuclei to profile the changes in the chromatin regulatory landscape that occur during liver repopulation.

Sample ID	Condition	Index	Cumulate reads
Table 1. ATAC-seq Library Sequencing Summary			
SUN1-GFP-1	Quiescent	CGAGGCTG	119,120,180
SUN1-GFP-2	Quiescent	AAGAGGCA	111,970,248
3603	1-week repopulation	AATTCGTT	97,320,484
3604	1-week repopulation	GGCGTCGA	135,005,202
2383	4-week repopulation	GTAGAGGA	186,365,116
2385	4-week repopulation	TGCTGGGT	236,418,952
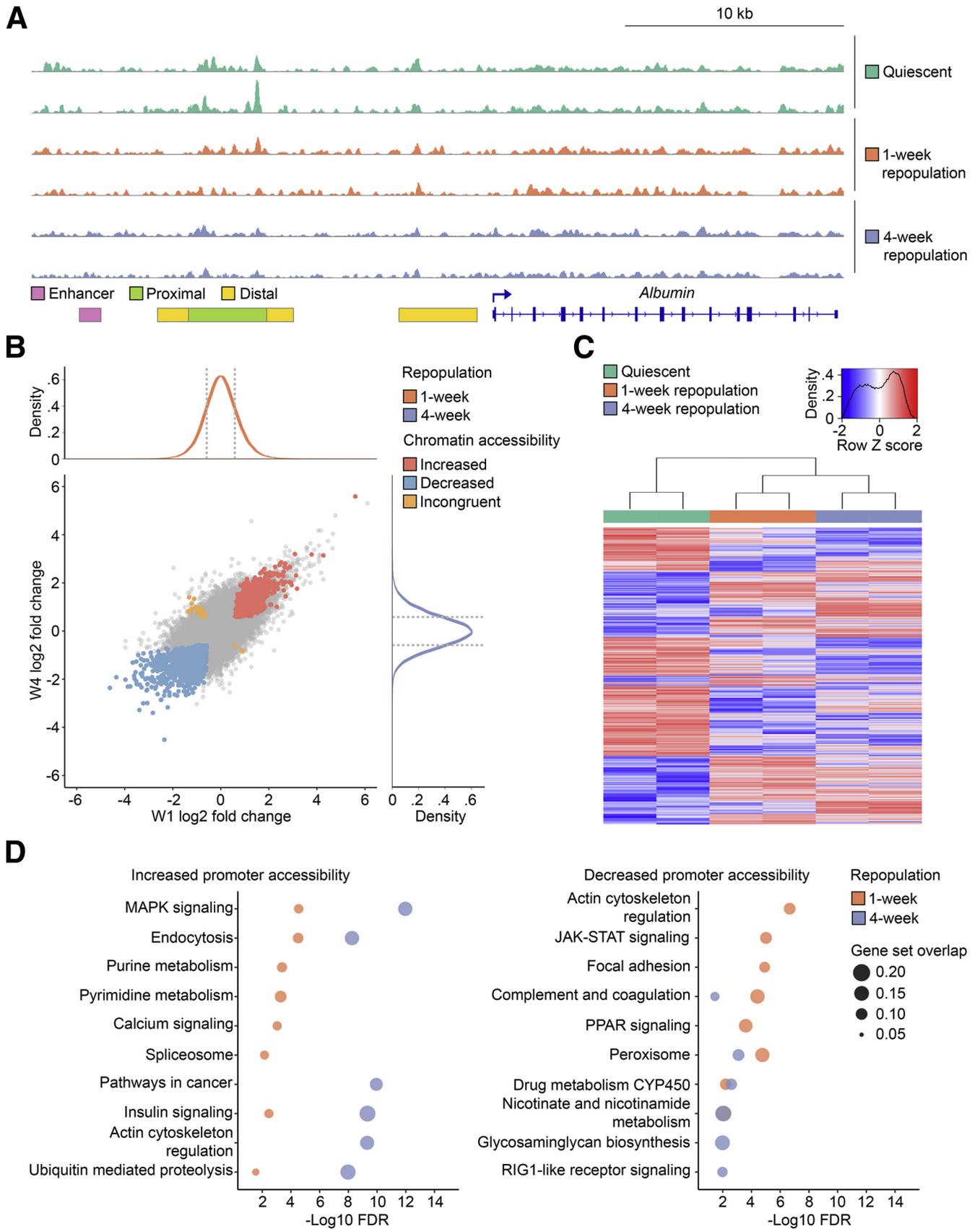
Immunofluorescence labeling showed the presence of GFP-tagged nuclear envelopes in FAH-positive cells (Figure 1C), illustrating the specificity of using SUN1-GFP+ nuclei as a marker to identify repopulating hepatocytes. Interestingly, FAH and GFP signals were not homogeneous across all replicating cells, possibly because of the different copy numbers of plasmids taken in after hydrodynamic tail-vein injection of the SUN1-GFP construct.[22] In addition, because the Sleeping Beauty transposon system shows little insertion site preference,[23] the loci in which the DNA fragments are integrated can affect expression levels of FAH and SUN1-GFP.[24]

ATAC-Seq Detects Differentially Accessible Chromatin Regions

All ATAC-seq libraries were sequenced to approximately 100 million reads to ensure ample coverage across the genome followed by quality assessment to verify the robustness of the data (Table 1). We observed consistent ATAC-seq signals across various loci such as the *Alb* gene, which showed a progressive decrease in accessibility at the enhancer region during repopulation (Figure 2A). To identify differentially accessible chromatin regions, fragments smaller than 150 bp, termed *nucleosome-free reads*, were used for peak calling. We identified 16,043 differentially accessible regions between quiescent and repopulating hepatocytes (Figure 2B, Supplementary Table 1), of which 5359 showed increased accessibility in 1-week and 5102 in 4-week repopulating hepatocytes, whereas 3580 regions showed decreased accessibility in week 1 and 5304 in week 4. Hierarchical clustering of the differentially accessible sites showed a clear separation of repopulating and quiescent hepatocytes (Figure 2C), corroborating previous transcriptome studies that 1-week and 4-week repopulating hepatocytes have a similar expression profile distinct from quiescent hepatocytes.[15] Replicates also clustered within the same condition, illustrating the reproducibility between biological replicates. Comparing accessibility regulated in the same direction in both time points (congruent), 1241 peaks were increased congruently and 2033 were decreased congruently (Figure 2B). Of note, only 28 regions showed accessibility changes in opposite directions in week 1 and week 4 (incongruent), reflecting the similarity in the chromatin profile between the 2 repopulation time points.

Next, we focused on differentially accessible promoter elements. Differential ATAC-seq regions within 1 kb upstream and downstream of transcription start sites (TSS) were determined and Kyoto Encyclopedia of Genes and Genomes pathway[25] analysis was performed (Figure 2D). As expected, pathways involved in cell growth and proliferation were enriched among the genes with increased accessibility in the promoter regions during repopulation, including mitogen-activated protein kinase signaling[26] and cancer pathways. Interestingly, purine and pyrimidine metabolism were enriched only in genes with increased promoter accessibility at week 1, but not at week 4, suggesting early activation of DNA synthesis immediately after liver injury in early stages of repopulation. This observation is consistent with previous comparison of the *Fah*−/− and partial hepatectomy (PHx) models showing that the transcriptome of 1-week repopulating hepatocytes in the *Fah*−/− mouse is closest to that of 36 and 48 hours after PHx,[15] at which the highest rate of DNA synthesis occurs in this model.[27] On the other hand, genes involved in hepatocyte functions such as complement and coagulation and metabolic pathways showed significantly decreased promoter accessibility at both regeneration time points. Our pathway enrichment analysis substantiates prior studies of gene expression profiles and extends the findings to chromatin accessibility in that proliferation pathways are activated while liver functions are inhibited during repopulation.[12,15]

Integration of Chromatin Accessibility and Gene Activity Infers Regulatory Mechanisms

To evaluate the association of chromatin landscape and gene expression, we used our prior TRAP sequencing (TRAP-seq) study[15] as a data set of transcriptomic changes in repopulating hepatocytes. Genes with ATAC-seq signals and TRAP-seq reads that changed in the same direction at the same time point were identified as concordant genes (Figure 3A, Supplementary Table 2). We observed significant overlap of the concordant genes with ATAC-seq and TRAP-seq ($P < 1E-16$ for all 1-week concordant genes and 4-week concordantly activated genes; $P = .03$ for 4-week concordantly inhibited genes), although there was no significant overlap of genes with increased expression at 1 week and decreased chromatin accessibility at 4 weeks ($P = .39$). Kyoto Encyclopedia of Genes and Genomes pathway[25] analysis suggested enrichment of cell growth and replication in the week 1 concordantly activated genes, and overrepresentation of biosynthesis and metabolism in both week 1 and week 4 concordantly inhibited genes (Figure 3B and C). In addition, pathway enrichment supported previous observations that activation of the glutathione metabolic network is essential for reactive oxygen species removal after PHx or recovery after toxic liver injury.[15,28,29] We conclude that changes to the chromatin structure underlie the up-regulation of genes involved in cell proliferation and

---

Figure 2. (*See previous page*). **Chromatin accessibility changes during liver repopulation are related to cell growth activation and metabolic inhibition.** (*A*) ATAC-seq shows reproducible signals across biological replicates and a decrease of peak intensity in the proximal regulatory region[103] of the *Alb* locus. (*B*) There were 16,043 significantly differential accessible regions identified in repopulating and quiescent hepatocytes (absolute fold change, ≥1.5; FDR, ≤0.05). Comparison of differential accessible regions identified at different time points during repopulation shows 3273 that changed in the same direction (congruent peaks), of which 1241 were increased congruently (*red dots*) and 2033 were decreased congruently (*blue dots*). (*C*) Hierarchical clustering of all differentially accessible regions shows that biological replicates have a similar chromatin landscape. (*D*) Kyoto Encyclopedia of Genes and Genomes pathway analysis of differential accessible promoters with increased (*left*) and decreased (*right*) accessibility in repopulating hepatocytes. CYP450, cytochrome P450; JAK-STAT, Janus kinase/signal transducers and activators of transcription; MAPK, mitogen-activated protein kinase; RIG1, retinoic acid inducible gene 1.
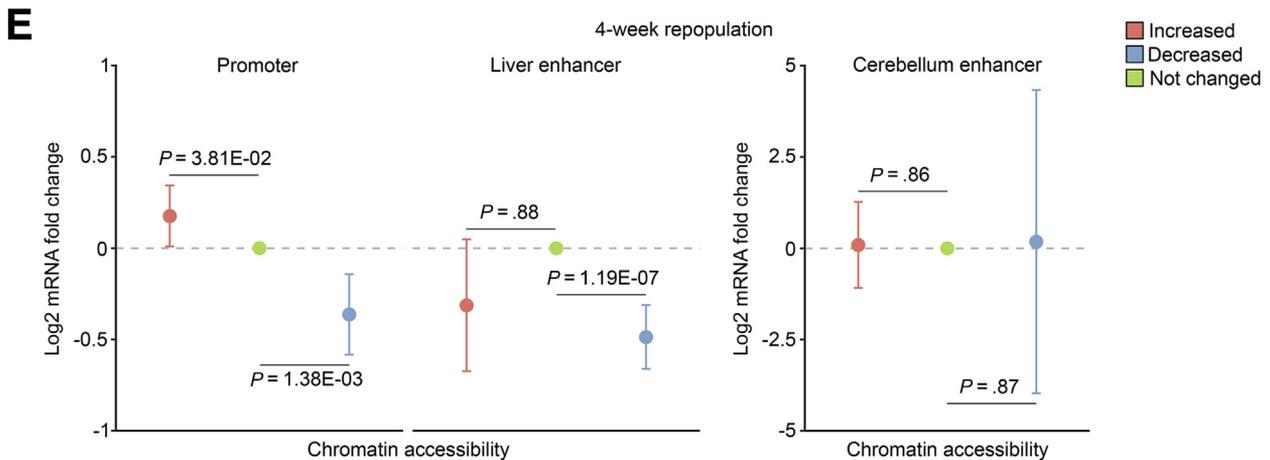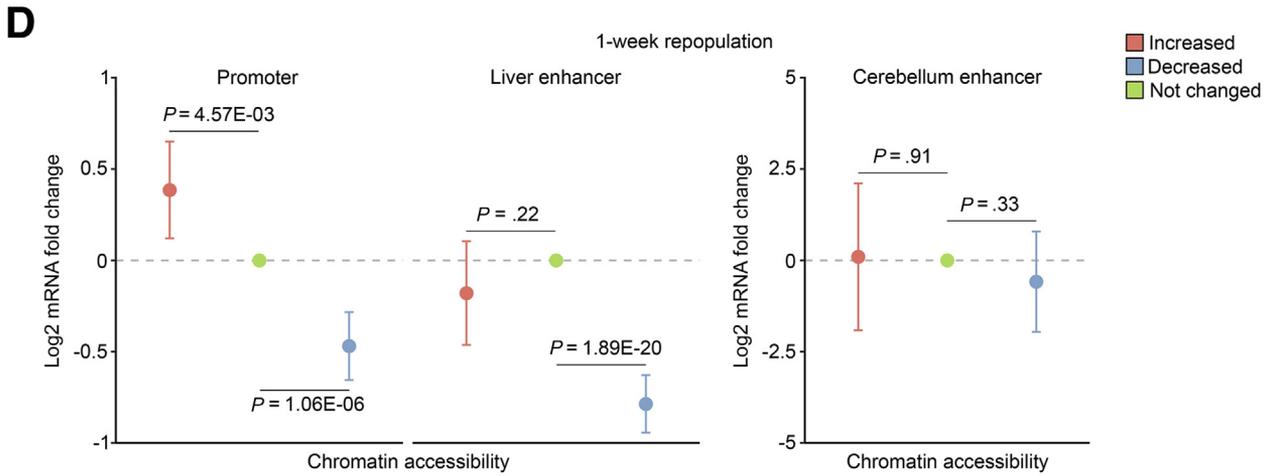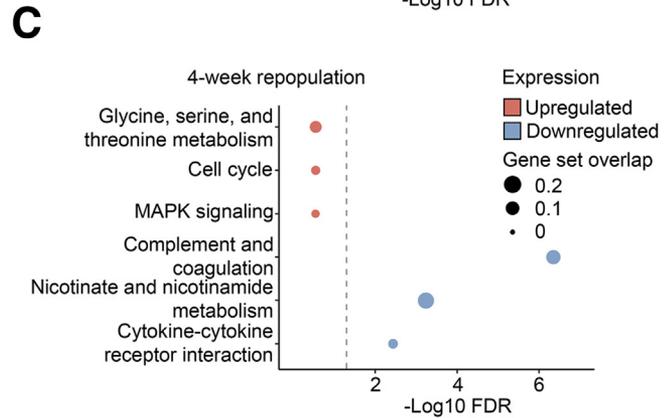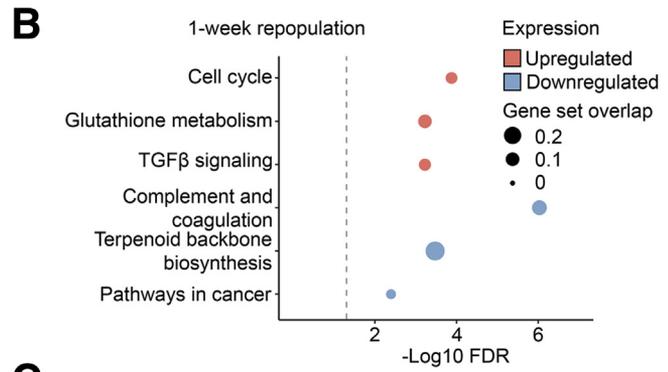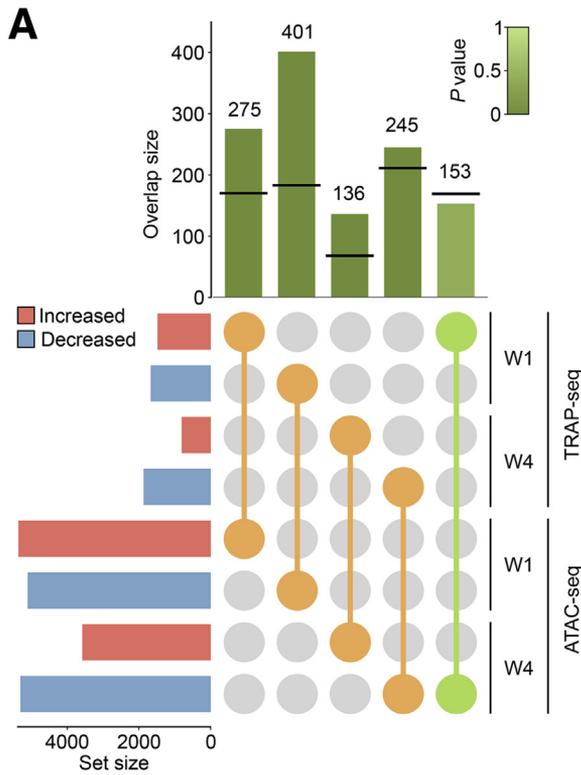
down-regulation of genes associated with metabolic processes.

Next, we sought to investigate coregulatory networks in repopulating hepatocytes. All ATAC-seq peaks identified were first separated into increased, decreased, or unchanged accessibility, with a cut-off value of absolute fold change of 1.5 or higher and a false-discovery rate (FDR) of 0.05 or less, followed by subdivision into regulatory regions of promoters, liver-specific enhancers, or cerebellum-specific enhancers as a negative control.[30] Promoter peaks were annotated to the nearest genes, and the corresponding transcript levels at the same time point were extracted from TRAP-seq data.[15] We then compared the gene expression levels in the differentially accessible promoters with those in the unchanged promoters (Figure 3D and E). The normalized log_2 fold change was positive ($P = 7.47E-03$ in week 1 and 3.81E-02 in week 4) with increased and negative ($P = 1.06E-06$ in week 1 and 1.38E-03 in week 4) with decreased promoter accessibility at both time points, showing a significant association of promoter openness and transcriptional activity. Differentially accessible liver enhancer peaks were categorized similarly, putative enhancer-regulated genes were extrapolated,[30] corresponding target gene expression was extracted,[15] and the transcript level changes were compared with those of genes with unchanged enhancer accessibility. Interestingly, decreased liver enhancer accessibility was highly correlated with decreased gene activity ($P = 1.89E-20$ in week 1 and 1.19E-07 in week 4), although no significant expression changes ($P = .22$ in week 1 and .88 in week 4) were associated with increased enhancer openness. Although the exact mechanism explaining this lack of correlation requires further evaluation, we posit that target genes regulated by enhancers in the quiescent liver already are highly expressed in mature differentiated hepatocytes.[12,15] An increase in liver enhancer accessibility hence does not further increase the expression of these genes significantly. Another likely explanation for the lack of significant association between increased liver enhancer accessibility and activation of target genes could be the recruitment of repressors instead of activators to the regulatory elements to decrease expression.[31–33] Finally, refinement of the computationally predicted enhancer-promoter pairs with experimental approaches could result in a more accurate correlation of enhancer accessibility and transcriptional activity. Importantly, cerebellum enhancers showed no significant correlation with the changes in transcript levels and chromatin accessibility in the repopulating liver, as expected (Figure 3D and E, right). Our integrated ATAC-seq and TRAP-seq analysis showed that gene activation is regulated by increased promoter accessibility, presumably allowing recruitment of transcriptional activators and RNA polymerase II to the TSS, whereas gene inhibition may be governed by both decreased promoter and enhancer openness, preventing long-range enhancer–promoter interactions.[34]

Differential Chromatin Accessibility Predicts Transcription Factors Involved in Liver Repopulation

Dynamic coordination of chromatin structure and transcription factors is required to fine-tune gene expression. Chromatin organization influences access of the transcriptional apparatus by regulating binding sequence accessibility[35] and transcription factor binding stability[36]; conversely, transcription factors affect access of remodelers to the chromatin[35] and histones.[37] To identify DNA binding transcription factors that connect differential chromatin accessibility and gene expression, we performed de novo motif profiling at differentially accessible promoters and liver enhancers.[30]

We found enrichment of the ETS transcription factor ELK1 motif in promoters with increased accessibility in both 1-week (FDR, 1E-76) and 4-week (FDR, 1E-41) repopulating hepatocytes (Figure 4A and B, Table 2). ELK1 binds to the serum response element upon mitogen-activated protein kinase phosphorylation[38] to activate immediate early genes such as *Fos* and components of the basal transcriptional machinery.[39] Furthermore, ELK1 supports cell-cycle entry during liver regeneration because *Elk1*[-/-] mice show reduced hepatocyte proliferation after PHx.[40] We postulate that promoters become more accessible after acute liver injury to permit increased ELK1 occupancy, enabling hepatocyte repopulation.

Among the regions with increased accessibility during liver repopulation, surprisingly, the CCCTC-binding factor

Figure 3. (*See previous page*). Association of expression levels and chromatin accessibility implicates divergent regulatory mechanisms for gene activation and inhibition. (A) Differential gene expression data were obtained from a previous study that implemented translating ribosome affinity purification followed by RNA-sequencing (TRAP-seq).[15] The upset plot shows overlap of ATAC-seq regions and TRAP-seq genes that are changed significantly in the same direction at the same time points (concordant genes) in repopulating hepatocytes. The Fisher exact test was performed to calculate the significance of overlapping targets. *Top*: The *horizontal black lines* in the *green bars* indicate the number of overlaps expected by chance. (B and C) Kyoto Encyclopedia of Genes and Genomes pathway analysis of concordantly activated and repressed genes in (B) 1-week and (C) 4-week repopulating hepatocytes. *Dashed lines* denote a FDR of 0.05. (D and E) Association of changes in chromatin accessibility and gene expression in (D) 1-week and (E) 4-week repopulating hepatocytes indicate that promoter accessibility changes are related to both gene activation and inhibition, while only decreased liver enhancer accessibility is correlated significantly with decreased expression of putative target genes.[30] Cerebellum enhancers and their putative targets do not show any significant relationship to chromatin accessibility and gene expression changes in the liver. One-sample *t* tests were performed to identify the differences in normalized log_2 fold change in differentially accessible and unchanged chromatin regions. *Vertical lines* denote the 95% CI of normalized log_2 fold change in peaks with increased and decreased accessibility. MAPK, mitogen-activated protein kinase; mRNA, messenger RNA; TGFβ, transforming growth factor β.
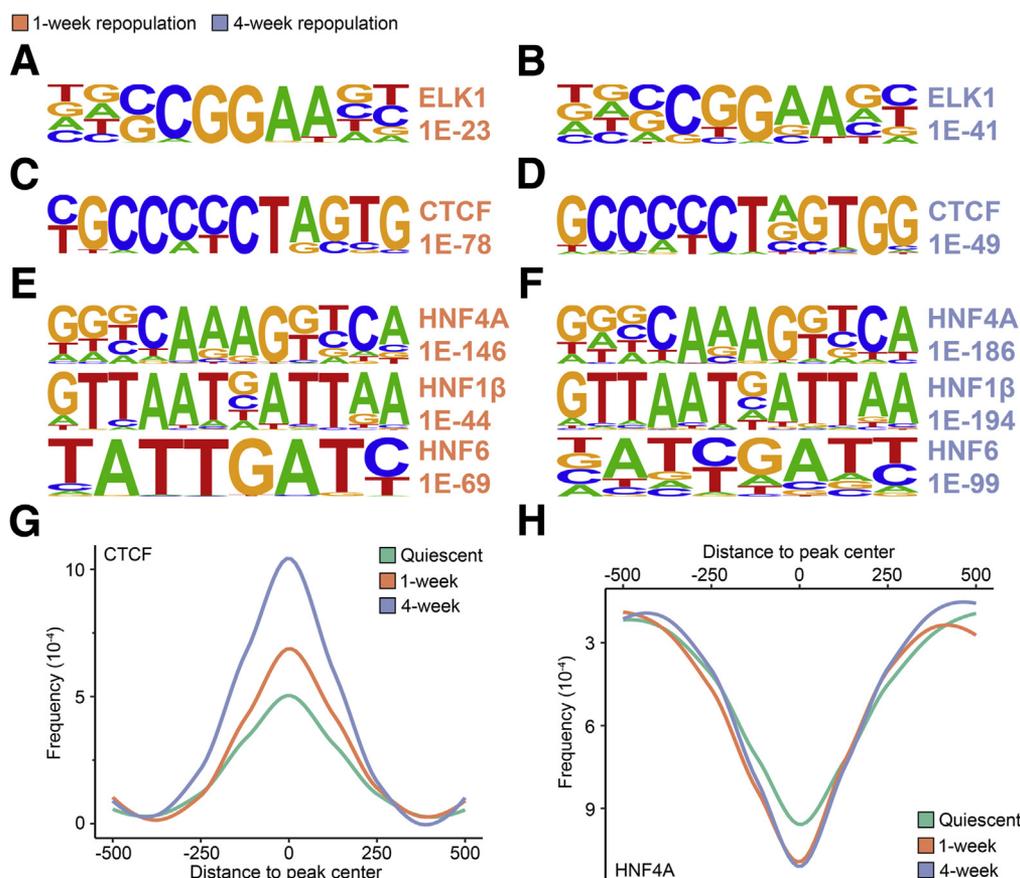
Figure 4. Enrichment analysis identifies transcription factor motifs overrepresented at differential accessible promoters and enhancers.[30] Up to 3 representative motifs common between 1- and 4-week repopulating hepatocytes are shown. For a complete list of enriched motifs identified in each condition and genomic region, refer to Table 2. (*A* and *B*) The ELK1 motif is enriched in promoter regions that became more open in repopulating hepatocytes. Other motifs not shown include KLF3, GFY, SP2, and ATF3. (*C* and *D*) The CTCF motif is overrepresented in liver enhancers with increased accessibility in both (*C*) 1-week and (*D*) 4-week repopulating hepatocytes. HNF1β is enriched specifically in 1-week repopulating hepatocytes. (*E* and *F*) Motifs of liver-enriched transcription factors HNF4α, HNF1β, and HNF6 are enriched in enhancers with decreased accessibility during (*E*) 1-week and (*F*) 4-week liver repopulation. (*G* and *H*) Motif frequency of the differential accessible peaks for (*G*) CTCF and (*H*) HNF4α show enrichment of the transcription factor motifs at the enhancer peak center in repopulating hepatocytes. (*A–F*) Numbers presented denote the FDR.

(CTCF) motif was highly enriched (FDR, 1E-78 in week 1 and 1E-49 in week 4) (Figure 4*C* and *D*). CTCF plays numerous roles in transcriptional regulation to function as a transcriptional activator[41] or repressor,[42] an insulator to block enhancer–promoter interactions,[43] a chromatin structure organizer to form topologically associated domains,[44] a modulator of long-range chromatin looping,[45] and even a mediator of local RNA polymerase II, pausing to regulate alternative exon use.[46] CTCF is recruited to the *Axin1* promoter as a transcriptional repressor by the long noncoding RNA associated with liver regeneration (LALR1) after PHx, leading to activation of Wnt/β-catenin signaling to promote hepatocyte proliferation.[47] However, the function of CTCF in liver regeneration is not fully understood.

In addition, we found the hepatocyte nuclear factor 4α (HNF4α) binding motif to be associated significantly with liver enhancers with decreased accessibility during liver regeneration (FDR, 1E-146 in week 1 and 1E-186 in week

4) (Figure 4*E* and *F*). HNF4α is a master regulator atop the transcriptional cascade of hepatocyte differentiation[48,49] and a crucial factor that maintains hepatocytes in the differentiated state.[50] Importantly, HNF4α suppresses liver proliferation, because mice with conditional deletion of *Hnf4a* show increased hepatocyte bromodeoxyuridine incorporation and Ki67 expression.[51] HNF4α also directly inhibits cell growth and replication pathways, as shown by the up-regulation of cell cycle and proliferation genes upon acute HNF4α loss.[51,52] Moreover, motifs of other liver-enriched transcription factors also were overrepresented at enhancers that became less accessible in repopulating hepatocytes, including HNF1β and HNF6[53] (Figure 4*E* and *F*).[53] We examined the locations for CTCF and HNF4α motifs within regions of dynamic chromatin accessibility and found that they are present in the center of these regions with CTCF at those with increased accessibility (*P* = 2.70E-04 in week 1 and 1.97E-13 in week 4), and HNF4α at those with decreased accessibility

Table 2. ATAC-seq Peaks Motif Enrichment Analysis

Motif	P value	Log P value	Targets, n (%)	Background, n (%)	Best match/details	Notes
GGGCGGGGCYWV	1.00E-39	-90.003359	498.0 (45.36)	11,864.1 (26.78)	KLF3(Zf)/MEF-Klf3-ChIP-Seq(GSE44748)/Homer(0.942)	
NDSCGGAANY	1.00E-25	-59.311416	464.0 (42.26)	12,116.1 (27.34)	Elk1(ETS)/Hela-Elk1-ChIP-Seq(GSE31477)/Homer(0.959)	Shown in Figure 4*A*
GGGAATTGTAGT	1.00E-21	-48.87299	89.0 (8.11)	1089.0 (2.46)	GFY(?)/Promoter/Homer(0.975)	
CCACTAGAGGGC	1.00E-18	-43.414469	121.0 (11.02)	1959.4 (4.42)	CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer(0.861)	
CAGGGAGGCGGT	1.00E-14	-34.185766	126.0 (11.48)	2359.2 (5.32)	PB0076.1_Sp4_1/Jaspar(0.615)	
CGCGTTGTCT	1.00E-14	-32.421255	395.0 (35.97)	11,281.3 (25.46)	Smad3(MAD)/NPC-Smad3-ChIP-Seq(GSE36673)/Homer(0.620)	
GGKKYGGAGT	1.00E-13	-30.614303	88.0 (8.01)	1452.7 (3.28)	KLF3(Zf)/MEF-Klf3-ChIP-Seq(GSE44748)/Homer(0.636)	

NOTE. The selection criteria were as follows: (1) *P* value < 1.00E-12; (2) best match/details: de novo motif is in the center of known motif; and (3) motifs commonly identified (up to 3) in both 1- and 4-week repopulation hepatocytes are shown in Figure 4.

(P = .59 in week 1 and 2.48E-03 in week 4) (Figure 4*G* and *H*).

In summary, de novo motif analysis of differentially accessible ATAC-seq regions suggests increased occupancy of ELK1 and CTCF at chromatin regions that become more accessible, and decreased binding of liver-enriched transcription factors at liver enhancers that become less accessible during repopulation.

HNF4α Occupancy Is Decreased in Liver-Specific Enhancers During Repopulation

We postulated that decreased HNF4α binding allows repopulating hepatocytes to assume a less differentiated and pro-proliferative state and performed chromatin immunoprecipitation sequencing (ChIP-seq) on quiescent and 4-week repopulating livers to examine genome-wide HNF4α occupancy during the repopulation process. We observed 508 peaks with decreased and only 14 peaks with increased occupancy in repopulating livers (Figure 5*A*, Supplementary Table 3). Remarkably, 42% (214) of lost HNF4α occupancy occurred within previously defined liver enhancers,[30] while 23% (119) fell into distal intergenic regions, and 10% (52) were within 1 kb upstream and downstream of the TSS (promoter) (Figure 5*B*). These data corroborate the differentially accessible chromatin analysis of transcription factor motifs that had identified enrichment of the HNF4α consensus sequence at enhancers with decreased accessibility in repopulating hepatocytes (Figure 4*H*).

Next, we integrated ATAC-seq, ChIP-seq, and TRAP-seq data sets,[15] and identified hepatocyte-enriched genes crucial for establishing liver functions including complement and coagulation (*Cfb*, *F2*), biosynthesis (*Itih1*, *Acsl1*, *Pgrmc1*), and metabolism (*Ugt1a5*, *Mthfs*, *Rdh10*)[54] as correlated with decreased HNF4α enhancer occupancy during regeneration (Figure 5*C* and *E*). To explore the

mechanism responsible for decreased HNF4α occupancy during liver repopulation, we next turned to the TRAP-seq data set[15] to inspect *Hnf4a* expression levels in quiescent and replicating hepatocytes. Remarkably, we found a 50% reduction of *Hnf4a* transcripts in 4-week repopulating hepatocytes (FDR, 4.16E-3) compared with the quiescent liver (Figure 5*D*, Supplementary Table 4). Taken together, these results implicate decreased chromatin accessibility and reduced *Hnf4a* expression as contributors to the suppression of hepatocyte-specific genes and down-regulation of liver biosynthetic functions during repopulation.

CTCF Promoter Occupancy Is Increased in the Repopulating Liver

To extend the computational finding of enriched CTCF motif at promoters with increased accessibility, we performed ChIP-seq in quiescent and 4-week repopulating livers. CTCF occupancy was increased at 1382 sites in the repopulating liver, while only 2 peaks showed decreased binding (Figure 6*A*, Supplementary Table 5). To characterize the role of increased CTCF occupancy during liver repopulation, we first evaluated its potential insulator function by calculating an "insulator strength score"[55] at all gained binding sites. Genomic regions with increased CTCF occupancy with divergent flanking promoters within 50 kb were identified and the normalized expression levels corresponding to the genes were extracted from our TRAP-seq data (Figure 6*B*).[15] Surprisingly, gene pairs with increased CTCF binding were not significantly more enriched for differential gene expression than random gene pairs (P = .9) (Figure 6*C*), suggesting that CTCF is unlikely to act as an insulator during liver repopulation.

Remarkably, the vast majority (1026; 74%) of the gained CTCF peaks fell within 1 kb upstream and downstream of the
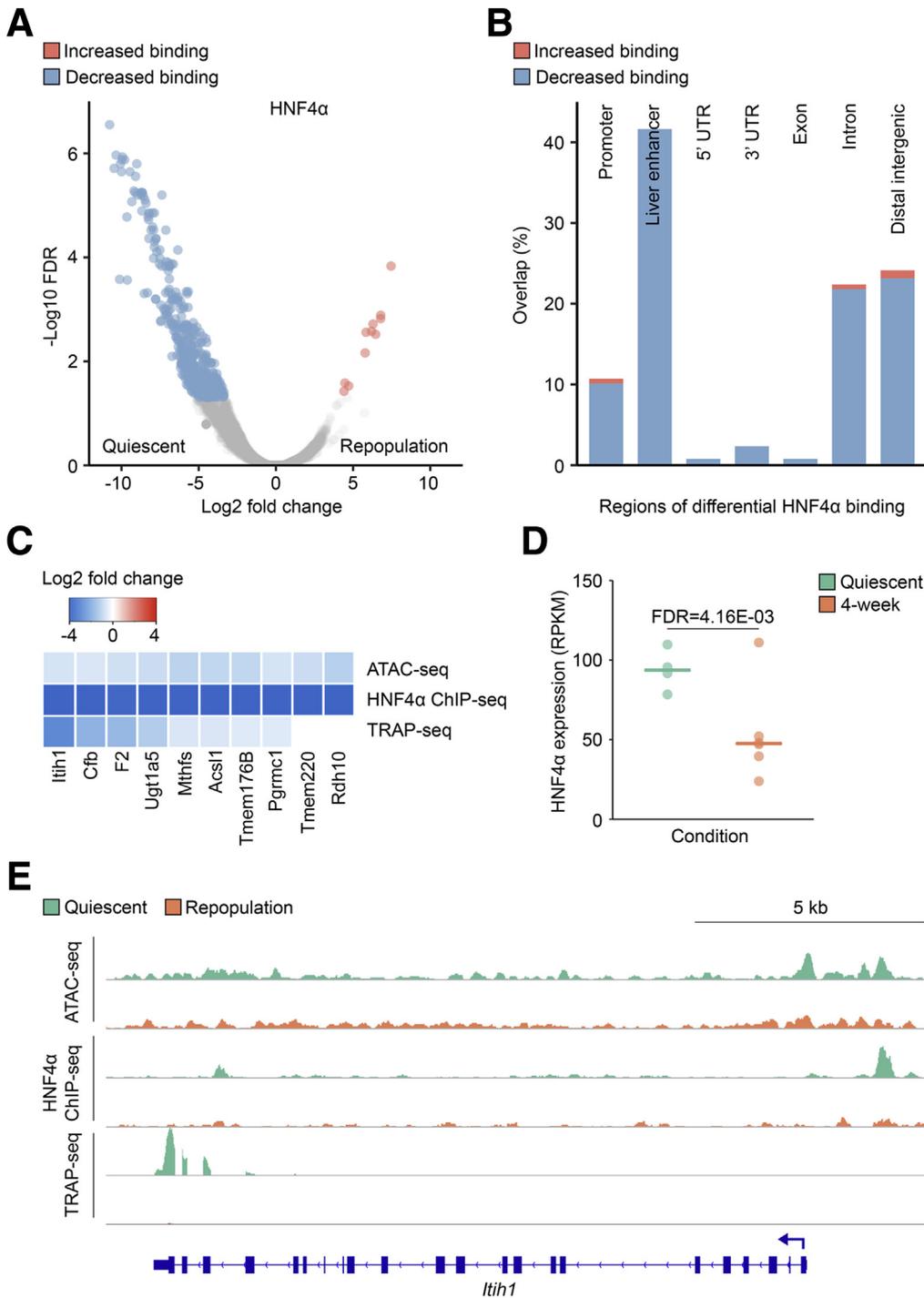
Figure 5. HNF4α binding is decreased in the repopulating liver. (*A*) The 508 genomic regions show decreased, and only 14 show increased, HNF4α occupancy in the regenerating liver. (*B*) Forty percent of peaks with decreased HNF4α binding overlap with liver-enriched enhancers (liver enhancer),[30] and 25% are in distal intergenic regions that contain ubiquitous enhancers (distal intergenic). (*C*) Integrative analysis of chromatin accessibility (ATAC-seq), HNF4α binding (ChIP-seq), and gene expression (TRAP-seq)[15] changes suggest that the suppression of liver functions including complement, biosynthesis, and metabolic pathways during liver regeneration are associated with reduced HNF4α occupancy. (*D*) HNF4α expression is downregulated in repopulating hepatocytes (n = 4 for quiescent hepatocytes, n = 6 for repopulating hepatocytes).[15] (*E*) Representative tracks (n = 2 for ATAC-seq and ChIP-seq, n = 4 for TRAP-seq) of chromatin accessibility, HNF4α occupancy, and transcript levels at *Itih1*, the locus with the strongest decrease of HNF4α occupancy. RPKM, reads per kilobase of transcript, per million mapped; UTR, untranslated region.

TSS (promoter) (Figure 6*D*). To examine the targets of increased CTCF occupancy, all differentially bound peaks were annotated to the nearest genes and their corresponding expression changes were obtained from our TRAP-seq data set.[15,25,56] We found 545 (39%) peaks associated with chromatin modification, transcription regulation, and cancer (Figure 6*E*), while 656 (47%) sites with increased CTCF binding were associated with inhibition of genes in cell death regulation, stress response, and morphogenesis. Together,

our network analysis suggests a diverse role for CTCF in transcriptional regulation in which increased CTCF occupancy supports hepatocyte replication and prevents cell death during liver repopulation, possibly by enabling binding of both activating and repressing cofactors.

CTCF is known to show divergent roles in activating and repressing transcription by recruiting various protein partners in a context-dependent manner.[57] To identify these cofactors, we performed motif analysis for the regions differentially bound by CTCF (Figure 6*F*). As

expected, the CTCF motif was highly enriched (FDR, 1E-26) at all differential binding sites, confirming the specificity of the anti-CTCF antibody for immunoprecipitation. At sites where CTCF binding corresponded to gene activation, we observed significant enrichment for the zinc finger and BTB domain-containing protein 3 (ZBTB3) (FDR, 1E-10) and nuclear transcription factor Y (NF-Y) (FDR, 1E-10) binding motifs (Figure 6F). ZBTB3 is considered a likely factor binding 5' of CTCF because of its frequent enrichment approximately 10 bp upstream of CTCF motifs in the human genome.[58] Furthermore, expression of ZBTB3 is induced by the accumulation of reactive oxygen species to promote cancer cell growth and prevent apoptosis via the activation of antioxidant gene expression in cell lines.[59] Whether CTCF directly interacts with or indirectly recruits ZBTB3 is as yet unclear, but the proteins are likely to interact based on their close proximity at promoters. NF-Y binds to the CCAAT box present at approximately 30% of the promoters[60] and is required for cell-cycle progression, DNA synthesis, and proliferation in mouse embryonic fibroblasts.[61] In addition, reconstituted in vitro transcription reactions showed that binding of NF-Y disrupts the nucleosome structure at promoters containing the NF-Y recognition sequence.[62] Recruitment of NF-Y hence could induce local nucleosome repositioning to allow increased accessibility of the transcriptional apparatus to activate gene expression.

On the other hand, the Yin Yang 1 (YY1) binding motif was enriched (FDR, 1E-13) at sites where increased CTCF occupancy corresponded with decreased gene expression (Figure 6F). YY1 regulates embryogenesis, cell differentiation, and tumorigenesis,[63,64] as well as enhancer–promoter interactions analogous to long-range chromatin looping mediated by CTCF.[65] YY1 functions as a transcriptional repressor via recruitment of the polycomb repressor complex, resulting in trimethylation of histone H3 lysine 27.[66,67] It is also a cofactor of CTCF in regulating X chromosome inactivation, although the exact mechanism remains unclear.[68] Given these observations, it is likely that direct or indirect co-binding of CTCF and YY1 at promoters induces transcriptional repression or disrupts enhancers to promoter interactions to down-regulate target genes.

When examining gene expression, we found the levels of ZBTB3 and YY1 were not changed significantly in repopulating hepatocytes (Supplementary Table 4). Three NF-Y proteins showed varying changes in transcript levels, with unchanged NF-YA, down-regulated NF-YB in 1-week, and down-regulated NF-YC in 4-week repopulating hepatocytes, albeit all with modest changes of less than 2-fold. These observations do not rule out the possibility of posttranslational modifications that might alter the abundance or localization of transcription factors.

To analyze if transcription factors colocalize to CTCF-occupied promoters with differential gene expression during liver regeneration, we performed ZBTB3 and YY1 ChIP–quantitative polymerase chain reaction (qPCR) on quiescent and 4-week repopulating livers. We observed a significant increase of ZBTB3 occupancy at the *Ctnna2* ($P =$ .023) and *Smad3* ($P = $ .025) promoters, 2 genes with increased promoter accessibility, increased CTCF binding, and up-regulated expression during liver regeneration (Figure 6G and H). Regarding YY1 occupancy, there was a significant increase at the *Bcl2l11* ($P = $ .029) promoter, a gene with increased promoter accessibility, enhanced CTCF occupancy, and decreased transcript levels (Figure 6I). With the limited loci tested, we conclude that ZBTB3 is recruited to open chromatin regions occupied by CTCF to activate gene expression during liver regeneration. On the other hand, increased YY1 binding to select promoters with increased CTCF binding could regulate transcriptional repression in repopulating hepatocytes. These results suggest that increased chromatin accessibility correlates with enhanced CTCF occupancy that recruits coactivators or corepressors to fine-tune target gene expression to induce replication and prevent apoptosis during liver repopulation (Figure 6J). Future experiments that use co-immunoprecipitation and high-throughput sequencing technologies to analyze interactions between CTCF and cofactors as well as genome-wide binding patterns of the coregulators will aid in the understanding of mechanisms underlying CTCF modulation.

Liver Regeneration Is Accompanied by Nucleosome Remodeling

Most eukaryotic DNA is packaged around histone protein octamers into nucleosomes to regulate chromatin organization and transcriptional control. Nucleosome properties such as positioning and turnover rates can affect the binding of transcription factors and access of the transcriptional machinery.[69] The nucleosome landscape adjacent to the TSS is of particular interest because nucleosomes adopt a specific phasing pattern immediately upstream and downstream.[70] Hence, nucleosome organization could act as an additional layer of transcriptional regulation in repopulating hepatocytes.

We inferred nucleosome positioning from nucleosome-containing sequences by extracting ATAC-seq reads longer than 150 bp (Figure 7A). Nucleosomes surrounding the TSS were defined as *-1 nucleosomes* within 350 bp upstream and as *+1 nucleosomes* within 250 bp downstream, and the distance between the +1 to -1 nucleosomes was defined as the *nucleosome-free region*. When compared with quiescent hepatocytes, there was a median downstream shift of 9 bp in 1-week ($P = $ 2.60E-13) and an upstream shift of 19 bp in 4-week ($P < $ 1E-15) repopulating hepatocytes for the -1 nucleosomes, although there was no significant shift in +1 nucleosome positioning (Figure 7B, Supplementary Table 6). As a result, there was a global increase of promoter openness in 4-week repopulating hepatocytes as the distance between +1 to -1 nucleosomes increased, while the nucleosome-free region was shorter in 1-week regenerating liver compared with the quiescent state. The difference in genome-wide promoter openness in repopulating hepatocytes at various time points suggests that accessibility of divergent functional regions could be regulated
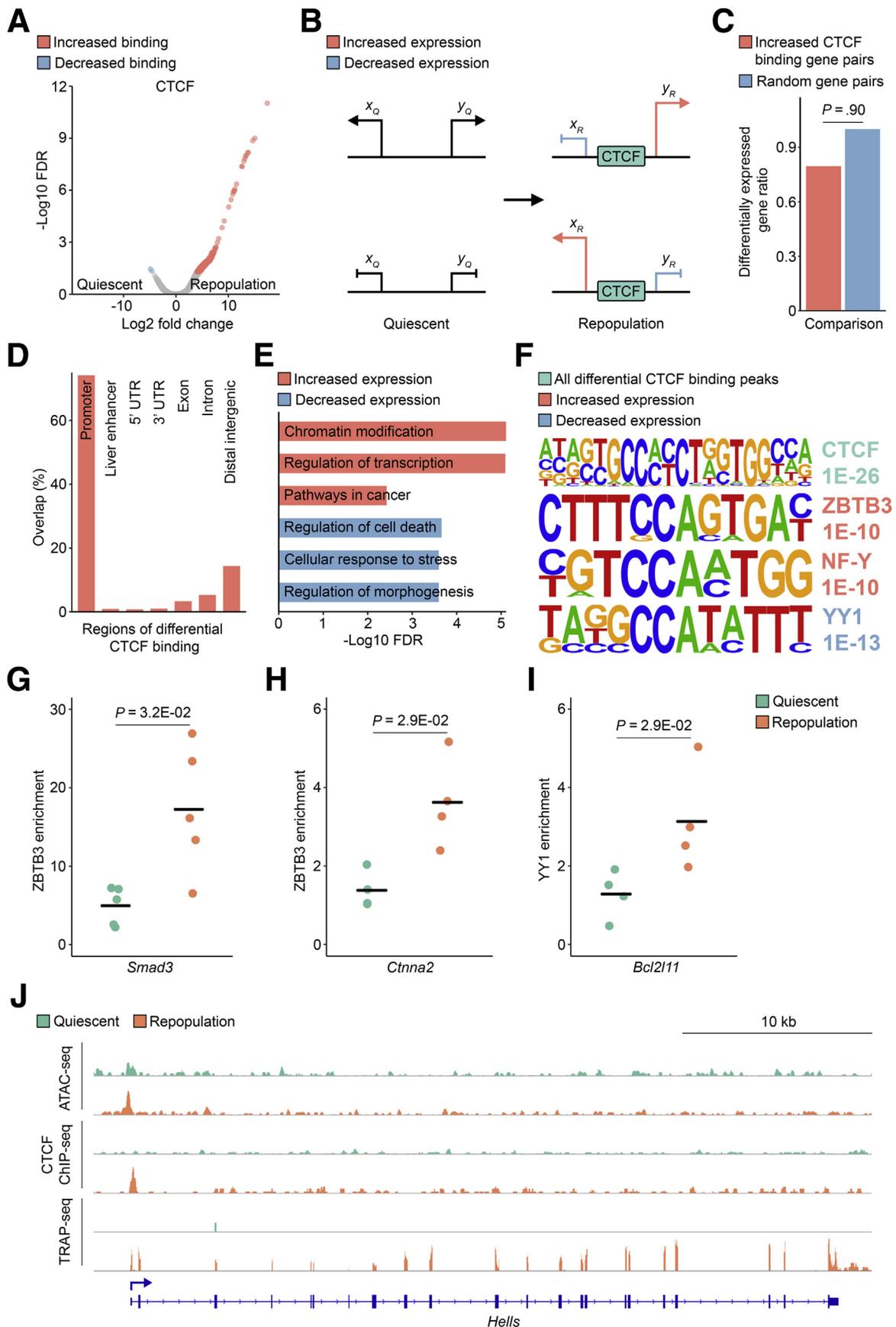
differentially during liver regeneration. Indeed, the nucleosome-free region constitutes only 17.5% of regions with increased accessibility in week 1, but 45.6% in week 4 (Figure 7C), whereas 39.0% of week 1 and only 26.9% of week 4 regions that became more open were in the distal intergenic regions (Figure 7D). On the other hand, chromatin regions with decreased accessibility show a similar distribution between the nucleosome-free region and distal intergenic regions. These observations indicate that the increase of chromatin accessibility occurs mainly at distal genomic areas in 1-week and around the TSS in 4-week repopulating hepatocytes.

To evaluate the association of TSS accessibility and gene expression, we extracted the top 500 up-regulated and down-regulated genes in repopulation[15] and calculated the change in the length of the nucleosome-free region between quiescent and regenerating hepatocytes as a surrogate for differential TSS accessibility. We only observed a significant increase ($P = 1.15E-2$) of the +1 to -1 nucleosome distance in genes activated in week 4 when compared with quiescent hepatocytes, although no significant change in the nucleosome-free region was present in genes up-regulated in week 1 or genes down-regulated in weeks 1 and 4 (Figure 7E and F). It is likely that eviction or repositioning of the -1 nucleosomes could expose transcription factor binding sequences and allow access of the transcriptional machinery to the TATA box for gene activation in regenerating hepatocytes.[71] Altogether, analysis of the nucleosome structure implies nucleosome reorganization could affect gene activation, but not inhibition, during liver repopulation.

Discussion

Gene regulation is tightly controlled by a complex network integrating transcription factor binding and transcriptional apparatus assembly, chromatin structure, epigenetic modifications, and even intrachromosomal and interchromosomal interactions.[9,10] In this study, we investigated the association of chromatin accessibility, nucleosome properties, transcription factor occupancy, and gene expression[15] to delineate the multidimensional framework of transcriptional regulation in the repopulating liver. By implementing the INTACT method[19] to express SUN1-GFP in the $Fah^{-/-}$ model, we successfully performed cell type–specific isolation of only repopulating hepatocyte nuclei followed by ATAC-seq to identify changes of the chromatin landscape (Figures 1 and 2). Integration of TRAP-seq[15] with ATAC-seq determined that gene activation corresponds with increased promoter openness, while gene inhibition is linked to decreased promoter and enhancer accessibility (Figure 3C). We also corroborated previous findings that cell cycle, DNA synthesis, proliferation, and glutathione metabolism are activated whereas complement and coagulation, biosynthesis, and metabolic pathways are inhibited during liver repopulation (Figures 2D and 3B and C).[12,15] In addition, de novo motif analysis identified enrichment of CTCF and HNF4α binding sequences in regions with increased and decreased accessibility in repopulating hepatocytes, respectively (Figure 4). We further validated the differential occupancy of both factors in the repopulating liver with ChIP-seq and observed decreased HNF4α binding at liver enhancers[30] (Figure 5) and increased CTCF binding at promoters (Figure 6). Integrated ATAC-seq, ChIP-seq, and TRAP-seq analysis suggests that CTCF recruits cofactors to activate genes involved in chromatin organization and replication and inhibit genes in the regulation of cell death (Figure 6E–J). On the other hand, loss of HNF4α occupancy at liver enhancers decreases the expression of hepatocyte-enriched genes that are crucial in establishing liver homeostasis and function (Figure 5C–E).

In general, 40% of CTCF binding sites occur in intergenic regions distant to the TSS, whereas 35% of CTCF sites are found in promoters.[30,44] Interestingly, the vast majority (75%) of sites with increased CTCF occupancy are located within promoters in the repopulating liver (Figure 6D). In fact, CTCF can function as a direct transcriptional repressor at the *Myc* promoter[72] and as an activator of the amyloid precursor protein promoter,[73] strengthening the notion that CTCF plays a more localized role as a transcriptional regulator in the repopulating liver via recruitment of cofactors. Up-regulation of CTCF in liver cancer is associated with poor survival, likely through the activation of forkhead box M1, to stimulate cell growth and tumor metastasis.[74] The CTCF–forkhead box M1 axis could be triggered during liver regeneration to promote hepatocyte proliferation.[75] Increased CTCF activity at the *Myc* promoter[76] or decreased CTCF repression at the *Myc* enhancer[77] both have

Figure 6. (*See previous page*). **CTCF binding is increased at promoters in the repopulating liver.** (*A*) The 1306 peaks show increased, whereas only 2 peaks show decreased, CTCF occupancy during repopulation. (*B*) Schematic to test the insulator function of increased CTCF binding to differentially regulate expression of the flanking genes.[55] (*C*) Promoters flanking sites of increased CTCF occupancy are not more enriched for differentially expressed genes compared with random gene pairs in the genome. A Fisher exact test was used to test the differentially expressed gene ratios from the 2 groups of gene pairs. (*D*) A total of 75% of the genomic regions with increased CTCF binding are within 1 kb upstream and downstream of the TSS (promoter), and only 13 peaks overlap with liver enhancers.[30] (*E*) Enriched pathways of increased chromatin accessibility, CTCF occupancy, and increased (red) or decreased (blue) gene expression during liver repopulation. (*F*) Motif enrichment analysis identifies an overrepresentation of CTCF motif in differentially bound regions, ZBTB3, and NF-Y motifs at sites with increased CTCF occupancy associated with gene activation, and the YY1 motif at sites with increased CTCF occupancy associated with gene inhibition. Numbers denote FDR. (*G* and *H*) ZBTB3 occupancy is increased in the repopulating liver at the (*G*) *Smad3* and (*H*) *Ctnna2* promoters, 2 genes with increased CTCF occupancy and expression during regeneration. (*H*) YY1 occupancy is increased in the repopulating liver at the *Bcl2l11* promoter, a gene with increased CTCF binding and decreased expression during regeneration. (*I*) Representative tracks (n = 2 for ATAC-seq and ChIP-seq, n = 4 for TRAP-seq) of chromatin accessibility, CTCF occupancy, and transcript levels at *Hells*, the locus with the strongest increase in CTCF binding. UTR, untranslated region.
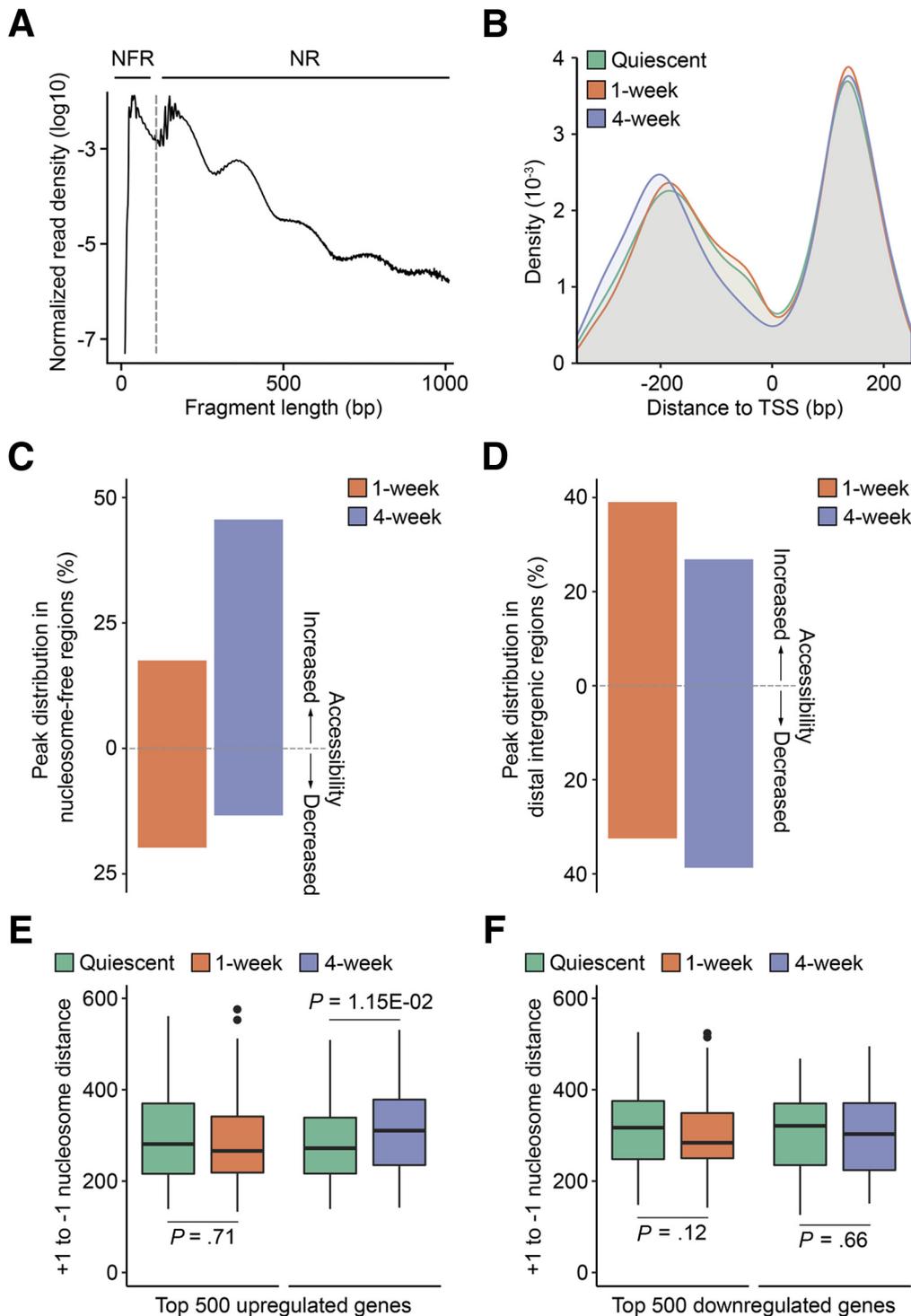
Figure 7. Decreased nucleosome density is associated with increased gene expression[15] in repopulating hepatocytes. (*A*) Schematic for identifying nucleosome positioning information with NucleoATAC.[102] (*B*) Globally, -1 nucleosomes have an upstream shift away from the TSS in 4-week repopulating hepatocytes, whereas +1 nucleosome positioning is constant during liver repopulation. (*C* and *D*) Distribution of regions with differential accessibility in (*C*) the nucleosome-free region that is within 350 bp upstream and 250 bp downstream of the TSS and (*D*) distal intergenic regions in 1- and 4-week repopulating hepatocytes. (*E*) The top 500 up-regulated genes have an increased +1 to -1 nucleosome distance in 4-week but not 1-week repopulating hepatocytes when compared with quiescent hepatocytes. (*F*) The top 500 down-regulated genes are not associated significantly with changes in +1 to -1 nucleosome distance in repopulating compared with quiescent hepatocytes. Permutation tests with 10,000 iterations were used to compare the nucleosome distance in repopulating and quiescent hepatocytes. NFR, nucleosome-free reads; NR, nucleosomal reads.

been observed in cancer cells that lead to increased MYC expression. The high tumor mutational burden of CTCF results in abnormal occupancy,[78,79] and thus the cofactors and targets of CTCF could be different in the regenerating liver and liver cancer. The multitude of CTCF functions warrants further investigation to understand its contribution to mediating chromatin structure and organization in the context of liver repopulation. Specifically, CTCF also acts as an insulator to block enhancer–promoter interactions,[43] a factor that promotes long-range chromatin looping,[45] and a TAD boundary protein that defines expression domains for tight transcriptional control.[44] Future experiments to detect changes in chromatin interactions via chromosome conformation capture[80] would be valuable in determining whether differential CTCF occupancy affects 3-dimensional chromatin organization during liver repopulation.

A   Quiescent hepatocytes
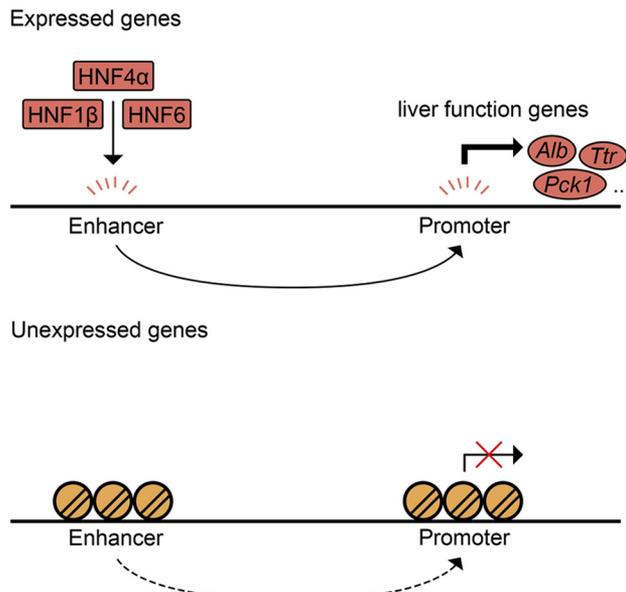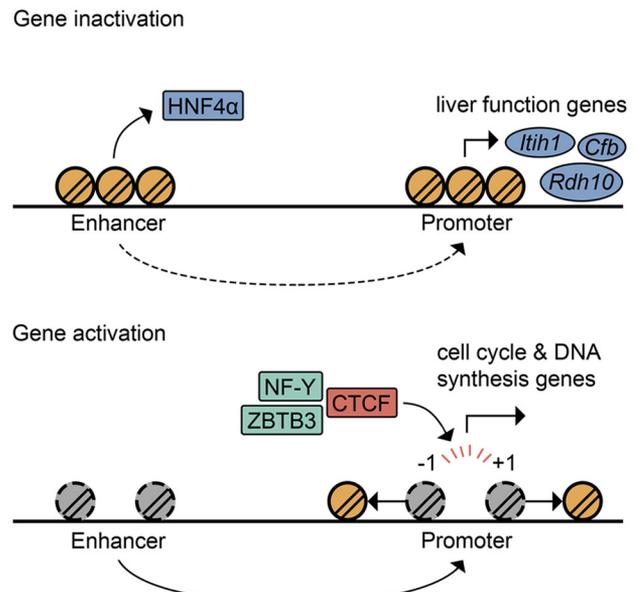
B   Repopulating hepatocytes

Figure 8. Model of transcriptional regulation in repopulating hepatocytes. (*A*) Access to enhancers allows liver-enriched transcription factors to maintain quiescent hepatocytes in the differentiated state (*top*). In contrast, chromatin-dense enhancers and promoters prevent transcription factor binding to inhibit gene expression of cell-cycle genes (*bottom*). (*B*) During liver repopulation, decreased accessibility of liver enhancers[30] in conjunction with more closed promoters prevents binding of transcription factor and assembly of the transcriptional machinery at hepatocyte-specific liver function genes, resulting in a less differentiated transcriptomic and epigenomic profile in the repopulating cells (*top*). Conversely, the promoter regions of cell-cycle genes become more open, with increased +1 to -1 distance and increased CTCF occupancy at the promoter, allowing increased expression of genes involved in the cell-cycle and DNA synthesis pathways (*bottom*).

The mechanisms of increased CTCF and decreased HNF4α binding in the repopulating liver are also not fully understood. In the current study, we infer that a more open chromatin state at specific promoters correlates with the accessibility of CTCF to its binding sites; however, we have not assessed causality. Previous work found that enrichment of thymidine at the 18th position in the CTCF motif reduces its affinity, where low-affinity sites are more sensitive to loss of CTCF binding during mouse embryonic stem cell differentiation.[55] In addition, it is likely that changes in DNA methylation influence differential CTCF occupancy because methylated CpGs in the CTCF recognition site can prevent its binding.[81,82] Demethylation at specific promoter regions therefore could increase CTCF occupancy during liver repopulation. In the case of reduced HNF4α occupancy at liver-specific enhancers in the regenerating liver, part of this effect can be explained by reduced expression of HNF4α itself. Furthermore, HNF4α could be regulated posttranscriptionally via phosphorylation by kinases such as protein kinase A and C, as well as adenosine monophosphate–activated protein kinase to decrease its DNA binding activity or nuclear localization.[83] Activation of the mitogen-activated protein kinase signaling pathway also was shown to inhibit *Hnf4a* expression via activation of the transcription factor JUN.[83,84] The fact that enrichment of the DNA synthesis pathways is observed only in 1-week repopulating livers

and that the *Hnf4a* transcript level is unchanged in week-1 but reduced in week-4 hepatocytes strengthens the notion that activation of cell growth and proliferation occur early after the initiation of liver repopulation, followed by a later reduction of *Hnf4a* transcription. Future studies using, for instance, targeted degradation of CTCF[85] or HNF4α could be implemented to identify potential promoters and inhibitors of liver repopulation. Technologies such as complementary DNA[8] or clustered regularly interspaced short palindromic repeats[86,87] screens also could be used to evaluate the effectors downstream of CTCF activation and HNF4α inhibition.

In summary, we propose the following model to explain the transcriptional adaptations that accompany liver repopulation (Figure 8): during hepatocyte replication, the promoters of selected genes become more open owing to an increased distance between histones at +1 to -1, increasing accessibility for CTCF, transcription factor recruitment, and transcriptional machinery assembly to activate genes that regulate cell cycle, DNA synthesis, and proliferation pathways. On the other hand, decreased enhancer accessibility in conjunction with suppression of *Hnf4a* expression reduces or prevents HNF4α binding, and possibly that of other hepatocyte nuclear factors, to liver enhancers, resulting in repression of hepatocyte metabolic and biosynthetic function genes.

Methods

Plasmid Construction

The generation of the pKT2/Fah-Sun1-Gfp//SB plasmid was described previously.[15] The nuclear envelope SUN1-tagged GFP (SUN1-GFP) plasmid was a generous gift from Dr Jeremy Nathans (Johns Hopkins University, Baltimore, MD). We amplified the SUN1-GFP insert by PCR amplification with the primers MfeI-Sun1-F and BsiW1-Sun1-R and subcloned it into the vector pKT2/Fah-mCa//SB[7] to construct pKT2/Fah-Sun1-Gfp//SB. This construct uses the Sleeping Beauty transposase for stable transgene integration into the genome. The plasmid was prepared with the GenElute HP Plasmid Maxiprep Kit (NA0310-1KT; MilliporeSigma) for endotoxin-free maxi-scale DNA extraction and purification.

Mouse Studies

Fah$^{-/-}$ mice were maintained on 7.5 mg/L NTBC (Swedish Orphan Biovitrum, Stockholm, Sweden) in the drinking water. Hydrodynamic tail-vein injection[86] of 10 μg of pKT2/Fah-Sun1-Gfp//SB was performed followed by NTBC withdrawal for 1 week (n = 2) or 4 weeks (n = 2) to induce liver repopulation.[15] The *Rosa*$^{LSL-Sun1-GFP}$ mice[19,88] were kindly provided by Dr Jeremy Nathans and were tail-vein–injected with AAV8.TBG.PI.Cre.rBG (Penn Vector Core, Philadelphia, PA) at 1×10^{11} virus particles per mouse to ablate the loxP-stop-loxP cassette only in hepatocytes. Livers from these mice were harvested 1 week after viral injection and served as quiescent controls. All studies were performed in 8- to 12-week-old mice.

Immunofluorescence Staining

Liver lobes were isolated, fixed in 4% paraformaldehyde overnight at 4°C, embedded in paraffin, and sectioned. Tissue sections were deparaffinized with xylene and rehydrated with serial incubation of 100%, 95%, 80%, and 75% ethanol followed by phosphate-buffered saline (PBS). Antigen retrieval was performed in Tris/EDTA buffer (10 mmol/L Tris, 1 mmol/L EDTA, pH 9.2) in a pressure cooker (2100 Antigen Retriever; Aptum Biologics Ltd, Southampton, UK) and cooled to room temperature. Slides then were blocked with blocking buffer (PBS, 1% bovine serum albumin) for an hour followed by overnight incubation of antibodies in the blocking buffer at 4°C in a humidified chamber. Three washes of PBS were performed the next day followed by incubation with secondary antibodies at room temperature for 2 hours. Goat anti-GFP antibody (ab6673, 1:300; Abcam, Cambridge, United Kingdom) and rabbit anti-FAH antibody (ab81087, 1:600; Abcam) were used to label repopulating hepatocytes from *Fah*$^{-/-}$ mice after 1 and 4 weeks of repopulation and all hepatocytes from *Rosa*$^{LSL-GFP-L10a}$ mice injected with AAV8-TBG-Cre. 4',6-diamidino-2-phenylindole (DAPI) (B1098, 1:10,000; BioVision) was used to label nuclei.

Hepatocyte Nuclei Isolation

Liver was homogenized in 10 mL hypotonic buffer (10 mmol/L Tris-HCL, pH 7.5, 2 mmol/L MgCl$_2$, 3 mmol/L CaCl$_2$) on ice. The homogenate was filtered with a 100-μm filter and sedimented at $400 \times g$ for 10 minutes at 4°C. Ten milliliters of hypotonic buffer with 10% glycerol was used to resuspend the pellet followed by dropwise addition of 10 mL cell lysis buffer (hypotonic buffer, 10% glycerol, 1% octylphenoxypolyethoxyethanol CA-630). The homogenate was incubated for 5 minutes on ice and sedimented at $600 \times g$ for 5 minutes at 4°C. Nuclei were washed with lysis buffer again and quantified in a hemocytometer. Isolated nuclei were labeled with an Alexa Fluor 647 anti-GFP antibody (338006, clone FM264G, 1:25; BioLegend, San Diego, CA) for 30 minutes and 2 μg/mL DAPI immediately before sorting. After gating for the DAPI-positive signal, nuclei double-positive for GFP and AF647 were sorted with a BD FACSAria II (BD Biosciences, San Jose, CA), and only tetraploid hepatocyte nuclei were collected for further experiments.

ATAC-Seq Library Generation

ATAC-seq libraries were generated as previously described.[20] Briefly, transposition was performed on 25,000 sorted tetraploid nuclei at 37°C for 30 minutes followed by DNA purification with the MinElute Reaction Cleanup Kit (28206; Qiagen, Hilden, Germany). DNA fragments were PCR preamplified for 5 cycles initially, and one tenth of the volume (5 μL) was removed for qPCR amplification for 20 cycles. A plot of R vs cycle number was generated and the number of cycles required to reach one third of the maximum R was determined for each sample. The pre-amplified ATAC-seq libraries then were amplified for the calculated additional cycles. Agencourt AMPure XP beads (A63881; Beckman Coulter) were used for size selection to generate the final libraries.[89] Library quality was assessed with an Agilent High-Sensitivity DNA Bioanalyzer (5067-4626; Agilent Technologies), and quantity was measured with KAPA Library Quantification Kits (KK4835; KAPA Biosystems).

ATAC-Seq Peak Calling

ATAC-seq libraries were paired-end sequenced on an Illumina HiSeq 4000 (Illumina, San Diego, CA) with 50, 75, or 100 reads. Reads then were trimmed to 50 bp with Cutadapt[90] and peaks were called with the ATAC-Seq/DNase sequencing pipeline (https://github.com/kundajelab/atac_dnase_pipelines). Briefly, the trimmed fastq files were aligned to the mouse genome (mm10) with Bowtie2[91] followed by removal of PCR duplicates and mitochondrial reads. Bam files of the same biological sample from various technical replicates then were merged with Samtools[92] and duplicated reads were removed. The filtered reads were shifted 5 bp for + strands and 4 bp for - strands to adjust for the transposase binding sites.[20] Nucleosome-free reads were identified with the R package ATACseq quality control (QC) using a random Forest classifier[93] followed by peak calling with MACS2.[94] Artifact signals then were removed according to the mm10 empiric blacklist regions.[95] The irreproducible discovery rate framework was used to compare all pairs of biological replicates to identify

reproducible peaks that passed a threshold of 10% for all pairwise analyses. The conservative peak set for each sample was identified by selecting the longest peak list from all pairs that passed the 10% irreproducible discovery rate cutoff value.

ATAC-Seq Peak Quality Assessment

To ensure the ATAC-seq peaks generated from the sorted nuclei were of high quality, the R package ATACseqQC[93] was used for assessment. We first visualized the insert size distribution to confirm the presence of distinct periodicity of approximately 175 bp associated with nucleosome patterning in all samples, indicating the DNA fragments are protected by integer multiples of nucleosomes.[20] The signal intensity of nucleosome-free reads and nucleosomal reads also was averaged across all TSS to examine evidence that no overfragmentation was introduced during hepatocyte nuclei isolation, sorting, or ATAC-seq library preparation.

ATAC-Seq Differential Peak Analysis

The R package ATACseqQC[93] was used to split the aligned bam files into nucleosome-free reads and nucleosomal reads. The R package DiffBind[96] was used to identify differential accessible peaks from the nucleosome-free reads. The overlapping regions from the ATAC-seq peak sets for each sample were identified and merged into nonoverlapping regions. Read counts for each region were quantified with a dba.count (score = DBA_SCORE_TMM_READS_FULL, fragmentSize=0, bScaleControl=F, filter=0, bRemoveDuplicates=F, bUseSummarizeOverlaps=T). Peaks identified in both biological replicates in the same conditions were used for differential analysis with dba.analyze (method=DBA_EDGER, bSubControl=F, bTagwise=T) in conjunction with edgeR.[97] Peaks with an absolute fold change of ≥ 1.5 and FDR of ≤ 0.05 were identified as significant differentially accessible regions.

Integrative Analysis of TRAP-Seq and ATAC-Seq Data

To identify chromatin accessibility and gene expression that changed in the same direction at the same time point (concordant genes), the differentially accessible peaks were first annotated to the nearest TSS with the R package ChIPseeker.[98] Genes with differential expression during liver repopulation were obtained from a previous study that used translating ribosome affinity purification followed by RNA-sequencing (TRAP-seq).[15] The concordant ATAC-seq peaks and TRAP-seq genes were identified and the expected overlap and significance was calculated with a hypergeometric test. To evaluate the association of chromatin accessibility and gene expression changes, all chromatin regions were stratified into regions with increased, decreased, or unchanged accessibility, with the cut-off value of an absolute fold change of ≥ 1.5 and FDR of ≤ 0.05. For promoter accessibility and gene activity association analysis, regions within 1 kb upstream and downstream of the TSS were identified and annotated to the nearest genes with the R package ChIPseeker.[98] The corresponding expression change at the same time point was extracted from TRAP-seq[15] and normalized by subtracting the mean $\log_2$ fold change of the unchanged from the increased and decreased chromatin accessibility groups. The normalized expression fold change of the nearest genes in the differentially accessible promoters was compared with that in the unchanged accessibility promoters with a 1-sample t test. For enhancer accessibility and gene expression association studies, liver- and cerebellum-specific enhancers and their putative targets were obtained from a previous study.[30] Briefly, regions with the presence of H3K4me1 but the absence of H3K4me3 ChIP-seq peaks were identified as putative enhancers and refined with a chromatin-signature–based enhancer predictor. Enhancer–promoter units were identified by calculating the correlation of H3K4me1 and RNA polymerase II ChIP-seq peak strength along each chromosome. All possible promoter and enhancer pairs with a greater than 0.23 Spearman correlation coefficient were identified as linked enhancer–promoter units. Gene expression fold changes were normalized as described earlier, and the normalized gene expression fold change of the enhancer target genes in the differentially accessible enhancers was compared with that in the unchanged accessibility enhancers with a 1-sample t test.

Transcription Factor Motif Enrichment Analysis

ATAC-seq peaks are separated into promoter and liver enhancer[30] regions and Homer[99] is used to identify enrichment of de novo motifs with the function findMotifsGenome.pl (mm10 -size given). Motifs with a P value of lower than 1E-12 are considered significant to reduce the number of false positives. FDR also is calculated with each significant motif. To ensure the identified motifs are enriched in ATAC-seq peaks with different accessibility, motif frequency surrounding 500 bp upstream and downstream of the peak center from all identified irreproducible discovery rate peaks in quiescent hepatocytes and differentially accessible regions in repopulating cells was extracted. The difference in motif frequency distribution of regenerating and quiescent samples then was calculated with a Kolmogorov–Smirnov test.

ChIP-Seq Library Generation

A total of 100 mg of quiescent (n = 2) and repopulating (n = 2) liver tissue was finely chopped with a razor blade and cross-linked in 1% formaldehyde for 10 minutes followed by the addition of 2.5 mol/L glycine and incubation for 5 minutes at room temperature. Tissues were sedimented, washed with cold PBS, and Dounce-homogenized in cold ChIP cell lysis buffer (10 mmol/L Tris-HCl pH 8.0, 10 mmol/L NaCl, 3 mmol/L MgCl₂, 0.5% IGEPAL CA-630, and protease inhibitor) on ice. After incubation at 4°C for 5 minutes, nuclei were pelleted and resuspended in nuclear lysis buffer (50 mmol/L Tris-HCl pH 8.1, 1% sodium dodecyl sulfate [SDS], 5 mmol/L EDTA, and protease inhibitor). Nuclei were sonicated with a Bioruptor (Diagenode, Denvill, NJ) for 2 rounds of 7.5 minutes each. A total of 10 μg sheared DNA was incubated with anti-CTCF (2 μg, 07-729;

Millipore, Burlington, MA) or anti-HNF4α (2 μg, ab181604; Abcam) antibodies in dilution buffer (16.7 mmol/L Tris-HCl pH 8.1, 167 mmol/L NaCl, 0.01% SDS, 1.1% Triton-X 100, and protease inhibitor) at 4°C overnight. Protein A agarose beads also were washed with cold dilution buffer 3 times and incubated with blocking buffer (10 mg/mL bovine serum albumin, ChIP dilution buffer, and protease inhibitor) at 4°C overnight. Sheared DNA incubated with antibody and blocked protein A agarose were incubated at 4°C for 1 hour the next day and washed at room temperature with buffers Tris-SDS-EDTA I (20 mmol/L Tris-HCl pH 8.1, 150 mmol/L NaCl, 2 mmol/L EDTA, 0.1% SDS, and 1% Triton X-100), Tris-SDS-EDTA II (20 mmol/L Tris-HCl pH 8.1, 500 mmol/L NaCl, 2 mmol/L EDTA, 0.1% SDS, and 1% Triton X-100), ChIP buffer III (10 mmol/L Tris-HCl pH 8.1, 0.25 mol/L LiCl, 1 mmol/L EDTA, 1% NP-40, and 1% deoxycholate), and Tris-EDTA (10 mmol/L Tris-HCl pH 8.1, 1 mmol/L EDTA). Chromatin was eluted with elution buffer (1% SDS, 0.1 mol/L NaHCO$_3$) twice and incubated with 0.2 mol/L NaCl at 65°C overnight to reverse the cross-links. Digestion was performed with 10 mg/mL proteinase K in 40 mmol/L Tris-HCl pH 7.5 and 10 mmol/L EDTA to purify CTCF- or HNF4α-bound and input DNA. ChIP-seq libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit for Illumina (E7645S; New England BioLabs, Ipswich, MA) and Agencourt AMPure XP beads were used for size selection to generate the final libraries. Library quality was assessed with an Agilent (Beverly, MA) High-Sensitivity DNA Bioanalyzer (5067-4626; Agilent Technologies), and quantity was measured with KAPA Library Quantification Kits (KK4835; KAPA Biosystems).

ChIP-Seq Data Analysis

ChIP-seq libraries were sequenced on an Illumina HiSeq 4000 (Illumina) with 100 single-end reads and aligned to the mm10 genome with STAR.[100] Bam files from various technical replicates of the same biological sample were merged with Samtools.[92] Peak calling was performed with Homer[99] and differential occupancy analysis was performed with the R package DiffBind.[96] Read counts for each peak were quantified with dba.count (score=DBA_SCORE_TMM_MINUS_FULL, bUseSummarizeOverlaps=TRUE) and differential analysis was identified with dba.analyze (method=DBA_EDGER, bSubControl=T, bTagwise=F) in conjunction with edgeR.[97]

ChIP-qPCR

ChIP was performed with 5 μg of anti-ZBTB3 (ab106536) and 2 μg of YY1 (ab109237) antibodies with 10 μg of sheared DNA from quiescent and 4-week repopulating livers as described earlier. Input and immunoprecipitated DNA were purified with phenol–chloroform extraction followed by qPCR with the primer sets ZBTB3-ChIP-Ctnna2-qPCR-F1 and -R1, ZBTB3-ChIP-Smad3-qPCR-F1 and -R1, YY1-ChIP-Bcl2l11-qPCR-F1 and -R1, YY1-ChIP-Igf2r-qPCR-F1 and -R1, and 40S-F2 and -R2. All primer sequences are listed in Supplementary Table 7.

CTCF Differential Expression Insulator Analysis

Increased CTCF occupancy during liver repopulation could prevent distal regulatory regions to activate only 1 of the flanking promoters surrounding a CTCF binding site, therefore leading to a larger difference in gene expression levels. We define this differential expression insulator function, in which a gene pair is expressed as either high or low without the presence of CTCF, but only 1 flanking gene showed a decrease in gene expression after binding of CTCF. An insulator strength score was calculated for all significantly gained (fold change, $\geq$1.5; FDR, $\leq$0.05) CTCF peaks in the repopulating liver as previously described.[55] Briefly, CTCF sites with divergent flanking promoters within 50 kb were identified and the corresponding gene expression levels from quiescent and 4-week repopulating hepatocytes were extracted from published TRAP-seq.[15]

Low expressors, in which reads per kilobase of transcript per million mapped–normalized read counts are 0 across all samples, were filtered followed by calculation of a rank percentile based on the reads per kilobase of transcript per million mapped for each gene. Let x_Q and y_Q be the expression percentile in the quiescent hepatocytes, and x_R and y_R are the expression percentile in the 4-week repopulating hepatocytes. The insulator strength score is calculated by taking the maximum value of $x_Q \times y_Q \times x_R \times (1 - y_R)$ and $x_Q \times y_Q \times (1 - x_R) \times y_R$. A differential expression insulator function will have one of the following effects: (1) increased x_R and decreased y_R: in this case, $x_Q \times y_Q \times x_R \times (1 - y_R)$ will be the largest, or (2) decreased x_R and increased y_R: in this case, $x_Q \times y_Q \times (1 - x_R) \times y_R$ will be the largest. Gained CTCF sites with the top 25% insulator strength scores were categorized as strong insulators. Random gene pairs not flanked by CTCF within 50 kb were used as controls, and a differential expression insulator score for each gene pair was calculated as described earlier. The number of significant (FDR, $\leq$ 0.05) and nonsignificant (FDR, $>$ 0.05) differential expressions of the flanking genes were identified for all strong insulators from increased CTCF binding and random genomic regions. Finally, we used the Fisher exact test to examine the likelihood of gained CTCF sites to contain more significantly changed genes when compared with that of control regions.

Nucleosome Location Analysis With ATAC-Seq

MAC[94] was used to identify broad peaks from all aligned bam files including nucleosome-free reads and nucleosome-containing reads from ATAC-seq. Broad peaks then were processed with BEDtools[101] to extend the peaks (BEDtools slop -b 200), sorted by genomic positions (sort -k1,1 -k2,2n), and overlapping reads were merged (BEDtools merge). The nucleosome position was identified with NucleoATAC[102] from the aligned bam and broad peak files. The closest nucleosomes with respect to TSS were identified, and those within 350 bp upstream and 250 bp downstream of the TSS were identified as the -1 and +1 nucleosomes, respectively.

Nucleosome Positioning Analysis

The distance of +1 to -1 nucleosomes was calculated for each transcript. We used the Kolmogorov–Smirnov test to compare the +1 and -1 nucleosome distribution differences between quiescent and repopulating hepatocytes, respectively. To analyze the association between gene activity and nucleosome positioning, transcriptomic changes in repopulating hepatocytes[15] were first stratified into 3 categories: top 500 up-regulated (fold change, $\geq$1.5; FDR, $\leq$0.05), top 500 down-regulated (fold change, $\geq$1.5; FDR, $\leq$0.05), and unchanged (absolute fold change, <1.5; or FDR, >0.05) genes. The distances between the +1 to -1 nucleosomes were calculated for each gene and differential positioning was performed by comparing the distance in quiescent vs regenerating hepatocytes in the up-regulated, down-regulated, and unchanged gene expression groups, respectively, with a permutation test (n = 10,000).

Statistical Analysis

EdgeR[97] was used for all high-throughput sequencing data analyses. For the integrative TRAP-seq and ATAC-seq analysis, a hypergeometric test was used for identifying the significance of overlapping gene sets, and a 1-sample t test was used to compare the difference between normalized gene expression fold change in differentially accessible promoter and enhancer peaks, respectively. A Kolmogorov–Smirnov test was performed for global distribution change of +1 and -1 nucleosome positioning and a permutation test (n = 10,000) was performed to test the change in +1 to -1 nucleosome distance of genes with differential expression.

Study Approval

The animal experiments performed in this study were reviewed and approved by the Institutional Animal Care and Use Committee of the Office of Animal Welfare at the University of Pennsylvania.

References

1. Trefts E, Gannon M, Wasserman DH. The liver. Curr Biol 2017;27:R1147–R1151.
2. Michalopoulos GK, DeFrances MC. Liver regeneration. Science 1997;276:60–66.
3. Lee W. Etiologies of acute liver failure. Semin Liver Dis 2008;28:142–152.
4. Grompe M, al-Dhalimy M, Finegold M, Ou CN, Burlingame T, Kennaway NG, Soriano P. Loss of fumarylacetoacetate hydrolase is responsible for the neonatal hepatic dysfunction phenotype of lethal albino mice. Genes Dev 1993;7:2298–2307.
5. Russo PA, Mitchell GA, Tanguay RM. Tyrosinemia: a review. Pediatr Dev Pathol 2001;4:212–221.
6. Overturf K, Al-Dhalimy M, Tanguay R, Brantly M, Ou C-N, Finegold M, Grompe M. Hepatocytes corrected by gene therapy are selected in vivo in a murine model of hereditary tyrosinaemia type I. Nat Genet 1996; 12:266–273.
7. Wangensteen KJ, Wilber A, Keng VW, He Z, Matise I, Wangensteen L, Carson CM, Chen Y, Steer CJ, McIvor RS, Largaespada DA, Wang X, Ekker SC. A facile method for somatic, lifelong manipulation of multiple genes in the mouse liver. Hepatology 2008; 47:1714–1724.
8. Wangensteen KJ, Zhang S, Greenbaum LE, Kaestner KH. A genetic screen reveals Foxa3 and TNFR1 as key regulators of liver repopulation. Genes Dev 2015;29:904–909.
9. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. Cell 2013;152:1237–1251.
10. Lelli KM, Slattery M, Mann RS. Disentangling the many layers of eukaryotic transcriptional regulation. Annu Rev Genet 2012;46:43–68.
11. Su AI, Guidotti LG, Pezacki JP, Chisari FV, Schultz PG. Gene expression during the priming phase of liver regeneration after partial hepatectomy in mice. Proc Natl Acad Sci U S A 2002;99:11181–11186.
12. White P, Brestelli JE, Kaestner KH, Greenbaum LE. Identification of transcriptional networks during liver regeneration. J Biol Chem 2005;280:3715–3722.
13. Yang D, Liu Q, Yang M, Wu H, Wang Q, Xiao J, Zhang Y. RNA-seq liver transcriptome analysis reveals an activated MHC-I pathway and an inhibited MHC-II pathway at the early stage of vaccine immunization in zebrafish. BMC Genomics 2012;13:319.
14. Min JS, DeAngelis RA, Reis ES, Gupta S, Maurya MR, Evans C, Das A, Burant C, Lambris JD, Subramaniam S. Systems analysis of the complement-induced priming phase of liver regeneration. J Immunol 2016; 197:2500–2508.
15. Wang AW, Wangensteen KJ, Wang YJ, Zahm AM, Moss NG, Erez N, Kaestner KH. TRAP-seq identifies cystine/glutamate antiporter as a driver of recovery from liver injury. J Clin Invest 2018;128:2297–2309.
16. Huang J, Schriefer AE, Yang W, Cliften PF, Rudnick DA. Identification of an epigenetic signature of early mouse liver regeneration that is disrupted by Zn-HDAC inhibition. Epigenetics 2014;9:1521–1531.
17. Sato Y, Katoh Y, Matsumoto M, Sato M, Ebina M, Itoh-Nakadai A, Funayama R, Nakayama K, Unno M, Igarashi K. Regulatory signatures of liver regeneration distilled by integrative analysis of mRNA, histone methylation, and proteomics. J Biol Chem 2017; 292:8019–8037.
18. Heiman M, Kulicke R, Fenster RJ, Greengard P, Heintz N. Cell type–specific mRNA purification by translating ribosome affinity purification (TRAP). Nat Protoc 2014;9:1282–1291.
19. Mo A, Mukamel EA, Davis FP, Luo C, Henry GL, Picard S, Urich MA, Nery JR, Sejnowski TJ, Lister R, Eddy SR, Ecker JR, Nathans J. Epigenomic signatures of neuronal diversity in the mammalian brain. Neuron 2015; 86:1369–1384.
20. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 2013;10:1213–1218.

21. Shin S, Wangensteen KJ, Teta-Bissett M, Wang YJ, Mosleh-Shirazi E, Buza EL, Greenbaum LE, Kaestner KH. Genetic lineage tracing analysis of the cell of origin of hepatotoxin-induced liver tumors in mice. Hepatology 2016;64:1163–1177.

22. Bell JB, Podetz-Pedersen KM, Aronovich EL, Belur LR, McIvor RS, Hackett PB. Preferential delivery of the Sleeping Beauty transposon system to livers of mice by hydrodynamic injection. Nat Protoc 2007; 2:3153–3165.

23. Aronovich EL, McIvor RS, Hackett PB. The Sleeping Beauty transposon system: a non-viral vector for gene therapy. Hum Mol Genet 2011;20:R14–R20.

24. Garrison BS, Yant SR, Mikkelsen JG, Kay MA. Post-integrative gene silencing within the Sleeping Beauty transposition system. Mol Cell Biol 2007;27:8824–8833.

25. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 2016; 44:D457–D462.

26. Zhang W, Liu HT. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res 2002;12:9–18.

27. Weglarz TC, Sandgren EP. Timing of hepatocyte entry into DNA synthesis after partial hepatectomy is cell autonomous. Proc Natl Acad Sci U S A 2000; 97:12595–12600.

28. Huang ZZ, Li H, Cai J, Kuhlenkamp J, Kaplowitz N, Lu SC. Changes in glutathione homeostasis during liver regeneration in the rat. Hepatology 1998;27:147–153.

29. Riehle KJ, Haque J, McMahan RS, Kavanagh TJ, Fausto N, Campbell JS. Sustained glutathione deficiency interferes with the liver response to TNF-α and liver regeneration after partial hepatectomy in mice. J Liver Disease Transplant 2013;1.

30. Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, Ren B. A map of the cis-regulatory sequences in the mouse genome. Nature 2012;488:116–120.

31. Park K, Atchison ML. Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta), that binds to the immunoglobulin kappa 3' enhancer and the immunoglobulin heavy-chain mu E1 site. Proc Natl Acad Sci U S A 1991; 88:9804–9808.

32. Walton KM, Rehfuss RP, Chrivia JC, Lochner JE, Goodman RH. A dominant repressor of cyclic adenosine 3',5'-monophosphate (cAMP)-regulated enhancer-binding protein activity inhibits the cAMP-mediated induction of the somatostatin promoter in vivo. Mol Endocrinol 1992;6:647–655.

33. Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet 2012;13:613–626.

34. Daugherty AC, Yeo RW, Buenrostro JD, Greenleaf WJ, Kundaje A, Brunet A. Chromatin accessibility dynamics reveal novel functional enhancers in C. elegans. Genome Res 2017;27:2096–2107.

35. Li M, Hada A, Sen P, Olufemi L, Hall MA, Smith BY, Forth S, McKnight JN, Patel A, Bowman GD, Bartholomew B, Wang MD. Dynamic regulation of transcription factors by nucleosome remodeling. Elife 2015;4.

36. Ballaré C, Castellano G, Gaveglia L, Althammer S, González-Vallinas J, Eyras E, Le Dily F, Zaurin R, Soronellas D, Vicent GP, Beato M. Nucleosome-driven transcription factor binding and gene regulation. Mol Cell 2013;49:67–79.

37. Strenkert D, Schmollinger S, Sommer F, Schulz-Raffelt M, Schroda M. Transcription factor-dependent chromatin remodeling at heat shock and copper-responsive promoters in Chlamydomonas reinhardtii. Plant Cell 2011;23:2285–2301.

38. Besnard A, Galan-Rodriguez B, Vanhoutte P, Caboche J. Elk-1 a transcription factor with multiple facets in the brain. Front Neurosci 2011;5:35.

39. Boros J, Donaldson IJ, O'Donnell A, Odrowaz ZA, Zeef L, Lupien M, Meyer CA, Liu XS, Brown M, Sharrocks AD. Elucidation of the ELK1 target gene network reveals a role in the coordinate regulation of core components of the gene regulation machinery. Genome Res 2009; 19:1963–1973.

40. Wuestefeld T, Pesic M, Rudalska R, Dauch D, Longerich T, Kang T-W, Yevsa T, Heinzmann F, Hoenicke L, Hohmeyer A, Potapova A, Rittelmeier I, Jarek M, Geffers R, Scharfe M, Klawonn F, Schirmacher P, Malek NP, Ott M, Nordheim A, Vogel A, Manns MP, Zender L. A Direct in vivo RNAi screen identifies MKK4 as a key regulator of liver regeneration. Cell 2013;153:389–401.

41. Klenova EM, Nicolas RH, Paterson HF, Carne AF, Heath CM, Goodwin GH, Neiman PE, Lobanenkov VV. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. Mol Cell Biol 1993;13:7612–7624.

42. Lobanenkov VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, Polotskaja AV, Goodwin GH. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. Oncogene 1990;5:1743–1753.

43. Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell 1999;98:387–396.

44. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 2012;485:376–380.

45. Splinter E, Heath H, Kooren J, Palstra R-J, Klous P, Grosveld F, Galjart N, de Laat W. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev 2006;20:2349–2354.

46. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature 2011; 479:74–79.

47. Xu D, Yang F, Yuan J-H, Zhang L, Bi H-S, Zhou C-C, Liu F, Wang F, Sun S-H. Long noncoding RNAs associated with liver regeneration 1 accelerates hepatocyte

proliferation during liver regeneration by activating Wnt/β-catenin signaling. Hepatology 2013;58:739–751.

48. Sladek FM, Zhong WM, Lai E, Darnell JE. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. Genes Dev 1990; 4:2353–2365.

49. Duncan SA, Manova K, Chen WS, Hoodless P, Weinstein DC, Bachvarova RF, Darnell JE. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. Proc Natl Acad Sci U S A 1994;91:7598–7602.

50. Babeu J-P, Boudreau F. Hepatocyte nuclear factor 4-alpha involvement in liver and intestinal inflammatory networks. World J Gastroenterol 2014;20:22–30.

51. Bonzo JA, Ferry CH, Matsubara T, Kim J-H, Gonzalez FJ. Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4α in adult mice. J Biol Chem 2012;287:7345–7356.

52. Walesky C, Gunewardena S, Terwilliger EF, Edwards G, Borude P, Apte U. Hepatocyte-specific deletion of hepatocyte nuclear factor-4α in adult mice results in increased hepatocyte proliferation. Am J Physiol Gastrointest Liver Physiol 2013;304:G26–G37.

53. Schrem H, Klempnauer J, Borlak J. Liver-enriched transcription factors in liver function and development. Part I: the hepatocyte nuclear factor network and liver-specific gene expression. Pharmacol Rev 2002;54:129–158.

54. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt KD. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 2016; 44:D733–D745.

55. Plasschaert RN, Vigneau S, Tempera I, Gupta R, Maksimoska J, Everett L, Davuluri R, Mamorstein R, Lieberman PM, Schultz D, Hannenhalli S, Bartolomei MS. CTCF binding site sequence differences are associated with unique regulatory and functional trends during embryonic stem cell differentiation. Nucleic Acids Res 2014;42:774–789.

56. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–29.

57. Zlatanova J, Caiafa P. CTCF and its protein partners: divide and rule? J Cell Sci 2009;122:1275–1284.

58. Boyle AP, Song L, Lee B-K, London D, Keefe D, Birney E, Iyer VR, Crawford GE, Furey TS. High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. Genome Res 2011; 21:456–464.

59. Lim J-H. Zinc finger and BTB domain-containing protein 3 is essential for the growth of cancer cells. BMB Rep 2014;47:405–410.

60. Dolfini D, Gatta R, Mantovani R. NF-Y and the transcriptional activation of CCAAT promoters. Crit Rev Biochem Mol Biol 2012;47:29–49.

61. Bhattacharya A, Deng JM, Zhang Z, Behringer R, de Crombrugghe B, Maity SN. The B subunit of the CCAAT box binding transcription factor complex (CBF/NF-Y) is essential for early mouse development and cell proliferation. Cancer Res 2003;63:8167–8172.

62. Oldfield AJ, Yang P, Conway AE, Cinghu S, Freudenberg JM, Yellaboina S, Jothi R. Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. Mol Cell 2014;55:708–722.

63. Donohoe ME, Zhang X, McGinnis L, Biggers J, Li E, Shi Y. Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. Mol Cell Biol 1999;19:7237–7244.

64. Kurisaki K, Kurisaki A, Valcourt U, Terentiev AA, Pardali K, Ten Dijke P, Heldin C-H, Ericsson J, Moustakas A. Nuclear factor YY1 inhibits transforming growth factor beta- and bone morphogenetic protein-induced cell differentiation. Mol Cell Biol 2003; 23:4494–4510.

65. Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, Abraham BJ, Cohen MA, Nabet B, Buckley DL, Guo YE, Hnisz D, Jaenisch R, Bradner JE, Gray NS, Young RA. YY1 is a structural regulator of enhancer-promoter loops. Cell 2017;171:1573–1588. e28.

66. Atchison L, Ghias A, Wilkinson F, Bonini N, Atchison ML. Transcription factor YY1 functions as a PcG protein in vivo. EMBO J 2003;22:1347–1358.

67. Wilkinson FH, Park K, Atchison ML. Polycomb recruitment to DNA in vivo by the YY1 REPO domain. Proc Natl Acad Sci U S A 2006;103:19296–19301.

68. Donohoe ME, Zhang L-F, Xu N, Shi Y, Lee JT. Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. Mol Cell 2007;25:43–56.

69. Bai L, Morozov AV. Gene regulation by nucleosome positioning. Trends Genet 2010;26:476–483.

70. Shivaswamy S, Bhinge A, Zhao Y, Jones S, Hirst M, Iyer VR. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. PLoS Biol 2008;6:e65.

71. Svaren J, Hörz W. Transcription factors vs nucleosomes: regulation of the PHO5 promoter in yeast. Trends Biochem Sci 1997;22:93–97.

72. Filippova GN, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G, Neiman PE, Collins SJ, Lobanenkov VV. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of

avian and mammalian c-myc oncogenes. Mol Cell Biol 1996;16:2802–2813.

73. Yang Y, Quitschke WW, Vostrov AA, Brewer GJ. CTCF is essential for up-regulating expression from the amyloid precursor protein promoter during differentiation of primary hippocampal neurons. J Neurochem 1999; 73:2286–2298.

74. Zhang B, Zhang Y, Zou X, Chan AW, Zhang R, Lee TK-W, Liu H, Lau EY-T, Ho NP-Y, Lai PB, Cheung Y-S, To K-F, Wong HK, Choy KW, Keng VW, Chow LM, Chan KK, Cheng AS, Ko BC. The CCCTC-binding factor (CTCF)-forkhead box protein M1 axis regulates tumour growth and metastasis in hepatocellular carcinoma. J Pathol 2017;243:418–430.

75. Xiang D, Liu C-C, Wang M-J, Li J-X, Chen F, Yao H, Yu B, Lu L, Borjigin U, Chen Y-X, Zhong L, Wangensteen KJ, He Z-Y, Wang X, Hu Y-P. Non-viral FoxM1 gene delivery to hepatocytes enhances liver repopulation. Cell Death Dis 2014;5:e1252.

76. Pu H, Zheng Q, Li H, Wu M, An J, Gui X, Li T, Lu D. CUDR promotes liver cancer stem cell growth through upregulating TERT and C-Myc. Oncotarget 2015; 6:40775–40798.

77. Schuijers J, Manteiga JC, Weintraub AS, Day DS, Zamudio AV, Hnisz D, Lee TI, Young RA. Transcriptional dysregulation of MYC reveals common enhancer-docking mechanism. Cell Rep 2018;23:349–360.

78. Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, Tanaka H, Taniguchi H, Kawakami Y, Ueno M, Gotoh K, Ariizumi S-I, Wardell CP, Hayami S, Nakamura T, Aikata H, Arihiro K, Boroevich KA, Abe T, Nakano K, Maejima K, Sasaki-Oku A, Ohsawa A, Shibuya T, Nakamura H, Hama N, Hosoda F, Arai Y, Ohashi S, Urushidate T, Nagae G, Yamamoto S, Ueda H, Tatsuno K, Ojima H, Hiraoka N, Okusaka T, Kubo M, Marubashi S, Yamada T, Hirano S, Yamamoto M, Ohdan H, Shimada K, Ishikawa O, Yamaue H, Chayama K, Miyano S, Aburatani H, Shibata T, Nakagawa H. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. Nat Genet 2016;48:500–509.

79. Umer HM, Cavalli M, Dabrowski MJ, Diamanti K, Kruczyk M, Pan G, Komorowski J, Wadelius C. A significant regulatory mutation burden at a high-affinity position of the CTCF motif in gastrointestinal cancers. Hum Mutat 2016;37:904–913.

80. Belton J-M, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. Hi-C: a comprehensive technique to capture the conformation of genomes. Methods 2012; 58:268–276.

81. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 2000;405:482–485.

82. Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, Lee K, Canfield T, Weaver M, Sandstrom R, Thurman RE, Kaul R, Myers RM, Stamatoyannopoulos JA. Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Res 2012;22:1680–1688.

83. Vető B, Bojcsuk D, Bacquet C, Kiss J, Sipeki S, Martin L, Buday L, Bálint BL, Arányi T. The transcriptional activity of hepatocyte nuclear factor 4 alpha is inhibited via phosphorylation by ERK1/2. PLoS One 2017;12: e0172020.

84. Simó R, Barbosa-Desongles A, Hernandez C, Selva DM. IL1β down-regulation of sex hormone-binding globulin production by decreasing HNF-4α via MEK-1/2 and JNK MAPK pathways. Mol Endocrinol 2012;26:1917–1927.

85. Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. Cell 2017;169:930–944.e22.

86. Wangensteen KJ, Wang YJ, Dou Z, Wang AW, Mosleh-Shirazi E, Horlbeck MA, Gilbert LA, Weissman JS, Berger SL, Kaestner KH. Combinatorial genetics in liver repopulation and carcinogenesis with a novel in vivo CRISPR activation platform. Hepatology 2018; 68:663–676.

87. Kieckhaefer JE, Maina F, Wells RG, Wangensteen KJ. Liver cancer gene discovery using gene targeting, Sleeping Beauty, and CRISPR/Cas9. Semin Liver Dis 2019;39:261–274.

88. Luo C, Keown CL, Kurihara L, Zhou J, He Y, Li J, Castanon R, Lucero J, Nery JR, Sandoval JP, Bui B, Sejnowski TJ, Harkins TT, Mukamel EA, Behrens MM, Ecker JR. Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. Science 2017;357:600–604.

89. Ackermann AM, Wang Z, Schug J, Naji A, Kaestner KH. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. Mol Metab 2016;5:233–244.

90. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 2011;17:10.

91. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–359.

92. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078–2079.

93. Ou J, Liu H, Yu J, Kelliher MA, Castilla LH, Lawson ND, Zhu LJ. ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. BMC Genomics 2018;19:169.

94. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9:R137.

95. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74.

96. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin S-F, Palmieri C, Caldas C, Carroll JS. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 2012;481:389–393.

97. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010; 26:139–140.

98. Yu G, Wang L-G, He Q-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 2015;31:2382–2383.

99. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 2010;38:576–589.

100. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.

101. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010; 26:841–842.

102. Schep AN, Buenrostro JD, Denny SK, Schwartz K, Sherlock G, Greenleaf WJ. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome Res 2015;25:1757–1770.

103. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Girón CG, Gil L, Gordon L, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek SH, Juettemann T, To JK, Laird MR, Lavidas I, Liu Z, Loveland JE, Maurel T, McLaren W, Moore B, Mudge J, Murphy DN, Newman V, Nuhn M, Ogeh D, Ong CK, Parker A, Patricio M, Riat HS, Schuilenburg H, Sheppard D, Sparrow H, Taylor K, Thormann A, Vullo A, Walts B, Zadissa A, Frankish A, Hunt SE, Kostadima M, Langridge N, Martin FJ, Muffato M, Perry E, Ruffier M, Staines DM, Trevanion SJ, Aken BL, Cunningham F, Yates A, Flicek P. Ensembl 2018. Nucleic Acids Res 2018;46:D754–D761.