Boston University Theses & Dissertations

Boston University Theses & Dissertations

2019

Chromatin accessibility and epigenetic changes induced by xenobiotic and hormone exposure in young adult mouse liver

https://hdl.handle.net/2144/39470 Downloaded from DSpace Repository, DSpace Institution's institutional repository

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

AND

COLLEGE OF ENGINEERING

Dissertation

CHROMATIN ACCESSIBILITY AND EPIGENETIC CHANGES INDUCED BY XENOBIOTIC AND HORMONE EXPOSURE IN YOUNG ADULT MOUSE

LIVER

by

ANDY RAMPERSAUD

B.S., St. Lawrence University, 2010 M.S., Boston University, 2012

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2019

© Copyright by ANDY RAMPERSAUD 2019

All rights reserved

Approved by

First Reader

David Waxman, Ph.D. Professor of Biology, Medicine, and Biomedical Engineering; Program in Bioinformatics

Second Reader

Trevor Siggers, Ph.D. Assistant Professor of Biology

DEDICATION

I would like to dedicate this work to my family members past and present. In memory of Mahadeo Rampersaud, my father, whose constant encouragement inspired me to pursue higher education. To my mother; Nalinee, my brother; Sean, sister-in-law; Lisa, and niece; Stacy, whose never ending support helped me finish this degree. Finally, to my wife Pinky and son Arnav, who remind me that the best part of my day is the time I spend with my family.

ACKNOWLEDGMENTS

I would like to thank my research advisor, Dr. David Waxman, for his guidance, patience, prompt feedback and the opportunity to work in the lab starting with my wetlab rotation during the summer of 2010. I appreciate all your efforts to mentor and help me mature as a scientist.

I would also like to thank past and present thesis advisory committee members for helpful discussions regarding my research: Dr. Cassandra Smith, Dr. Thomas Tullius, Dr. Mark Kon, and Dr. Frank Naya. I thank Dr. Trevor W. Siggers for helpful research discussions and feedback as my thesis second reader.

I would like to thank past and current members of the Waxman Lab for the many years of support and encouragement to succeed in graduate school. My first meeting as a lab member was in May 2011, I've had the pleasure of working with so many people during this time, too many to list here. I'd like to thank Dr. Aarathi Sugathan for the initial help with doing bioinformatics in a research lab setting. I particularly would like to thank Dr. George Steinhardt, who has been with me since the very beginning of the graduate program. George and I shared our first year cubicle space, took the same classes, we were members of the same Challenge project, ended up joining the same lab and sharing the same office space. I'm so thankful for the times we helped each other with getting through the ups and downs of graduate school and maintaining a work/life balance. I've learned so much from fellow bioinformatics students: Dr. Tisha Melia, Gracia Bonilla,

and Kritika Karri. A special "thank you" to Cindy Marmol for going above and beyond as our lab administrator. I can always rely on Cindy for helpful advice both in and out of graduate school. Lastly, I'd like to thank the many others in the lab and BU community with whom I had the pleasure to interact with regularly.

I'd like to acknowledge the Bioinformatics program administrators for helping me throughout this time: Caroline Lyman, Johanna Vasquez, and Dave King. I thank the bioinformatics Systems Administrator, Mary Ellen Gipson-FitzPatrick, for helping me learn how to use a computational server. I also thank the Research Computing staff for assisting me with using the shared computing cluster; an invaluable resource for doing bioinformatics.

To my wife, Pinky, thank you for your unwavering patience and all the sacrifices you have made to stay by my side. To my son, Arnav, you motivate me to finish this dissertation.

To my older brother; Sean, sister-in-law; Lisa, and niece; Stacy: thank you for all the support and encouragement. To my mom; Nalinee, I am so grateful for your love and support throughout these many years. To my beloved father, Mahadeo, thank you for providing me with the opportunity to pursue my education. I know that you would be happy to know that I finished my PhD.

CHROMATIN ACCESSIBILITY AND EPIGENETIC CHANGES INDUCED BY XENOBIOTIC AND HORMONE EXPOSURE IN YOUNG ADULT MOUSE LIVER

ANDY RAMPERSAUD

Boston University Graduate School of Arts and Sciences

and

College of Engineering, 2019

Major Professor: David Waxman, PhD. Professor of Biology and Medicine, and Biomedical Engineering; Program in Bioinformatics

ABSTRACT

Transcription factors activated by exogenous or endogenous stimuli alter gene expression with major effects on chromatin accessibility and the epigenome. This thesis investigates that impact of environmental chemical and hormonal exposure on liver chromatin accessibility in a mouse liver model. Exposure to the constitutive androstane receptor (CAR)-specific agonist ligand 1,4-bis-[2-(3,5-dichloropyridyloxy)]benzene (TCPOBOP) induces localized changes in chromatin accessibility at several thousand DNase hypersensitive sites (DHS). Activating histone marks, associated with enhancers and promoters, were induced by TCPOBOP and were highly enriched at opening DHS. Opening DHS were highly enriched for CAR binding sites and nuclear receptor direct repeat-4 motifs. These DHS were also enriched for the CAR heterodimeric partner RXRA, binding by CEBPA and CEBPB, and motifs for other liver-specific factors. Thus, TCPOBOP alters the enhancer landscape through changes in histone marks and by mechanisms linked to induced CAR binding. In other studies, the impact of pituitary growth hormone (GH) secretion patterns on chromatin accessibility changes associated with sex-biased liver gene expression was examined. In adult male liver, the transcription factor STAT5 is directly activated by each successive plasma GH pulse. In female liver, STAT5 is persistently activated by the near-continuous stimulation by plasma GH. A majority of the ~4,000 GH-regulated, sex-biased DHS have chromatin marks characteristic of enhancers and were enriched for proximity to sex-biased gene promoters. Chromatin accessibility is thus a key feature of sex-differential gene expression. Two major classes of male-biased DHS were identified: dynamic male-biased DHS, almost all bound by STAT5, which undergo repeated cycles of chromatin opening and closing induced by each GH pulse; and static male-biased DHS, whose accessibility is unaffected GH/STAT5 pulses and whose sex bias results from these chromatin sites being more closed in female liver. Sites with STAT5 binding showed greater chromatin opening, many of which also contain the STAT5 motif. Finally, the effect of a single GH pulse on hypophysectomized male mouse liver was investigated to identify DHS responsive to the male, pulsatile-GH, secretion pattern. These studies demonstrate that widespread epigenetic changes associated with target gene expression are induced by xenobiotics and hormones regulating liver gene expression.

TABLE OF CONTENTS

DEDICATION iv		
ACKNOWLEDGMENTS v		
ABSTRACTvii		
TABLE OF CONTENTSix		
LIST OF FIGURES xiv		
LIST OF ABBREVIATIONS xvii		
CHAPTER 1 - Introduction 1		
1.1 Abstract		
1.2 The impact of TCPOBOP exposure on the epigenetic landscape in mouse liver2		
1.3 Growth Hormone regulation of male and female chromatin accessibility in		
mouse liver		
1.3.1 GH activation of STAT5 and its role in sexually dimorphic gene expression 6		
1.4 Thesis goals and hypotheses		
CHAPTER 2 - Widespread epigenetic changes to the enhancer landscape of mouse liver		
induced by xenobiotic agonist ligand of nuclear receptor CAR 10		
2.1 Abstract		
2.2 Introduction		
2.3 Methods		
2.3.1 Animal studies		

	2.3.2	Cross linking and sonication of liver chromatin	. 16
	2.3.3	ChIP-seq laboratory methods	. 18
	2.3.4	Sequence analysis	20
	2.3.5	Chromatin mark ChIP-seq data analysis	. 21
	2.3.6	Liver DHS analysis	22
	2.3.7	Mapping of DHS and ChIP-seq peaks to TCPOBOP-responsive genes	23
	2.3.8	Chromatin mark responses at DHS	. 24
	2.3.9	DHS classification	25
	2.3.10	Enhancer class enrichment analysis	27
	2.3.11	DHS motif analysis	28
	2.3.12	Receiver Operating Characteristics (ROC) analysis	30
	2.3.13	CAR ChIP-seq data analysis	30
	2.3.14	CAR ChIP-seq enrichment analysis	. 33
	2.3.15	Enrichment of TCPOBOP-responsive genes for nearby CAR binding si	ites
		35	
	2.3.16	TCPOBOP-responsive genes dependent on MET/EGFR signaling	36
	2.3.17	RXR and CEBP ChIP-seq data analysis	37
	2.3.18	Peak normalization	38
2	2.4 R	esults	38
	2.4.1	ГСРОВОР-induced changes in chromatin accessibility may be transient,	
	sustair	ned or late	38
	2.4.2	ГСРОВОР-induced changes in active histone marks	40

2.4.3	3 Many ΔDHS are activated promoters and enhancers that map to TCPOBOP-	
respo	oonsive genes	41
2.4.4	4 Enrichment of nuclear receptor DR4 motif at ΔDHS	44
2.4.5	5 CAR binding at TCPOBOP-responsive DHS	46
2.4.6	6 CAR binding at TCPOBOP-responsive genes	48
2.4.7	7 DHS opening in the absence of CAR binding	49
2.4.8	8 TCPOBOP-induced binding of RXRA, CEBPA and CEBPB at TCPOBOP-	
respo	oonsive ΔDHS	51
2.4.9	9 Enriched motifs for associated co-factors	52
2.5	Discussion	54
2.6	Conclusions	62
CHAPTER 3 – Chromatin Accessibility Alterations Due to Plasma Growth Hormone		
Pulses in	Adult Male Mouse Liver 1	20
3.1	Abstract 1	20
3.2	Introduction1	21
3.3	Materials and Methods 1	25
3.3.1	1 Animal treatments and DNase-seq analysis 1	25
3.3.2	2 DNase-seq libraries, sequencing, and data analysis 1	26
3.3.2 3.3.3	 2 DNase-seq libraries, sequencing, and data analysis	26 27
3.3.2 3.3.3 3.3.4	 DNase-seq libraries, sequencing, and data analysis	.26 .27 .28
3.3.2 3.3.3 3.3.4 3.3.5	 DNase-seq libraries, sequencing, and data analysis	.26 .27 .28 .28

	3.3.7	Sex-specific genes and hypox response classification
	3.3.8	STAT5 binding and motif analysis 134
	3.3.9	Enrichment Analysis
	3.3.10	Chromatin state map analysis
	3.3.11	DNase-I cut site aggregate plots
	3.3.12	DHS peak normalization140
3	.4 R	esults
	3.4.1	Sex-biased DHS are enhancer regions that target sex-specific genes 141
	3.4.2	DHS responsive to endogenous pulses of GH and STAT5 in male liver 142
	3.4.3	Dynamic and static male-biased DHS 143
	3.4.4	Dynamic and static male-biased DHS and the presence of the STAT5 motif
		145
	3.4.5	Dynamic DHS are associated with stronger STAT5 binding and associated
	chroma	atin opening
	3.4.6	Chromatin accessibility changes following hypox and impact of a single
	replace	ement pulse of GH in male liver
	3.4.7	Many endogenous GH pulse-opened DHS are also opened by a single GH
	pulse i	n hypophysectomized male mouse liver
	3.4.8	A single GH pulse is sufficient for opening chromatin in hypophysectomized
	male n	nice
	3.4.9	A single GH pulse induces robust chromatin at dynamic DHS in
	hypopł	nysectomized male mouse liver 153

3.4.10	Chromatin state differences between static and dynamic male-biased DHS	
	155	
3.5 D	iscussion	
3.5.1	Sex-biased DHS are enhancer regions with distal gene targets 159	
3.5.2	Endogenous pulses of GH are responsible for chromatin accessibility	
change	es in male mouse liver	
3.5.3	Male-biased DHS can be classified into GH-pulse dynamic and static sites	
	160	
3.5.4	STAT5 binds directly to dynamic DHS 162	
3.5.5	Several thousand genomic regions respond to a single pulse of GH in	
hypopl	nysectomized male mouse liver 163	
3.5.6	A single GH pulse is sufficient for short-term restoration of chromatin	
accessi	bility in hypox male mice	
3.5.7	Chromatin state differences between dynamic and static male-biased DHS	
	167	
APPENDIX	X	
Appendix	x 1 – List of supplemental tables related to Chapter 2	
Appendix	x 2 – List of supplemental tables related to Chapter 3 225	
LIST OF JOURNAL ABBREVIATIONS		
REFERENC	CES	
CURRICUI	LUM VITAE	

LIST OF FIGURES

Fig.	2.1	TCPOBOP-stimulated chromatin opening and closing
Fig.	2.2	Enhancer and promoter DHS classes
Fig.	2.3	DR4 and half-site motifs in Δ DHS sequences
Fig.	2.4	Enrichment of transcription factor binding at TCPOBOP-responsive DHS
	and ge	nes
Fig.	2.5	ChIP-seq binding sites for mCAR, hCAR, RXR, and active histone marks at
	∆DHS	regions of TCPOBOP-inducible genes
Fig.	2.6	mCAR and hCAR binding sites and Δ DHS regions nearby TCPOBOP-
	inducil	ble genes involved in DNA synthesis and cell cycle regulation
Fig.	2.7	CAR binding to opening and to static DHS76
Fig.	2.8	Enrichment of de novo-discovered motifs at TCPOBOP-responsive ΔDHS .
		78
Fig.	S2.1	TCPOBOP opened DHS and induced genes in male and female liver 80
Fig.	S2.2	Comparisons of DHS width between subsets of opening Δ DHS regions 82
Fig.	S2.3	TCPOBOP exposure induces sex-differences in chromatin accessibility 84
Fig.	S2.4	Enrichment of ΔDHS vs. static DHS for containing activating chromatin
	marks	87
Fig.	S2.5	Identification and enrichment of enhancer and promoter DHS 89
Fig	S2.6	De novo and constructed motifs from TCPOBOP opened DHS sequences. 93
Fig.	S2.7	Receiver operating characteristic (ROC) curve analysis of constructed motifs
	in TCF	POBOP responsive DHS

Fig.	S2.8	Comparison of MACS14 score distributions for mouse (mCAR) and human
	(hCAR	CAR binding sites
Fig.	S2.9	Enrichment of ΔDHS vs. sDHS for containing bound CAR (analysis based
	on CA	R ChIP-exo data from Niu et al, 2018) 101
Fig.	S2.10	Enrichment for TCPOBOP-responsive genes (differentially expressed
	genes,	DEGs) for containing a CAR binding site (ChIP-exo) in the same TAD 105
Fig.	S2.11	21 de novo motifs, discovered from sets of ΔDHS that open, clustered by
	sequen	ce similarity
Fig.	S2.12	Enrichment of DHS that open for containing de-novo discovered motifs
		113
Fig.	S2.13	Enrichment of CAR bound opening and static DHS for containing de-novo
	discov	ered motifs 116
Fig.	discove 3.1	ered motifs
Fig. Fig.	discove 3.1 3.2	ered motifs
Fig. Fig. Fig.	discove 3.1 3.2 3.3	ered motifs
Fig. Fig. Fig. Fig.	discove 3.1 3.2 3.3 3.4	ered motifs
Fig. Fig. Fig. Fig. Fig.	discove 3.1 3.2 3.3 3.4 3.5	ered motifs
Fig. Fig. Fig. Fig. Fig.	discove 3.1 3.2 3.3 3.4 3.5 genes.	ered motifs
Fig. Fig. Fig. Fig. Fig.	discove 3.1 3.2 3.3 3.4 3.5 genes. 3.6	ered motifs
Fig. Fig. Fig. Fig. Fig.	discove 3.1 3.2 3.3 3.4 3.5 genes. 3.6 hypox	ered motifs
Fig. Fig. Fig. Fig. Fig.	discove 3.1 3.2 3.3 3.4 3.5 genes. 3.6 hypox 3.7	ered motifs

Fig. 3.9	Chromatin state analysis of static and dynamic male-biased DHS 187	
Fig. 3.10	Venn diagram and flow chart of the DHS that open due to an endogenous	
GH pu	ılse (2,373 sites) 192	
Fig. 3.11	Boxplot analysis and DNase-I cut site aggregate plots of the 7 DHS subsets	
	194	
Fig. S3.1	STAT5 ChIP-seq signal intensity distributions for the indicated sets of	
dynam	nic and static DHS 196	
Fig. S3.2	Discovered motifs from DHS sequences	
Fig. S3.3	Summary of sex-specific genes used in this study 204	
Fig. S3.4	Summary of DHS counts for the GH time course study	

LIST OF ABBREVIATIONS

°C	degrees Celsius
3-h or 27-h	
AR	Androgen receptor
AUC	Area under the curve
BED	Browser Extensible Data
bp	Base pairs
CAR	Constitutive androstane receptor
CCRP	CAR retention protein
CEBPA	CCAAT enhancer-binding protein alpha
CEBPB	CCAAT Enhancer Binding Protein Beta
cGH	Continuous growth hormone
ChIP	Chromatin immunoprecipitation
ChIP-exo	(ChIP) combined with lambda exonuclease digestion
CTCF	CCCTC-binding factor
CUX2	Homeobox protein cux-2
Сур2	Cytochrome P450, CYP2 family
Сур2b10	Cytochrome P450, family 2, subfamily b, polypeptide 10
CYP2B6	Cytochrome P450 family 2 subfamily B member 6
Сур2с39	Cytochrome P450, family 2, subfamily c, polypeptide 39
Сур2с55	Cytochrome P450, family 2, subfamily c, polypeptide 55
DAVIDDataba	se for annotation, visualization and integrated discovery

DHS	DNase hypersensitive sites
DMSO	Dimethyl sulfoxide
DR4 motifs	Direct-repeat 4 motifs
DREME	Discriminative Regular Expression Motif Elicitation
eDHS	Enhancer DHS
EGFR	Epidermal growth factor receptor
EMSA	Electrophoretic mobility shift assay
ER	Estrogen receptor
ER	Everted repeat
ES	Enrichment scores
FASTQ	file format for sequence read data
FDR	
FHx	Hypophysectomized female mice
FIMO	Find Individual Motif Occurrences
FOX	Forkhead box proteins
Foxa1/2	Forkhead box protein A1/2
GEO	Gene Expression Omnibus
GH	Growth hormone
H3K27ac	Lysine-27 acetylated histone-H3
H3K27me3	Lysine-27 trimethylated histone-H3
H3K4me3	Lysine-4 trimethylated histone-H3
hCAR	Human CAR

HCC	
HNF	Hepatocyte nuclear factors
HSP90	
Нурох	
i.p	Intraperitoneal injection
IR	Inverted repeat
Jak2	Janus kinase 2
JAK-STAT	Janus kinases-(STATs)
kb	Kilo base pair
KEGG	Kyoto Encyclopedia of Genes and Genomes
IncRNA	Long non-coding RNAs
MACS2	
mCAR	Mouse CAR
MEME	Multiple EM for Motif Elicitation
mg/kg	milligram per kilogram
MHx	Hypophysectomized male mice
MHx+GHH	Iypophysectomized male mice treated with growth hormone
mM	millimolar
Mup	Major urinary proteins
ng	nanogram
NLS	Nuclear localization signal
Nr1i3	Nuclear Receptor Subfamily 1 Group I Member 3

nt	nucleotide position
P450s	Cytochrome P450 enzymes
РВ	
PBREM	Phenobarbital (PB)-responsive enhancer module
PCA	
PP2a	
PPARA	Peroxisome proliferator-activated receptor-alpha; Ppara
PXR	
qPCR	Quantitative PCR
RACK	
RiPPM	reads-in-peaks-per-million
RNA-seq	
ROC	
RPKM	reads per kilobase per million
RXRA	
S	seconds
sDHS	Static DHS
SICER	Spatial Clustering for Identification of ChIP-Enriched Regions
STAT5	
SWI/SNF	SWItch/Sucrose Non-Fermentable is a nucleosome remodeling complex
TADs	
ТСРОВОР	

Thr-38	
Tomtom	Motif Comparison Tool
Tyr-694	Tyrosine residue 694
Tyr-699	Tyrosine residue 699
ΔDHS	Differential DNase-hypersensitive sites
μL	micro liter

CHAPTER 1 - Introduction

1.1 Abstract

Regulation of liver gene expression is important for detoxification, metabolism and maintaining homeostasis. Various stimuli, either from environmental chemicals or endogenous circulating hormones, can influence gene regulation in the liver. Environmental chemicals pose a health risk to humans and wildlife through their multitude of exposure routes and subsequent alterations of gene expression. Their mechanism of action following exposure is poorly understood. One hypothesis is that these chemicals directly bind to nuclear receptors, primary found in the liver, to exert their effects on gene transcription. Constitutive androstane receptor (CAR), studied here, is a member of the nuclear receptor superfamily, which contains several other receptors including those that function as xenobiotic and steroid hormone receptors. Aside from exogenous stimuli, growth hormone (GH) is a major regulator of sex-specific gene expression in mouse liver. GH secretion profiles are regulated by the pituitary gland and show fundamental differences between male and female liver. Signal transducer and activator of transcription (STAT)5, also studied here, is activated by GH and plays a key role in maintaining and establishing sexual dimorphic gene expression. Little is known about how sex-specific GH secretion establishes and maintains sex-specific chromatin accessibility. One hypothesis is that GH activated STAT5 directly binds to genomic regions responsible for regulating sex-specific gene expression. In my thesis work, I have explored the impact of environmental chemical exposure and levels of endogenous

hormone on mouse liver chromatin accessibility and the corresponding changes in the epigenetic landscape.

1.2 The impact of TCPOBOP exposure on the epigenetic landscape in mouse liver

Humans are exposed to thousands of environmental chemicals found in the soil, water, air or the manufacturing process of consumer products. It is important to understand how these foreign chemicals induce changes in liver gene expression and corresponding changes in chromatin accessibility and the epigenome. We have chosen to study the constitutive androstane receptor (CAR) and its potent, receptor-specific ligand 1,4-Bis[2-(3,5-Dichloropyridyloxy)]Benzene (TCPOBOP). CAR, once known to be an orphan receptor with no known endogenous ligands or function, is now known to directly bind several steroid hormones (Hernandez et al., 2009) and regulate bile acid (Lickteig et al., 2016), glucose and lipid homeostasis (Wada et al., 2009)in the liver. The main function of CAR is to facilitate the clearance of xenobiotics through direct binding with the chemical ligand (Suino et al., 2004) or indirect activation through the epidermal growth factor receptor (EGFR) signaling (Kobayashi et al., 2015), and induction of genes involved in drug metabolism. Although CAR activation is associated with positive physiological benefits (i.e. xenobiotic clearance), there are several negative effects linked to prolong CAR activation. For instance, CAR activation has been linked to liver steatosis (Mellor et al., 2016) and tumor promotion in mouse liver (Yamamoto et al., 2004).

One of the defining characteristics of CAR among all other nuclear receptors is its constitutive activity in the absence of ligand binding. In the absence of direct ligand binding, CAR can be indirectly activated. CAR activation by phenobarbital (PB) is achieved this indirect mechanism mediated by the EGFR (Kobayashi et al., 2015). To prevent persistent and aberrant CAR activity due to its constitutive function, the liver cell sequesters CAR in the cytoplasm by complexing CAR with heat shock protein 90 (HSP90), the cytoplasmic CAR retention protein (CCRP) (Timsit and Negishi, 2014) and by phosphorylating CAR at Thr-38 (Osabe and Negishi, 2011). Direct ligand binding causes conformational changes that free CAR from the protein complex with HSP90 and CCRP, leads to dephosphorylation of Thr-38 by the phosphatase PP2a, and exposes a nuclear localization signal (NLS) (Kanno et al., 2005). Dephosphorylation of CAR at Thr-38 is also a critical step in the indirect activation of CAR with PB; however, this is dephosphorylation is achieved through RACK and MEK/ERK signaling (Yang and Wang, 2014). Once activated by exogenous xenobiotics or endogenous steroid hormones in the cytoplasm, CAR translocates to the nucleus where it binds to and regulates transcription of genes involved in drug metabolism and excretion.

Cytochrome P450 genes are well known protein coding targets of CAR regulation in the liver. The effects of PB and TCPOBOP on the liver showed that cytochrome P450s are highly induced following drug treatment (Poland A, 1980), and eventually led to the discovery of CAR as a receptor and transcription factor. Since then, the *Cyp2* family of genes has been extensively studied for their robust response to CAR activation in mouse and human liver; specifically *Cyp2b10* in mice and its human orthologue *CYP2B6*

(Kobayashi et al., 2015). Increasing evidence also suggests that CAR regulates its gene targets by directly binding to response elements upstream of mouse *Cyp2b10* or human *CY2B6* (Tian et al., 2011). CAR activation of *Cyp2b10* in mice is regulated by a 51 bp sequence, conserved in mouse, rat, and human, that contains two nuclear receptor binding sites (defined by DR4 motifs) and was later named a phenobarbital responsive enhancer module (PBREM) (Honkakoski et al., 1998). Additional CAR inducible *Cyp2* genes include *Cyp2c39* and *Cyp2c55* (Jackson et al., 2006), thus highlighting the need to identify activated CAR binding sites genome-wide.

Many groups have attempted to characterize the genome-wide transcriptional changes in the liver by CAR activation. We successfully used RNA-seq to characterize global changes in the mouse liver transcriptome following exposure to the CAR-specific agonist ligand TCPOBOP. Analysis of dysregulated protein coding genes showed strong enrichment for KEGG pathways of xenobiotic and drug metabolism including upstream regulators associated with cell cycle dysregulation and hepatocellular carcinoma progression (Lodato et al., 2017). Although CAR activation is associated with dysregulation of protein coding and long non-coding RNA (lncRNA) genes, the chromatin accessibility and epigenetic changes that regulate CAR-induced gene expression were unknown. We investigated the distribution of TCPOBOP-responsive genes in topologically associating domains (TADs) and show that subsets of genes cluster within TADs. Further, TCPOBOP induces localized changes in mouse liver chromatin accessibility, many of which cluster in TADs together with TCPOBOP-responsive genes. Sites of chromatin opening were highly enriched for proximity to induced genes and chromatin closing was enriched nearby repressed genes due to TCPOBOP treatment (Lodato et al., 2018). Together, these findings indicate that TCPOBOP-activated CAR induces genomic regions, likely enhancer and promoter regulatory elements, to positively regulate xenobiotic and drug-metabolizing genes.

1.3 Growth Hormone regulation of male and female chromatin accessibility in mouse liver

Sexual dimorphic gene expression affects > 1,000 genes and is a consequence of distinct GH secretion patterns from the pituitary gland (Wiwi and Waxman, 2004). Many of the genes that are differentially expressed include those involved in xenobiotic and drug metabolism, such as cytochrome P450 genes (O'Connor and Waxman, 2006). These sex differences, particularly for drug metabolism genes, can have a significant impact on drug effectiveness and susceptibility to their side effects in men and women (Schwartz, 2007). Sexual dimorphic gene expression in the liver is associated with lipid metabolism and cardiovascular disease risk, and discrete genomic loci have implicated with polygenic dyslipidemia and coronary artery disease using a large collection of liver samples (Zhang et al., 2011). Sex differences in gene expression can affect susceptibility to other diseases: hepatic fibrosis, cirrhosis, and hepatocellular carcinoma (HCC) (Ruggieri et al., 2010). Sexual dimorphic levels of hormone receptors have been linked to disease development: estrogen receptor (ER) confers resistance to HCC in females (Naugler et al., 2007) and and rogen receptor (AR) confers susceptibility to HCC in males (Wu et al., 2010). Further, it has been shown that estrogen-dependent resistance to and androgenmediated facilitation of HCC depend on transcription factors Foxa1/2 (Li et al., 2012), which also function as important pioneer factors in modulation chromatin accessibility.

1.3.1 GH activation of STAT5 and its role in sexually dimorphic gene expression

Sexually dimorphic gene expression in rodents is controlled by the sex-specific pattern of release of GH from the pituitary gland. In rats, the male GH profile is regular and pulsatile, with sharp peaks of GH released at ~3.5 hour intervals with intervening periods of low or absent circulating GH. In female rats, GH release is more frequent, resulting in a near continuous presence of GH in circulation with the absence of GH-free intervals (Jansson et al., 1985). To examine the effects of GH regulation, hypophysectomy, the surgical removal of the pituitary gland, has been utilized to assess changes due to loss of GH along with other pituitary hormones (Wauthier and Waxman, 2008).

Hypophysectomy (hypox) has widespread effects on the liver transcriptome, in particular, abolishing sex-differences in gene expression. Microarray analysis of genome-wide gene expression showed that 90-94% of liver sex-biased gene expression in rats (Wauthier and Waxman, 2008) and mice (Wauthier et al., 2010) is lost following hypophysectomy. These gene expression changes could be the result of GH ablation by hypophysectomy but could also be due to the loss of other pituitary-dependent hormones. Studies of GH replacement, either by pulsatile hormone administration (Wauthier et al., 2010; Wauthier and Waxman, 2008) or by continuous hormone infusion (Holloway et al., 2006), have shown that the majority of sex differences in mouse and rat liver gene expression are in fact regulated by the sex-dependent plasma GH patterns.

The male plasma profile in mice is defined by regular pulses spaced by ~2.5-hour interpulse intervals, while female mice exhibit near-continuous low levels of GH established by more frequent (~1.4-hour intervals compared to males) pules of GH (MacLeod et al., 1991). Together, the lack of sustained GH-free intervals distinguishes the female from the male GH profile. In fact, pituitary-intact male mice treated with a continuous infusion of GH, delivered by an osmotic pump, show a dramatic feminization of liver gene expression; a large fraction of male-biased genes are down-regulated and female-biased genes are up-regulated (Holloway et al., 2006). Furthermore, there are corresponding changes in sex-specific chromatin accessibility and histone modifications associated with sex-biased epigenetic states (Lau-Corona et al., 2017).

Signal transducer and activator of transcription-5 (STAT5) is the main effector of transcriptional responses to GH in the liver. GH binding to the GH receptor activates STAT5 by tyrosine phosphorylation and activation of Janus kinase 2 (Jak2), a tyrosine kinase that phosphorylates STAT5 proteins: Tyr-699 of STAT5b, and Tyr-694 of STAT5a, which are two closely related (>90% sequence conserved) STAT5 genes. STAT5b is the major liver-expressed STAT5 form, and both forms are collectively referred to here as STAT5. Phosphorylated STAT5 dimerizes and translocates to the nucleus, where it binds to regulatory elements either indirectly or directly (recognition sequence: TTC-NNN-GAA) to regulate transcription of its target genes (Herrington et al., 2000). Given the cascade of signaling events described above, liver STAT5 activity levels mirror plasma levels of GH in individual male rats (Choi and Waxman, 2000), suggesting that STAT5's transcriptional regulatory function dynamically follows plasma GH levels. ChIP-seq analysis of genome-wide binding revealed a close correlation between sex-dependent STAT5 binding and sex-biased target gene expression. GH pulse activated STAT5 is strongly enriched for binding nearby male-biased genes in male mouse liver and for binding near-continuous GH-activated STAT5 nearby female-biased genes in female mouse liver (Zhang et al., 2012).

Furthermore, mouse knockout models show that the sex-dependent expression of > 1,000 liver expressed genes requires STAT5. In male *Stat5b* knockout mice, 90% of malebiased genes are repressed and 60% of female-biased genes are induced (or derepressed) (Clodfelter et al., 2007). These studies support the hypothesis that dynamic activity patterns of STAT5 lead to sexually dimorphic gene expression in male and female mouse liver.

1.4 Thesis goals and hypotheses

This thesis contains two major goals and hypotheses that are novel and will contribute to field of endocrine regulation of gene expression and epigenetics. In chapter 2, I build upon the gene expression and differential DNase-hypersensitive sites (Δ DHS) analysis to further classify these sites as active enhancer and promoter regions. I integrate chromatin accessibility, sequence analysis of these sites, histone modifications, and transcription factor binding datasets to better understand the actions of CAR through studies of TCPOBOP-responsive gene expression changes in mouse liver. Further, I show that CAR binding at these regulatory regions is associated with TCPOBOP-inducible gene expression. In chapter 3, I answer conceptually similar questions, but in the context of

sex-specific gene expression regulated by endogenous GH. To assess the role of GH in regulating gene expression and chromatin accessibility, I analyze datasets from intact male, hypox male and female, and hypox male mice given a single injection of GH. I examine the effect of hypophysectomy on chromatin accessibility in male and female mouse liver and show that a single pulse of GH is sufficient for temporary restoration of chromatin openness to levels seen in intact male mouse liver. Further, I show that STAT5 binding at GH-responsive DHS is associated with higher levels of open chromatin, which suggests that STAT5-mediate mechanisms for chromatin remodeling are an important aspect of sexually dimorphic gene expression.

CHAPTER 2 - Widespread epigenetic changes to the enhancer landscape of mouse liver induced by xenobiotic agonist ligand of nuclear receptor CAR

2.1 Abstract

Background: Many environmental chemicals induce hepatic drug, steroid and lipid metabolism and dysregulate genes linked to hepatocellular carcinogenesis by interacting with nuclear receptors; however, the epigenetic effects of such exposures and their relationship to gene dysregulation are poorly understood. Here, we used the model environmental chemical TCPOBOP, a highly specific halogenated agonist ligand of the nuclear receptor and xenobiotic sensor CAR (Nr1i3), to investigate early effects of xenobiotic exposure on the epigenome and chromatin states in mouse liver. Results: Global epigenetic maps were obtained for activating and repressive histone marks, and TCPOBOP-responsive liver transcription factor binding sites were identified and used to characterize changes to the liver epigenome and chromatin states following TCPOBOP exposure. TCPOBOP stimulated time-dependent chromatin opening or closing at a few thousand enhancer-marked genomic regions enriched at TCPOBOP-responsive genes and identified by differential DNase-hypersensitive site (DHS) analysis. DHS opened by TCPOBOP were significantly enriched for induced enhancer marks, CAR binding and CAR DR4 motifs, indicating CAR is likely to bind directly to these TCPOBOP-activated enhancers. CAR binding was also enriched at many constitutively open DHS nearby the inducible enhancer DHS. DNA replication/hepatocyte proliferation genes co-dependent on MET and EGF receptor signaling for induction by TCPOBOP were also enriched for

CAR binding, suggesting these genes are also induced by a direct binding mechanism. In contrast, a subset of opening DHS and many closing DHS mapping to TCPOBOP-responsive genes did not bind CAR, indicating an indirect mechanism for those chromatin accessibility changes. TCPOBOP-responsive DHS were also enriched for induced binding of RXRA, CEBPA and CEBPB, and for motifs for other liver factors that may contribute to liver-specific transcriptional responses to TCPOBOP exposure. Conclusions: These studies elucidate the impact of foreign chemical exposure on the enhancer landscape of mouse liver, and show that widespread epigenetic changes are induced by both direct and indirect mechanisms linked to CAR activation. The global maps of thousands of environmental chemical-induced epigenetic changes described here constitute a rich resource for further research on xenochemical effects on liver chromatin states and the epigenome.

2.2 Introduction

Nuclear receptor superfamily members CAR (constitutive androstane receptor; *Nr1i3*), PXR (pregnane X receptor, *Nr1i2*), and PPARA (peroxisome proliferator-activated receptor-alpha; *Ppara*) are hepatic sensors that facilitate metabolism of steroids, lipids and drugs and enable the elimination of foreign chemicals in mammalian systems (Dixit et al., 2005; Timsit and Negishi, 2007; Waxman, 1999). Persistent activation of CAR (Abe et al., 2018; Dong et al., 2015) and PPARA (Brocker et al., 2017) stimulates hepatocyte proliferation and induces liver growth and tumorigenesis in mice and rats (Lake, 2018; Misra et al., 2013). Diseases such as diabetes, inflammatory disease and liver diseases are also associated with activation of these receptors (Banerjee et al., 2015). All three nuclear receptors are ligand-activated transcription factors that bind similar DNA motifs and activate partially overlapping sets of genes (Aleksunes and Klaassen, 2012; Cui and Klaassen, 2016; Lodato et al., 2017; Ross et al., 2009), including long non-coding RNA genes (Dempsey and Cui, 2018; Lodato et al., 2017).

CAR is activated by structurally diverse environmental chemicals, including many consumer products, pharmaceuticals and industrial chemicals (Ashrap et al., 2017; Baldwin and Roling, 2009; Chang and Waxman, 2006; DeKeyser et al., 2011; Ito et al., 2012; Laurenzana et al., 2016; Oshida et al., 2015; Ren et al., 2010; Rooney et al., 2018). CAR is activated by dephosphorylation of Thr-38, which can be induced by direct binding of xenochemicals, such as the halogenated agonist ligand TCPOBOP (1,4-bis-[2-(3,5dichloropyridyloxy)]benzene), or by indirect activators, such as phenobarbital, which stimulates CAR Thr-38 dephosphorylation by disruption of signaling downstream of EGF receptor (EGFR) (Mutoh et al., 2013; Osabe and Negishi, 2011). Thr-38 dephosphorylation dissociates CAR homodimers and induces translocation of CAR to the nucleus, where CAR binds the nuclear receptor retinoid X receptor (RXR). The resultant CAR-RXR heterodimer binds DNA with a preference for direct repeats of AG(G/T)TCA hexamers spaced by 4 nucleotides, i.e., a DR4 motif (Frank et al., 2003; Kim et al., 2001), and activates target gene transcription (Negishi, 2017). Many activators of CAR can also activate PXR and/or PPARA (Al-Salman and Plant, 2012; Hurst and Waxman, 2004; Masuyama et al., 2000; Omiecinski et al., 2011; Ren et al., 2010; Tabb et al., 2004), which presents a major challenge when interpreting the complex gene expression changes that

these chemicals induce. This promiscuity of ligand binding is in part due to the large number of ligand binding site arrangements seen by X-ray (Buchman et al., 2018; Wallace and Redinbo, 2013; Wu et al., 2013) and solvent mapping analysis (Ngan et al., 2009).

Many environmental chemicals induce changes in histone modifications and DNA methylation (Cortessis et al., 2012; Eckstein et al., 2017; Helsley and Zhou, 2017; Ruiz-Hernandez et al., 2015; Strazzullo and Matarazzo, 2017). These changes may be preserved in epigenetic memory (Jimenez-Chillaron et al., 2015) and enable the recruitment of factors that control DNA compaction and chromatin accessibility, impacting transcription factor binding and transcriptional activation. Changes in DNA methylation at regulatory regions may also occur, but are often secondary to changes in chromatin accessibility (Stadler et al., 2011; Thurman et al., 2012). Environmental chemical exposures can induce both short-term and long-term (mitotically stable) (Reveron-Gomez et al., 2018) epigenetic changes, as well as trans-generational effects via the germ line (Guerrero-Bosagna and Skinner, 2012; Latchney et al., 2018; Nilsson et al., 2018; Xin et al., 2015). These epigenetic changes can have widespread physiological and pathophysiological consequences that impact development, aging and disease susceptibility (Christensen and Marsit, 2011; Edwards and Myers, 2007; Heindel et al., 2006; Jin et al., 2010; Nagy and Turecki, 2012). Relatively little is known about the mechanisms whereby exposure to foreign chemicals induce such changes to the epigenome (Parfett and Desaulniers, 2017; Suvorov and Waxman, 2015; Walker, 2016).

Changes in local chromatin accessibility, a key point of epigenetic regulation, can be identified by DNase hypersensitivity analysis, which we used to discover ~60,000-70,000 DNase hypersensitive sites (DHS) in mouse liver (Ling et al., 2010; Lodato et al., 2018). DHS include promoters, enhancers and insulators (Shu et al., 2011) and encompass ~90% of all binding sites for many liver transcription factors (Ling et al., 2010; Thurman et al., 2012). Chromatin states can be learned by combinatorial analysis of DHS with other epigenetic marks to identify functional genomic features, including active, poised and repressed promoters and enhancers (Ernst and Kellis, 2017; Libbrecht et al., 2015). Thus, lysine-27 acetylated histone-H3 (H3K27ac) marks active enhancer DHS (Creyghton et al., 2010), while lysine-4 trimethylated histone-H3 (H3K4me3) in combination with H3K27ac marks active promoter DHS (Shlyueva et al., 2014). In a study of mouse liver chromatin states (Sugathan and Waxman, 2013), we found that a subset of target genes of CAR, and of PXR, are in a poised state prior to nuclear receptor activation, as indicated by the combination of H3K27me3 (repressive) marks with H3K4me1 (activating) marks (Creyghton et al., 2010; Zentner et al., 2011). Furthermore, CAR target genes switch from a poised to an active state after persistent activation of CAR for several weeks (Chen et al., 2012; Lempiainen et al., 2011; Thomson et al., 2012). This epigenetic switch may predispose hepatocytes to the cancer-causing mutations associated with exposure to TCPOBOP and other non-genotoxic carcinogens (Dong et al., 2015; Phillips et al., 2009; Thomson et al., 2013a; Thomson et al., 2013b). However, comparatively little is known about the early epigenetic changes induced by CAR activation. One early response, for both CAR and PXR, is the rapid (within 3-h) induction of localized changes in chromatin
accessibility at several hundred genomic sites (Δ DHS regions), some proximal to CAR and PXR target genes (Lodato et al., 2018). By 27-h, TCPOBOP induces chromatin opening at ~2,000 sites and chromatin closing at ~600 sites in male mouse liver (Lodato et al., 2018). It is not known whether these early chromatin opening and chromatin closing events involve binding of CAR to liver chromatin (Niu et al., 2018; Tian et al., 2018), or how histone marks and chromatin states associated with these Δ DHS may change following TCPOBOP exposure and CAR binding.

Here, we examine early changes in active and repressive chromatin marks, and in transcription factor binding, in TCPOBOP-exposed mouse liver and their relationships to TCPOBOP-stimulated changes in chromatin accessibility, CAR binding and gene expression. Our findings identify distal opening DHS that comprise the CAR-activated enhancer landscape of TCPOBOP-exposed liver, as indicated by their strong enrichment for enhancer marks, CAR binding, DR4 and nuclear receptor (NR) half-site motifs, and proximity to TCPOBOP-responsive genes. CAR binding was also seen at many enhancer and promoter DHS that are constitutively open in mouse liver. We also found that many closing DHS, and a subset of opening DHS, do not bind CAR but map to CAR target genes and are enriched for changes in enhancer marks, indicating that TCPOBOP activation of CAR can indirectly induce epigenetic changes that are functionally linked to changes in gene expression. These global analyses give novel insights into the changes to the liver epigenome that are linked to dysregulation of gene expression in TCPOBOP exposed liver.

2.3 Methods

2.3.1 Animal studies

All animal work was performed in compliance with procedures approved by the Boston University Institutional Animal Care and Use Committee. Exposure of 7-week old male and female CD-1 mice (ICR strain) to TCPOBOP (3 mg/kg body weight, single i.p. injection in 1% DMSO in corn oil, or with vehicle control, tissue collection either 3-h or 27-h later, isolation of liver nuclei, DNase-I treatment to release DNA fragments from DHS regions followed by high throughput sequence analysis to identify liver DNase hypersensitive sites (DHS) were carried out as described elsewhere (Lodato et al., 2018). Liver nuclei and frozen liver tissue isolated from control and either 3-h or 27-h TCPOBOP-treated male mice were also used for ChIP-seq analysis of chromatin marks (histone marks) and transcription factor binding, as described below.

2.3.2 Cross linking and sonication of liver chromatin

Chromatin immunoprecipitation (ChIP) analysis of histone marks and transcription factor binding sites was carried out using sonicated liver chromatin, prepared in one of two ways: Method A, sonication of cross-linked liver nuclei; or Method B, sonication of whole liver tissue that was initially snap-frozen in liquid nitrogen and then cross-linked with formaldehyde. For <u>Method A</u>, to cross-link nuclei, ~100 million nuclei freshly isolated from male mouse livers were resuspended in 1 mL of cross-linking buffer (10 mM HEPES [pH 7.6], 25 mM KCl, 0.34 M sucrose, 2 mM MgCl₂, 0.15 mM 2mercaptoethanol) preheated to 30°C. Formaldehyde (37% stock solution) was added to give a final concentration of 0.8% (v/v), followed by incubation at 30°C for 9 min with periodic mixing to cross-link the nuclei. Glycine was then added to a final concentration of 0.1 M to quench the cross-linking reaction at room temperature. Cross-linked nuclei were layered on 3 mL of fresh homogenization buffer and centrifuged at 4°C for 30 min at 25,000 rpm. Cross-linked nuclear pellets were resuspended in RIPA buffer (50 mM Tris-HCl pH 8.1, 150 mM NaCl, 1% (v/v) NP-40, 0.5% (v/v) deoxycholic acid, Na salt; 0.1% (v/v) sodium dodecyl sulfate, and one Roche cOmplete Protease Inhibitor CocktailTM tablet per 50 mL (Sigma, catalog number 11697498001). Nuclei were then sonicated in a Bioruptor Twin sonicator on the high setting until chromatin was ~100-300 bp in length (80 to 100 cycles; 1 cycle = 30 s on and 30 s off). Sonicated material from vehicle and 3-h TCPOBOP-treated livers was then used for ChIP analysis.

For <u>Method B</u>, to cross-link whole liver tissue, a third of each liver (previously snapfrozen in liquid nitrogen and stored at -80°C) was retrieved and placed on ice for 5 min, followed by tissue disruption in a glass Dounce homogenizer with 4 mL of cross-linking buffer (50 mM HEPES [pH 7.5], 100 mM NaCl, 1 mM EDTA, and 0.5 mM EGTA) containing protease inhibitors (Thermo Scientific Pierce Protease Inhibitor Tablets, EDTA-free). The tissue homogenate was passed through a 70-micron cell strainer in a petri dish. Additional cross-linking buffer was added to the homogenizer to rinse out the remaining tissue and passed through the cell strainer. Formaldehyde was added to the homogenate to a final concentration of 1% and mixed rapidly in a conical tube. The mixture was rocked for 10 min at 22°C, followed by the addition of glycine to a final concentration of 0.125 M. The sample was rocked for 2 min at 22°C and then centrifuged at 4°C for 5 min at 2,500g. The pellet was washed twice in 10 mL of PBS, resuspended in 10 mL of lysis buffer 1 (50 mM HEPES [pH 7.5], 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% IGEPAL CA-630, 0.25% Triton X-100, and protease inhibitors) and rocked for 10 min at 4°C. The sample was then centrifuged at 4°C for 5 min at 2,000g, and the resultant pellet was resuspended in 10 mL of lysis buffer 2 (200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM Tris-HCl [pH 8.0], and protease inhibitors) and rocked for 5 min at 4°C. After another centrifugation, the pellet was resuspended in RIPA buffer to a final volume of 2 mL. The sample was transferred to a 15-mL TPX tube (Diagenode) containing 0.3 mL polypropylene beads (Diagenode), and sonicated for 35 cycles at 4°C (1 cycle = 30 s on and 30 s off) using a Bioruptor Pico sonicator (Diagenode). Sonicated material from vehicle-treated and 27-h TCPOBOP-treated livers was then used for ChIP analysis.

2.3.3 ChIP-seq laboratory methods

Dynabeads Protein A (Invitrogen #1002D) were used for all ChIP reactions for the 3-h TCPOBOP exposure group and Protein A/G Magnetic Beads ("Bimake beads", #B23030) were used for the 27-h exposure group. For ChIP analysis of H3K4me1, H3K4me3, and H3K27ac marks, 15 μ L of Dynabeads or 10 μ L of Bimake beads were used. For H3K27me3 marks, 10 μ L of Dynabeads or 5 μ L of Bimake beads were used. Beads were washed three times with 1 mL of blocking solution (0.5% (w/v) BSA in

PBS). 300 µL of fresh blocking solution was then added to each tube containing beads, followed by histone H3 antibody specific for K4me1 (1.2 µg, Abcam ab8895), K4me3 (3 μg, Abcam ab8580), K27ac (1.2 μg, Abcam ab4729), or K27me3 (2 μg, Abcam ab6002). Beads and antibody in blocking solution were incubated for 3-h at 4°C with rocking. Beads with bound antibody were washed three times with blocking solution and the following amounts of sonicated chromatin were then added: 10 μ g for H3K27me3; 15 μ g for H3K4me1, H3K4me3, and H3K27ac. ChIP reactions were incubated overnight at 4°C with rocking. The next day, ChIP samples were centrifuged at 2,000 rpm for 2 min at 4°C and washed with RIPA buffer three times. Next, ChIP samples were washed three times with RIPA buffer containing 0.5 M NaCl, followed by a single wash with TE buffer. After the final wash, 50 µL elution buffer (50 mM Tris-HCl [pH 8.1], 10 mM EDTA, 1% (v/v) SDS) was added and samples were incubated in a 65°C water bath for 35 min with periodic mixing. The supernatant containing eluted chromatin was adjusted to an NaCl concentration to 0.2 M and then incubated at 65°C to reverse the cross-links. RNase A (ThermoFisher E0531, 10 mg/mL) was then added to a final concentration of 0.12 mg/mL and the samples were incubated at 37°C for 30 min. Proteinase K (Bioline BIO-37084, 20 mg/mL) was added to a final concentration of 0.39 mg/mL and samples were incubated at 37°C for 2 h. Finally, DNA was purified using QIAprep 2.0 spin columns (Qiagen) according to the manufacturer's manual and eluted in a final volume of 50 μ L TE buffer.

2.3.4 Sequence analysis

ChIP-seq was carried out on three biological replicates (chromatin prepared from three individual male mouse livers) for each of four sample groups (livers of mice exposed to TCPOBOP for 3-h or for 27-h, and time-matched vehicle controls). Sequencing libraries were prepared from 5 ng of ChIP DNA material using the NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs). Sequencing was performed at the New York Genome Center (New York, New York) on an Illumina HiSeq instrument. Paired end sequence reads, 50 bp in length, were obtained for the ChIP-seq samples. A custom pipeline was established to process raw ChIP-seq sequencing files for histone marks and transcription factor binding site data sets. The pipeline inputs raw FASTQ files and outputs various quality control metrics, including FASTQC reports (FASTX-Toolkit v0.0.13.2), confirmation of read length, verification of the absence of read strand bias, quantification of contaminating adapter sequence (Trim galore v0.4.2), and distributions of insert-size lengths (for paired-end read sequencing), as calculated and visualized using Picard (v1.123). Reads were mapped to the mouse genome (release mm9) using Bowtie2 (v2.2.6) (Langmead et al., 2009). All discovered ChIP-seq peak sets were filtered to remove peaks that overlap ENCODE blacklisted regions (2012), as well as peaks comprised of >4 identical reads that did not overlap any other read ("straight peaks"). Raw and processed ChIP-seq data for control and TCPOBOP-treated liver samples are available at www.ncbi.nlm.nih.gov/gds under the following accession numbers: **GSE104060**, for activating chromatin marks H3K4me1, H3K4me3, and H3K27ac; GSE121921, for H3K27me3 marks; and GSE121915, for transcription factors RXRA,

CEBPA, and CEBPB. Sequencing statistics for each sample are summarized in Table_S1A. All genomic coordinates used in this study are based on mouse genome assembly mm9.

2.3.5 Chromatin mark ChIP-seq data analysis

ChIP-seq peaks were identified for the activating histone-H3 marks K4me1, K4me3 and K27ac using MACS2 (v2.1.0.20150731) with the option (--keep-dup); all other parameters were set to the default option. ChIP-seq peaks were identified for the repressive mark H3K27me3 using SICER (v1.1) (Xu et al., 2014) with the following parameters: fragment size 200 bp, window size 400 bp, gap size 2,400 bp and E-value threshold 100 (Sugathan and Waxman, 2013). MACS2 peaks discovered for H3K27ac, H3K4me1 and H3K4me3, and SICER regions identified for H3K27me3, were processed separately for the 3-h and for the 27-h TCPOBOP treatment and corresponding vehicle control datasets, using the methods described for H3K27ac peaks, as follows: First, a peak union list was generated by merging the peak calls from individual biological replicates for the corresponding treatment group. For example, a single peak union list for 3-h H3K27ac peaks was generated by merging all MACS2 peaks from each of the n = 6male mouse liver ChIP-seq samples (n = 3 vehicle control and n = 3 TCPOBOP exposures livers). Genomic regions that showed significantly differential ChIP-seq signal between TCPOBOP-exposed and vehicle control samples were discovered separately for each time point using diffReps (Shen et al., 2013), using the nucleosome option (200 bp window) and thresholds of TCPOBOP vs. vehicle control, with significance based on a

[fold-change] > 2 and FDR < 0.05 (Benjamini-Hochberg adjusted p-value) for diffReps normalized signal intensity values. BEDTools was used to determine the overlap between the H3K27ac peak union list for all six samples and the set of significantly differential H3K27ac mark regions identified by diffReps for TCPOBOP treatment vs. control liver (1,353 induced H3K27ac sites and 670 repressed H3K27ac sites at the 3-h time point). H3K27ac MACS2 peak union sites that did not overlap a diffReps region were annotated as static H3K27ac peaks (66,585 sites). Concatenation of the static H3K27ac peak regions with the diffReps-identified regions yielded the full set of 68,608 H3K27ac peaks used for downstream analysis. Corresponding analyses were performed for H3K4me1 and H3K4me3 MACS2 peaks and for H3K27me3 SICER regions. Table_S2 presents a summary of the numbers of TCPOBOP-responsive histone mark peaks, along with full listings of all chromatin mark peak regions with normalized read counts for each control and TCPOBOP treatment time point. Table_S3 lists the corresponding sets of normalized read counts that map to the set of 60,739 liver DHS regions described below.

2.3.6 Liver DHS analysis

A set of 60,739 merged DHS regions was obtained from the union of all DHS peak regions identified in male and female mouse liver after either 3-h or 27-h treatment with TCPOBOP or corn oil (vehicle) control (Lodato et al., 2018). A subset of these DHS regions showed a significant increase or decrease in chromatin accessibility following TCPOBOP exposure, as determined by diffReps analysis (Shen et al., 2013), and were defined as either robust or standard differential DHS (Δ DHS) based on whether the difference in normalized sequence read counts between TCPOBOP-treated and vehicle control livers was significant across the entire merged DHS region (robust ΔDHS) or was limited to the differential region identified by diffReps (standard ΔDHS) (Table_S4). DHS width comparisons between different sets of ΔDHS were based on the coordinates defined for the set of 60,739 DHS. For each DNase-seq dataset, FASTQ files from individual biological replicates were concatenated to produce combined samples. Thus, we generated a vehicle-treated combined sample, a TCPOBOP-treated combined sample, and an all-replicates (vehicle-treated + TCPOBOP-treated) combined sample. Foldchange comparisons between ΔDHS subsets were calculated based on normalized DHS read counts between vehicle-treated and TCPOBOP-treated combined samples. The combined list of robust + standard ΔDHS was used in all analysis, except where noted. A set of static DHS (i.e., constitutively open DHS) was used as background for enrichment calculations, and corresponds to the set of 55,866 DHS that did not undergo DHS opening or DHS closing at any of the four TCPOBOP exposure conditions examined (male and female livers, 3-h or 27-h TCPOBOP time points; Table_S4, columns F-I).

2.3.7 Mapping of DHS and ChIP-seq peaks to TCPOBOP-responsive genes

Genes that were significantly induced or repressed in male and/or female mouse liver after 3-h or 27-h TCPOBOP exposure were identified based on a gene list comprised of 24,197 RefSeq genes and 3,152 multi-exonic lncRNA genes (Lodato et al., 2017). Responsive genes were defined by a |fold-change| > 1.5 and adjusted p-value (FDR) < 0.001 for RefSeq genes, and by a |fold-change| > 2 and adjusted p-value (FDR) < 0.05 for lncRNA genes. Each of the 60,739 liver DHS, and flanking chromatin marks, i.e., marks within 400 bp of a DHS region, was assigned a single putative gene target corresponds to the closest gene transcription start site within the same TAD (topologically associating domain) (Supplementary Table_S1B from (Lodato et al., 2018)). Similarly, each CAR ChIP-seq peak was assigned to a single putative gene target based on the closest gene transcription start site within the same TAD as the CAR binding site.

2.3.8 Chromatin mark responses at DHS

For each of the 60,739 liver DHS regions, we characterized the impact of TCPOBOP exposure on each activating chromatin mark (H3K27ac, H3K4me1, H3K4me3) and repressive chromatin mark (H3K27me3) in male mouse liver, as shown in Table_S4. Each DHS was labeled as either induced, repressed, static (unchanged), or absent with respect to each chromatin mark, based on the presence or absence of the mark within 400 bp of the DHS, and then by considering the effect of TCPOBOP exposure on the DHSassociated mark, as determined by diffReps analysis (see above). A DHS was labeled induced with respect to a given chromatin mark if the diffReps-normalized signal intensity for the associated mark was significantly increased at one or both TCPOBOP time points. Similarly, a DHS was labeled repressed if the associated chromatin mark showed a significant decrease in intensity at one or both time points; and a DHS was labeled static (unchanged) with regard to an associated chromatin mark if the mark intensity was not significantly changed at either time point, or if it was unchanged at one time point and the mark was not detected by MACS2 (or SICER), i.e., was absent, at the other time point. A DHS was labeled absent with respect to a given chromatin mark if there were no overlapping or nearby (within 400 bp) MACS2 chromatin mark peaks at

either the 3-h or the 27-h TCPOBOP time point. Finally, a DHS was marked "conflict" if the associated chromatin mark intensity was significantly increased at one time point but was decreased at the other time point, or if it showed discrepant responses between activating chromatin marks.

2.3.9 DHS classification

DHS were designated enhancer DHS if they had an H3K27ac MACS2 peak within 400 bp; the H3K27ac peak could be either induced, repressed, or unchanged (static), at either or both TCPOBOP time points. In many cases, enhancer DHS also overlapped (within 400 bp) a MACS2 peaks for H3K4me1 and/or H3K4me3. An enhancer DHS was designated induced if the DHS was labeled as induced (as described above) with respect to either H3K27ac or H3K4me1, at either TCPOBOP time point. An enhancer DHS was designated repressed if the DHS was labeled as repressed with respect to either H3K27ac or H3K4me1 at either TCPOBOP time point. An enhancer DHS was designated repressed if the DHS was labeled as repressed with respect to either H3K27ac or H3K4me1 at either time point. The designation static enhancer DHS was given to DHS whose H3K27ac marks were unchanged following TCPOBOP exposure, provided that the H3K4me1 marks associated with the DHS were unchanged or absent. Poised DHS are DHS with a H3K4me1 peak within 400 bp that was unchanged, and without a K27ac peak.

A DHS was designated a promoter DHS if it had both a H3K27ac MACS2 peak and an H3K4me3 MACS2 peak within 400 bp of the DHS, and additionally, the DHS region was < 3 kb from the nearest RefSeq gene or multi-exonic lncRNA gene transcription start

site in the same TAD; the H3K4me3 peak could be either induced, repressed or static, at either or both TCPOBOP time points. Induced, repressed, static, and poised promoter DHS were designated based on the responses of the DHS-associated H3K27ac and H3K4me1 marks using the same definitions as the corresponding sets of enhancer DHS described above. A total of 18 H3K4me3-marked DHS that were < 3 kb from a gene transcription start site, but were without a H3K27ac mark, were excluded from the definition of promoter DHS. Thus, the promoter DHS examined here correspond to DHS with enhancer-like properties (i.e., presence of H3K27ac marks) but with the additional features of being proximal to a gene transcription start site and marked by H3K4me3. DHS marked by both H3K27ac and H3K4me3 that were > 3 kb from a transcription start site were designated enhancer DHS, as outlined above. DHS with no chromatin marks are those without a MACS2 peak within 400 bp, for all three activating chromatin marks (H3K27ac, H3K4me1, H3K4me3 peak) and at both TCPOBOP time points. 4,041 of the 60,739 liver DHS lacked these chromatin marks; a subset of these may be insulator DHS, as indicated by their overlap with binding sites for CTCF and cohesin (see Table_S4) (Matthews and Waxman, 2018).

An opening DHS was defined as a DHS showing significant chromatin opening at both time points of TCPOBOP exposure, or at one time point only, provided that the chromatin response was static at the other time point. Similarly, a closing DHS was one that showed significant chromatin closing at one or both time points. A DHS was designated as a static DHS with regard to chromatin opening/closing if there was no significant change in its DNase-seq normalized signal at both time points, as determined by diffReps analysis. Two DHS showed chromatin opening at one TCPOBOP time point and chromatin closing at the other time point and were labeled "conflict". A listing of all 60,739 DHS with the above classifications is provided in Table_S4.

2.3.10 Enhancer class enrichment analysis

Enrichment scores (ES) for a given enhancer class for being at DHS that open or close (for example, enrichment of induced enhancers compared to static enhancers at DHS that open) were calculated as follows. ES = (ratio A / ratio B), where ratio A = number ofenhancer sites at a Δ DHS, divided by the number of enhancer sites at a static DHS; and ratio B = the number of static enhancers at a Δ DHS, divided by the number of static enhancers at static (constitutively open) DHS. For example, of the 1,603 TCPOBOPinduced enhancers, 587 occur at DHS that open and 1,001 occur at static DHS (587/1001 = 0.5864); 15 other induced enhancers occur at DHS that close and were omitted from the enrichment calculation. Further, 779 static enhancers occur at DHS that open, and 29,562 static enhancers occur at static DHS (779/29562 = 0.0264), which gives ES = 22.3 (A/B = 0.5864/ 0.0264). Enrichments of enhancers at ΔDHS for mapping to a TCPOBOPresponsive gene (e.g. enrichment of induced enhancers at opening DHS that map to a TCPOBOP-induced gene compared to induced enhancers at static DHS that also map to an induced gene) were calculated as follows: ES = (ratio A / ratio B), where ratio A =number of enhancer sites at ΔDHS that map to induced genes, divided by the number of enhancer sites at Δ DHS that map to unresponsive genes; and ratio B = number of

enhancer sites at static DHS that map to differentially expressed genes, divided by the number of enhancer sites at static DHS that map to unresponsive genes. For example, the following enrichment calculation indicates the impact of induced enhancers at Δ DHS (compared to static DHS) for mapping to differentially expressed genes: of 587 induced enhancers at DHS that open, 192 map to TCPOBOP-induced genes and 390 map to unresponsive genes (192/390 = 0.4923); 5 other induced enhancers at DHS that open map to repressed genes, and were omitted from the enrichment calculation. Further, 28 of the 1,001 induced enhancers at static DHS map to induced genes, and 849 map to unresponsive genes (128/849 = 0.1508); 30 other induced enhancers at static DHS map to repressed genes and were omitted from the enrichment calculation. Overall, this gives ES = 3.3 (A/B = 0.4923/0.1508). Corresponding enrichments were calculated for enhancer classes at DHS that open/close and for induced/repressed genes in male liver. A Fisher's exact test, implemented in R, was used to calculate the significance of each enrichment.

2.3.11 DHS motif analysis

Motifs present in DHS sequences were identified using FIMO (v4.12.0) (Grant et al., 2011) with the option (--thresh .0005) to improve detection of short length motifs (e.g. hexameric sequences). De novo motif discovery was carried out using two separate algorithms: 1) the cisGenome (Ji et al., 2011) command "flexmodule_motif" was used with default options; and 2) MEME-ChIP (v4.12.0) (Machanick and Bailey, 2011) was implemented with the following options: -meme-mod "anr", to allow for any number of motif occurrences per sequence; -meme-nmotifs 10 and -dreme-m 10, to allow MEME

and DREME, respectively, to report up to 10 discovered motifs; and stringent E-value thresholds, -dreme-e 1e-10 and -filter-thresh 0.01, for reporting significant de novo motifs. MEME-ChIP de novo motifs with (MEME/DREME) E-value < E-10 were used for subsequent analysis. To determine motif enrichment, sequences were scanned using a combined list of cisGenome and MEME-ChIP de novo motifs and applied to the set of DHS that open in both male and female liver for both time points compared to a background set of 55,866 static DHS (see above). The motifStack package (Ou et al., 2018) was used to cluster motifs by sequence similarity. Redundant motifs in these motif clusters were removed by retaining the motif showing the strongest motif enrichment as the representative motif for the overall group. This yielded a final set of 21 motifs, which were used for downstream analysis. Tomtom (v4.12.0) (Gupta et al., 2007) was used to evaluate pairwise sequence similarity of each de novo motif against a database of 97 motif families (Macisaac et al., 2006) or against the JASPAR motif database of 579 annotated motifs (Khan et al., 2018). Enrichments of motif occurrence in ΔDHS were calculated for each ΔDHS set (e.g. enrichment of CAR motif, motif65, in the set of DHS that open in male liver after 3-h TCPOBOP exposure, etc.) as follows: Enrichment score (ES) = (ratio A / ratio B), where: ratio A = number of ΔDHS that contain the CAR motif, divided by the number of Δ DHS that lack the CAR motif; and ratio B = the number of static DHS that contain the CAR motif, divided by the number of static DHS that lack the CAR motif. For example, in male mice exposed to TCPOBOP for 3-h, there were 472 opening DHS, of which 311 contain the CAR motif and 161 lack the CAR motif (ratio A = 311/161 = 1.93), whereas 27,937 static DHS contain the CAR motif, and 27,929 static

DHS lack the CAR motif (ratio B = 27,937/27,929 = 1.00), which gives ES = 1.9 (A/B = 1.93/1.00).

2.3.12 Receiver Operating Characteristics (ROC) analysis

ROC curves and area under the curve values were calculated to determine the enrichment of de-novo and constructed motifs in Δ DHS subsets compared to a background of static DHS. A 500 bp region centered at the midpoint of each DHS was input for the foreground (e.g. Δ DHS) and background (e.g. static DHS) sets of sites for FIMO motif scans. FIMO reports a p-value (i.e. match score = -log(p-value)) for the motif match for each site in the foreground and background sequences. Data points for each ROC curve were generated by calculating the true positive rate and the false positive rate for several cutoffs of the match score. The true positive rate was defined as follows: (number of foreground sites with a score \geq match score cutoff) / (total number of foreground sites). The false positive rate was defined as follows: (number of background sites with a score \geq match score cutoff) / (total number of background sites). Significant differences in score distributions between foreground and background sets and area under the curve values were calculated using a Wilcoxon rank sum test, as implemented in R.

2.3.13 CAR ChIP-seq data analysis

ChIP-seq data for CAR binding in mouse liver (Niu et al., 2018), obtained by ChIP-exo, where the ChIP protein-bound DNA complex is trimmed to within a small number of nucleotides of the protein binding site by exonuclease digestion, was downloaded from

GEO, accession GSE112199. Thus, we obtained CAR binding site coordinates (BED files) and corresponding signal tracks (bigWig files) for ChIP-seq analysis of mouse CAR and human CAR protein expressed in livers of mice deficient in CAR after treatment with the direct activators of CAR (TCPOBOP for mouse CAR, and the human CAR agonist CITCO (6-(4-chlorophenyl)imidazo[2,1-b][1,3]thiazole-5-carbaldehyde-O-(3,4dichlorobenzyl)oxime) for human CAR), or after treatment with the indirect CAR activator phenobarbital (Niu et al., 2018). The UCSC Genome Browser (http://genome.ucsc.edu/) Utility liftOver command was used to convert BED and bigWig files from mouse genome assembly mm10 to mm9. Overlap analysis between these four sets of CAR binding site data and the set of 60,739 liver DHS (above) was performed using BEDTools to identify subsets of CAR binding sites at DHS, which are shown in Table_S5A, columns R-U. Significant differences between MACS14 score distributions among subsets of CAR binding sites (e.g. those outside vs. inside DHS) were calculated using a Wilcoxon Rank Sum test with a Benjamini-Hochberg p-value correction, as implemented in R.

Raw sequencing data from a second CAR ChIP-seq study (Tian et al., 2018) was downloaded from the Sequence Read Archive (accession SRP125957, sample SRR6337710 and samples SRR6337727-SRR63377236), to obtain CAR binding data for 8 week old female mouse liver exposed to vehicle control (n = 4) and 3-h TCPOBOP treatment (n = 7) livers. The custom ChIP-seq pipeline described above was used to process this data. MACS2 (v2.1.0.20150731) using the option (--keep-dup) and with all other parameters set to the default option, was used to discover ChIP-seq peaks for each individual replicate. Two of the biological replicates were determined to be low quality samples and were excluded from further analysis: control liver replicate-1, which yielded 18,068 MACS2 peaks, as compared to only 18 to 125 MACS2 peaks for each of the other three control liver replicates; and TCPOBOP liver replicate-5, which gave a very high percentage of reads in artefactual straight-read peaks (23.3% of all reads vs. < 0.2% for the other samples), likely indicating excessive qPCR duplicates. The retained ChIP-seq biological replicates were combined to create a single control sample (n=3) and a single TCPOBOP sample (n=6), which yielded 361 and 12,587 CAR ChIP-seq peaks, respectively, by MACS2 analysis. BEDTools "merge" was used to concatenate these peak lists to generate 12,728 CAR binding sites. diffReps was then used to identify sites of differential CAR ChIP-seq signal intensity between the control and TCPOBOP-treated livers. CAR diffReps regions that overlap either the set of 12,728 MACS2 peaks of CAR binding, or the set of 60,739 DHS, were combined, resulting in a set of 3,734 induced CAR binding sites and 103 repressed CAR binding sites. Of these, 3,021 TCPOBOPinduced CAR binding sites were found at a total of 2,901 DHS; these sites are listed in Table_S5B and summarized in Table_S5C, and were used for downstream analysis. For some of our analyses, we used a filtered set of 2,525 CAR binding sites that overlap H4K5ac peaks as reported by (Tian et al., 2018), of which 2,287 sites (90%) overlapped the set of 60,739 liver DHS (Table_S5A).

A set of 335 stringent CAR-free opening DHS was selected from the full set of 2,029 male liver opening DHS based on the absence of any overlapping CAR binding site in the five CAR ChIP-seq datasets described above, namely, mCAR binding in livers of mice stimulated with TCPOBOP or phenobarbital, hCAR binding in livers of mice stimulated with CITCO or phenobarbital (Niu et al., 2018), and endogenous mCAR binding using the set of CAR binding sites at DHS computed from the data of (Tian et al., 2018) as outlined above, or using the filtered set of 2,525 CAR binding sites reported by (Tian et al., 2018). A set of 455 robust CAR-bound opening DHS was selected from the full set of 2,029 opening DHS based on the presence of an overlapping CAR binding site in all five CAR ChIP-seq datasets. We used the same criteria to select static DHS characterized by robust CAR binding (n = 2,535 DHS) and static DHS characterized by stringent CAR-free DHS regions (n = 26,882 DHS) from the full set of 58,078 DHS that did not undergo DHS opening or DHS closing in male liver either 3-h or 27-h after TCPOBOP exposure (Table_S5A, column V).

2.3.14 CAR ChIP-seq enrichment analysis

Enrichments for containing a CAR binding site were calculated for each set of Δ DHS. CAR binding sites used in these analyses were those determined for mCAR and hCAR by (Niu et al., 2018), or were the set of 2,901 TCPOBOP-induced CAR binding sites that we identified using the ChIP-seq data of (Tian et al., 2018). Enrichments were calculated as follows: ES = (ratio A / ratio B), where ratio A = number of DHS that open following TCPOBOP exposure and that contain a CAR binding site, divided by the number of Δ DHS from the same Δ DHS set that lack a CAR binding site; and ratio B = the subset of 55,866 static DHS that contain a CAR binding site, divided by the number of static DHS that lack a CAR binding site. For example, 118 DHS that open in female mice treated with TCPOBOP for 3-h contain a CAR binding site, and 394 other DHS that open lack a CAR binding site (118/394 = 0.2995), whereas 2,441 static DHS contain a CAR binding site, and 53,425 static DHS lack a CAR binding site (2441/53425 = 0.0465), which gives ES = 6.6 (A/B = 0.2995/0.0465). Similar methods were used to calculate enrichments for the set of repressed CAR binding sites at DHS that close, and for the subset of induced CAR binding sites that also contain either a de novo discovered nuclear receptor (NR) half-site motif, a DR4-like motif or both motifs.

Enrichments of Δ DHS that contain a CAR binding site mapping to a TCPOBOPresponsive gene were calculated for each Δ DHS set, as follows: ES = (ratio A / ratio B), where ratio A = number of DHS that open and that contain a CAR binding site mapping to the corresponding set of TCPOBOP-induced genes, divided by the number of Δ DHS from that same set of Δ DHS whose putative gene target is unresponsive to TCPOBOP at that same time point; and ratio B = the number of static DHS that lack a CAR binding site and that map to the same set of TCPOBOP-induced genes, divided by the number of static DHS that lack a CAR binding site and whose putative gene target is unresponsive to TCPOBOP at that same time point. For example, in female mice after a 3-h TCPOBOP exposure, 61 DHS that open also contain a CAR binding site mapping to a 3-h TCPOBOP-induced gene, and 300 other DHS that open and contain a CAR binding site do not map to a gene that responded to 3-h TCPOBOP exposure (61/300 = 0.2033), whereas 493 static DHS that lack a CAR binding site map to a 3-h TCPOBOP-induced gene, and 48,461 static DHS that lack a CAR binding site do not map to a TCPOBOP-responsive gene (493/48461 = 0.0102), which gives ES = 19.9 (A/B = 0.2033/0.0102). Corresponding enrichment calculations were performed using the subset of CAR binding sites that also contain either a de novo discovered NR half-site motif, a DR4-like motif, or both motifs. Similar methods were used to calculate enrichments for the frequency at which robust CAR-bound opening DHS and stringent CAR-free opening DHS mapped to TCPOBOP-induced genes, as compared to the corresponding sets of robust CAR-bound and stringent CAR-free static DHS.

2.3.15 Enrichment of TCPOBOP-responsive genes for nearby CAR binding sites Enrichments of TCPOBOP-responsive RefSeq and multi-exonic lncRNA genes for containing a CAR binding site in the same TAD were calculated for each TCPOBOP time point, as follows: ES = (ratio A / ratio B), where ratio A = the number of differentially-expressed genes that respond to a given TCPOBOP exposure that contain a CAR binding site, divided by the number of differentially-expressed genes that lack a CAR binding site; and ratio B = the subset of TCPOBOP-unresponsive genes that contain a CAR binding site, divided by the number of unresponsive genes that lack a CAR binding site. For example, in male mice exposed to TCPOBOP for 3-h, 97 induced genes contain a binding site (97/41 = 2.37), whereas 6,599 unresponsive genes each contain a CAR binding site, and 20,577 genes lack CAR binding (6,599/20,577 = 0.3207), which gives ES = 7.4 (A/B = 2.37/0.3207).

2.3.16 TCPOBOP-responsive genes dependent on MET/EGFR signaling

DAVID Conversion tool was used to convert 1,248 Affymetrix probe IDs associated with genes that were either induced or repressed after 2 days TCPOBOP exposure (GEO dataset accession #GSE110695) (Bhushan et al., 2018) to 1,029 unique gene symbols, of which 962 genes were represented on our RefSeq gene list and used for downstream analyses. 369 of these TCPOBOP-responsive genes (38%) were differentially expressed in livers of 2 day TCPOBOP-exposed mice deficient in the MET proto-oncogene and inhibited with respect to EGF receptor signaling (MET/EGFR deficient livers) (Bhushan et al., 2018) compared to TCPOBOP-exposed control mice (top DAVID enriched terms: cell cycle, cell division, ES = 9.27, FDR = 1.54E-09), and 593 genes (62%), including essentially all the classic *Cyp* and other TCPOBOP-responsive drug-metabolizing enzyme genes, showed MET and EGFR-independent responses to TCPOBOP (top DAVID enriched term: Endoplasmic reticulum, ES = 10.0, FDR = 2.55E-08); these findings are fully consistent with (Bhushan et al., 2018). Further, 187 of the 369 genes were induced >2-fold by TCPOBOP, and further, their responses to TCPOBOP were reduced >2-fold in MET and EGFR signaling-deficient liver; 320 of the 593 gene set were induced >2-fold by TCPOBOP, but did not show a >2-fold decrease in response in the absence of MET/EGFR signaling. The enrichment of CAR binding sites mapping to both the 187 and the 320 gene sets was calculated as follows: ES = (ratio A / ratio B),

where ratio A = the number of induced genes with a CAR binding site, divided by the number of induced genes that lack a CAR binding site; and ratio B = the number of TCPOBOP-unresponsive genes that contain a CAR binding site, divided by the number of unresponsive genes that lack a CAR binding site. Table_S6 shows these gene lists, their associated CAR binding sites and DAVID enrichment analysis.

2.3.17 RXR and CEBP ChIP-seq data analysis

ChIP-seq datasets for the transcription factors RXRA, CEBPA, and CEBPB were analyzed using the nucleosome option in diffReps (200 bp window) to identify genomic regions with ChIP-seq signals that differed significantly between vehicle-treated and TCPOBOP-treated ChIP samples at each time point (diffReps regions with a |foldchange | > 2 in diffReps-normalized read counts and FDR < 0.05 (Benjamini-Hochberg) adjusted p-value)). Separately, MACS2 was used to identify transcription factor binding sites (i.e., ChIP-seq peaks) for each biological replicate. For each transcription factor and at each time point, ChIP-seq peak sets were merged across biological replicates using the BEDTools "merge" command to generate a peak union list. The overlap between diffReps regions and the corresponding MACS2 peak union list was then determined. Peak union sites that did not overlap a diffReps region were annotated as static ChIP-seq binding sites. Concatenation of the static binding sites with the diffReps-reported regions yielded a final set of transcription factor binding sites used for downstream analysis. The numbers of binding sites discovered for each transcription factor and those overlapping any of the 60,739 DHS regions are shown in Table_S1B.

2.3.18 Peak normalization

To visualize ChIP-seq regions in the UCSC genome browser (https://genome.ucsc.edu/), samples were normalized by using the number of sequence reads in each ChIP-seq peak per million mapped sequence reads (reads-in-peaks-per-million, RiPPM) as a scaling factor. Briefly, to obtain a comprehensive list of ChIP-seq peaks for each dataset, we first concatenated FASTQ files from individual biological replicates. Thus, we generated a vehicle-treated combined sample, a TCPOBOP-treated combined sample, and an all-replicates (vehicle-treated + TCPOBOP-treated) combined sample. ChIP-seq peaks identified by analyzing each individual sample were concatenated with the peaks identified by analyzing the combined samples by using BEDTools "merge" to generate a single list of non-overlapping ChIP-seq peaks. The fraction of reads in peaks for each sample was then calculated to obtain a scaling factor. Raw read counts were divided by this per-million scaling factor to obtain RiPPM normalized read counts.

2.4 Results

2.4.1 TCPOBOP-induced changes in chromatin accessibility may be transient, sustained or late

TCPOBOP induces both increases and decreases in chromatin accessibility, which can be identified as differential DNase hypersensitive sites (Δ DHS) between TCPOBOPexposed and control mouse livers (Lodato et al., 2018). We reasoned that classifying changes in DHS in a temporal and sex-specific manner would help explain regulation of TCPOBOP-responsive genes. We classified the opening chromatin regions as follows. Sustained early Δ DHS opened within 3-h of TCPOBOP exposure and were remained open through the 27-h time point (208 such sites in male liver, and 341 in female liver). Many more DHS were late Δ DHS, which did not open until 27-h (1,558 sites in male liver and 2,013 in female liver). Finally, transient Δ DHS were opened at 3-h, but not at 27-h (Fig. 2.1A). The sustained early Δ DHS showed a striking 43.5-fold enrichment for mapping to genes induced by TCPOBOP at 3-h, and up to 66% of these Δ DHS mapped to up-regulated genes by 27-h. Lower enrichments were seen for the other sets of Δ DHS (Fig. 2.1B, male liver; Fig. S2.1A, female liver). The sustained early and the late Δ DHS showed significantly greater increases in chromatin opening (Fig. 2.1C) and more robust responses to TCPOBOP than the transient Δ DHS (58-85% vs 20-40%; Fig. S2.1B, Fig. S2.1C).

 Δ DHS regions displayed a 60-fold range in chromatin accessibility (Fig. 2.1D, x-axis) and a wide range of responsiveness to TCPOBOP (Fig. 2.1D, y-axis). Many fewer DHS closed than were opened by TCPOBOP exposure at both time points (positive vs. negative y-axis values). The sustained early Δ DHS showed a significant increase in chromatin opening over time (Fig. 2.1C; Fig. 2.1D, shift in green dots to the right in lower vs. upper panels). Further, the width of the Δ DHS region was significantly greater for sustained early Δ DHS than for late-opening DHS (Fig. S2.2A). Sex differences in TCPOBOP-induced changes in chromatin accessibility were apparent (Fig. S2.3, Fig. S2.2B) and likely contribute to the sex differences in TCPOBOP-induced gene responses seen in mouse liver (Lodato et al., 2017). Δ DHS common to both sexes were more likely to show a robust response than Δ DHS specific to one sex (Fig. S2.1D, Fig. S2.1E).

2.4.2 TCPOBOP-induced changes in active histone marks

Based on the observation that TCPOBOP alters chromatin accessibility, we reasoned that changes in histone marks around these DHS would help with identifying these sites as TCPOBOP responsive promoter and enhancer regions. To further characterize TCPOBOP-induced changes in the epigenetic landscape, we used ChIP-seq to identify mouse liver genomic regions (sites) with histone marks that are significantly induced or repressed after TCPOBOP exposure. Histone mark changes were most common for H3K27ac, which identifies active enhancers, followed by H3K4me1, which marks active and poised enhancers, and H3K4me3, which marks active promoters (Fig. 2.2A). For all three active chromatin marks, histone mark induction at 27-h was strongly enriched for DHS that open, as compared to the corresponding background sets of TCPOBOPunresponsive (i.e., static) histone marks (Fig. 2.2B; H3K27ac, ES = 54.4, p =0; H3K4me1, ES = 29.4, p = 2.96E-43; and H3K4me2, ES = 120.4, p= 1.52E-29). Notably, 49 to 57% of each set of 27-h TCPOBOP-induced histone marks mapped to opening DHS. Genomic regions with repressed H3K27ac marks at 27-h were enriched for DHS that close at the 27-h time point (ES = 15, p = 4.26E-67) (Fig. 2B, bottom). Further, the opening DHS were significantly enriched for all three sets of induced marks at both time points, as compared to a background set of unchanging, constitutively open (static) DHS (Fig. S2.4). While a substantial fraction (33.4%) of 27-h TCPOBOP-induced ΔDHS were characterized by induced H3K27ac marks, only a small fraction of Δ DHS was associated with induced H3K4me1 or H3K4me3 marks. This finding reflects the comparatively small number of TCPOBOP-responsive H3K4me1 and H3K4me3 sites that we identified (Fig. 2.2A). A listing of all chromatin mark peak sets and corresponding annotations is provided in Table_S2 and their associations with DHS are shown in Table_S3.

2.4.3 Many ADHS are activated promoters and enhancers that map to TCPOBOPresponsive genes

To determine the relationship between ΔDHS regions and TCPOBOP-responsive enhancers and promoters, we used the above active histone mark datasets to classify the full set of 60,739 liver DHS, based on whether they were associated with promoter marks (12,957 DHS, 21.3%), enhancer marks (43,542 DHS, 71.7%) or no activating marks (4,041 DHS, 6.7%) (Fig. S2.5A); the latter DHS include insulators found at the TAD and intraTAD boundaries that anchor chromatin loops and segment each chromosome (Matthews and Waxman, 2018). The promoter and enhancer DHS were designated induced, repressed, static or poised, based on the patterns of response of their associated histone marks to TCPOBOP exposure (see Methods) (Fig. 2.2C). Enhancer marks (H3K27ac and/or H3K4me1) were induced by TCPOBOP at 1,603 enhancer DHS, and were repressed at 739 enhancer DHS. Promoter DHS showed similar patterns of active histone mark induction and repression by TCPOBOP (Fig. 2.2C). Thus, TCPOBOP induces or represses enhancer marks at ~3-5% of all enhancer and promoter DHS. The remaining 95-97% of enhancer DHS were either static enhancers (71%), i.e., H3K27ac marks were not significantly altered by TCPOBOP exposure, or poised enhancers (24%), i.e., they are deficient in H3K27ac marks but contain H3K4me1 marks that are unchanged (see Methods).

The increases in active marks at enhancer DHS and at promoter DHS were rarely accompanied by a loss of H3K27me3 marks, as shown by ChIP-seq analysis. Thus, H3K27me3 marks decreased significant at only 7 of the 1,603 enhancer DHS, and at only 3 of the 215 promoter DHS where active histone marks increased after TCPOBOP exposure. Furthermore, H3K27me3 marks increased at only 4 of the 739 enhancer DHS, and at only 3 of the 156 promoter DHS where TCPOBOP repressed active marks (Table_S4). Thus, a change in H3K27me3 (repressive) marks is not a common mechanism for regulation of TCPOBOP-responsive enhancer and promoter DHS.

Next, we investigated the link between TCPOBOP-induced chromatin opening and closing (Δ DHS) and the above changes in histone marks at enhancers and promoters (Fig. 2.2D, Fig. 2.2E). The set of 1,603 TCPOBOP-induced enhancer DHS showed very strong enrichment for being at DHS that open (ES = 22.3, p = 0), while the set of 739 TCPOBOP-repressed enhancers was significantly enriched for being at DHS that close (ES = 7.8, p = 4.34E-33) when compared to a background set of static enhancers at Δ DHS (Fig. S2.5B). Similarly, induced promoter DHS were strongly enriched (ES = 32.6, p = 1.74E-52) for being at DHS that open when compared to a background set of static open set of static promoters at Δ DHS (Fig. S2.5C), while repressed promoter DHS were enriched for

 Δ DHS that close (ES = 7.5, p = 2.64E-04). However, only a small fraction of all repressed promoter and enhancer DHS were at Δ DHS that close (Fig. 2.2D, Fig. 2.2E). No such enrichments were seen for poised enhancers or promoters, or for DHS without enhancer marks, which were almost always associated with static DHS (3,896 of 4,041 DHS; 96%).

To better understand the relationship between these strong positive associations between changes in chromatin accessibility and activating histone marks, we examined their impact on enrichments for TCPOBOP target genes. The strongest enrichment for TCPOBOP-induced genes (ES = 11.7, p = 5.47E-111) was seen when comparing to the set of 587 opening DHS with induced enhancer marks (Fig. 2.2D) to a background set of static enhancers with static enhancer marks (Fig. S2.5D). Enrichments for TCPOBOPinduced genes were reduced but still significant when we analyzed the set of 770 opening DHS with static enhancer marks (ES = 3.6, p = 1.15E-24; Fig. S2.5D). A similar level of enrichment was obtained when the 587 opening DHS with induced enhancer marks were compared to alternative background sets of enhancer DHS, comprised of opening DHS with unchanged enhancer marks, or static DHS with induced enhancer marks (ES = 3.2-3.3, p = E-18 to E-20; Fig. S2.5D, comparisons 3 and 4). This indicates that TCPOBOPstimulated chromatin opening and acquisition of activating enhancer marks both contribute to induction of TCPOBOP target genes. Supporting this conclusion, gene targets of opening DHS showed significantly greater increases in expression than gene targets of static DHS, and opening DHS with induced histone marks showed greater

increases in expression than opening DHS with static histone marks (Fig. 2.2F).

Corresponding conclusions regarding the contributions from both chromatin closing and loss of activating enhancers were apparent when we compared the enrichments for TCPOBOP-repressed genes of 66 closing enhancer DHS with repressed enhancer marks (ES = 6.6, p = 9.68E-06) with that of 382 closing enhancer DHS with static enhancer marks (ES = 3.6, p = 1.16E-09) (Fig. S2.5E).

2.4.4 Enrichment of nuclear receptor DR4 motif at ΔDHS

TCPOBOP-activated CAR and related nuclear receptor superfamily members bind their cognate DNA regulatory sequences as a heterodimer with RXR (Wagner et al., 2017; Waxman, 1999). The bound DNA sequence (DNA response element) is typically comprised of repeats of the consensus motif AG(G/T)TCA, which may be organized as a direct repeat (DR), an everted repeat (ER), or an inverted repeat (IR) (Wang et al., 2012). Using de novo motif discovery (Ma et al., 2014), we identified a DR motif with 4 nt spacing (DR4 motif) in both the 3-h male and the 3-h female Δ DHS sequence sets, while the 27-h Δ DHS sequences yielded a related NR half-site motif (Fig. 2.3A). The DR4 motif closely matches a canonical motif for CAR, PXR and vitamin D receptor (VDR) (motif 65; Fig. 2.3A, top). A majority of DHS that open either 3-h or 27-h after TCPOBOP treatment contain this motif, which is significantly enriched compared to a background set of static DHS sequences (Fig. 2.3B). No such enrichment was seen in the sets of closing DHS (Fig. 2.3C). The sustained early Δ DHS showed stronger CAR motif enrichment than the transient and late Δ DHS (Fig. 2.3D), consistent with their stronger

association with TCPOBOP-responsive genes (Fig. 2.1B). Motif enrichment was strongest in the set of 3-h TCPOBOP-induced Δ DHS identified in both male and female liver (i.e., male-female common Δ DHS). This enrichment declined by 27-h (sexindependent Δ DHS at 27-h vs. 3-h; Fig. 2.3E), which may explain our de novo identification of the DR4 motif in the 3-h but not the 27-h Δ DHS sequences (Fig. 2.3A) and suggests TCPOBOP activates secondary chromatin opening events that do not directly involve nuclear receptor binding to DR4 motifs. Lower enrichments in the 27-h Δ DHS sequences were also seen when scanning with the de novo discovered motifs (Fig. S2.6A).

Given the discovered NR half-site motifs (Fig. 2.3A), we considered whether opening DHS also show enrichment for NR motifs with different spacing or orientation than the DR4 motif. We scanned each opening DHS set using DR, ER, and IR motifs constructed from the de novo female liver NR half-site motif, with spacing ranging from 0 to 5 nt between half-sites (Fig. S2.6B). Strongest enrichments compared to a static DHS background set were obtained with the constructed DR4 motif: 5.5-fold enrichment for male-female common Δ DHS at 3-h, decreasing to 2.4-fold at 27-h (Fig. 2.3F). Significant, but more modest enrichments were seen for other motif orientations and spacings, most notably ER1, ER4 and DR3 (Fig. 2.3F, Fig. S2.6C). Closing DHS did not show significant enrichment for any NR half-site-based motifs (Fig. S2.6C). The strong preferential enrichment of the DR4 motif was validated by ROC curve analysis

comparing opening DHS to the background static DHS set (Fig. 2.3G, Fig. 2.3H, Fig. S2.7).

2.4.5 CAR binding at TCPOBOP-responsive DHS

We hypothesize that the direct binding of activated CAR is an important mechanism for explaining DHS regulation of TCPOBOP-responsive genes. We used two recently published ChIP-seq datasets to characterize CAR binding to TCPOBOP-responsive ΔDHS . In one study, genome-wide binding sites were identified by ChIP-exo for both mouse CAR (mCAR) and human CAR (hCAR) in transgenic male mouse livers. Mice were treated with TCPOBOP, to activate mCAR, or with the hCAR-specific agonist ligand CITCO, to activate hCAR. Alternatively, mCAR and hCAR were activated indirectly by treatment of the transgenic mice with phenobarbital (Niu et al., 2018). A majority (72-75%) of the mCAR and hCAR binding sites identified occur at one of the 60,739 liver DHS described in this study (Table_S1B). CAR binding within a DHS was significantly stronger than CAR binding outside of a DHS region (Fig. S2.8A), and hCAR showed more binding sites (Table_S1B) and stronger binding signals than mCAR (Fig. S2.8B), consistent with (Niu et al., 2018). Importantly, CAR binding was seen at both ΔDHS and at constitutively open, i.e., TCPOBOP-unresponsive (static) DHS, although binding was generally stronger at the opening DHS (Fig. S2.8B). Moreover, DHS that open following TCPOBOP exposure a highly significant 10 to 18-fold enrichment for containing mCAR binding sites when compared to a background set of static DHS, as was seen in both male and female mouse liver at both TCPOBOP time

points (Fig. 2.4A, Fig. S2.9A). Strikingly, mCAR binding was enriched 22.3-fold at sustained early Δ DHS vs. 6.1-fold at late Δ DHS, and correspondingly, mCAR binding was found at 76% of sustained early Δ DHS vs. only 46% of late Δ DHS (Fig. S2.9C; c.f. mCAR binding at 12% of static DHS). In contrast, mCAR binding was not enriched at closing DHS, except for a modest 1.4-fold enrichment at 27-h Δ DHS in female liver (Fig. S2.9B). We conclude that a large majority of CAR binding occurs at DHS that open following TCPOBOP exposure, and furthermore, that a substantial fraction, although not all DHS opening is closely associated with CAR binding.

We confirmed these conclusions using a second CAR ChIP-seq dataset, where binding sites for endogenous mCAR were identified in TCPOBOP-stimulated female mouse liver (Tian et al., 2018). Our analysis of this dataset (see Methods) identified 3,734 genomic sites where CAR binding is induced and 103 sites where it is repressed by TCPOBOP exposure (Table_S5B). DHS opened by TCPOBOP were strongly enriched (ES = 5.4-6.6 at 3-h) for induced CAR binding compared to a background set of static CAR binding sites (Fig. S2.9D). Even stronger enrichments were seen for the TCPOBOP-induced CAR binding sites that contained either a de novo NR half-site or a DR4-like motif (ES = 8.6-11.3) (Fig. S2.9D). DHS that close in male liver after 27-h TCPOBOP exposure were significantly enriched for sites where CAR binding is repressed, however, these repressed binding sites represent only 4% of closing DHS (Fig. S2.9E) and thus do not comprise a major mechanism for chromatin closing.

2.4.6 CAR binding at TCPOBOP-responsive genes

We hypothesize that proximal CAR binding plays a direct role in regulating TCPOBOP induced changes in gene expression. To ascertain the relationship between CAR binding and TCPOBOP-stimulated gene responses, we mapped the binding sites for mCAR and hCAR to the nearest gene transcription start site within the same TAD (putative gene targets of CAR). CAR binding was significantly enriched for TCPOBOP-responsive genes, with up to 87% of TCPOBOP-induced genes associated one or more CAR binding sites (Fig. S2.10A). Further, ~2-fold stronger enrichment was seen for genes induced by TCPOBOP after 3-h than after 27-h, when many secondary gene induction responses that do not involve direct binding by CAR are anticipated. Stronger enrichments for TCPOBOP-responsive genes at 3-h compared to 27-h were also seen when using the second CAR ChIP-seq dataset (Fig. S2.10D). Interestingly, up to 79% of 27-h TCPOBOP-repressed genes were targeted by a CAR binding site (Fig. S2.10B). Together, these findings support the proposal that CAR plays a direct role in DHS opening leading to gene induction, and perhaps also gene repression. Examples of CAR binding sites at Δ DHS proximal to induced genes are shown in Fig. 2.5.

Liver growth factor receptor signaling, via either the growth factor receptor MET or EGFR, is required for ~40% of all TCPOBOP-stimulated gene responses, in particular for induction of genes active in DNA replication and cell cycle regulation linked to hepatocyte proliferation (Bhushan et al., 2018). MET/EGFR signaling is not required, however, for TCPOBOP induction of *Cyp* and other drug-metabolizing enzyme genes (Bhushan et al., 2018). Given the MET/EGFR signaling requirement for TCPOBOP induction of the DNA replication and cell cycle genes, and the fact that these genes do not respond to TCPOBOP until 27-h (Lodato et al., 2017), we anticipated that their induction by TCPOBOP might be a secondary response to CAR activation, i.e., one that does not involve direct CAR binding. CAR binding site analysis indicated, however, that many of the TCPOBOP-induced DNA replication and cell cycle regulation genes are targets of CAR, as indicated by CAR ChIP-seq binding at gene proximal DHS regions (Fig. 2.6, Table_S6). Further, TCPOBOP-stimulated CAR binding was as strongly enriched at the MET/EGFR signaling-dependent gene set as it was at the MET/EGFRindependent gene set, with hCAR binding showing results very similar to mCAR (Fig. 2.6F, Fig. S2.10C).

2.4.7 DHS opening in the absence of CAR binding

While CAR binding was found at a large fraction of opening DHS (Fig. S2.9A), DHS opening also occurred at many sites where CAR does not bind, as was seen at three of the six Δ DHS upstream of the highly TCPOBOP-responsive gene *Cyp2b10* (Fig. 2.5A). To better characterize these sites, we identified 335 opening DHS that are stringently free of CAR binding, defined as opening DHS regions that do not overlap a binding site for mCAR or hCAR in any one of the five CAR ChIP-seq datasets that we examined (see Methods). We compared these 335 DHS to a set of 455 robust CAR-bound opening DHS, defined as opening DHS where CAR binding is seen in all five ChIP-seq datasets. Both sets of opening DHS were compared to corresponding background sets of DHS that are

constitutively open and do not undergo further chromatin opening, namely, static DHS that are stringently CAR-free (26,882 DHS), and static DHS with robust CAR binding (2,535 DHS), respectively (Fig. 2.7A, Table_S5A). The stringent CAR-free opening DHS exhibited a similar increase in chromatin opening as the robust CAR-bound opening DHS (Fig. 2.7B, left), but were weaker, in terms of normalized DHS signal intensity (Fig. 2.7B, right). Further, the stringent CAR-free opening DHS were depleted of NR half site and DR4 motif sequences, as were the stringent CAR-free static DHS (Fig. 2.7C) and consistent with the absence of CAR binding. The stringently CAR-free opening DHS were less frequently associated with either basal or TCPOBOP-stimulated enhancer marks (Table_S5D), but mapped to TCPOBOP-induced genes as frequently as robust CAR-bound opening DHS (25.1% of DHS vs. 28.4%; p > 0.2, Fisher's Exact test). Moreover, the stringently CAR-free opening DHS were enriched for mapping to TCPOBOP-induced genes (ES = 8.26, p< 0.0001 vs CAR-free static DHS) at least as strongly as robust CAR-bound opening DHS (ES = 5.04, p< 0.0001 vs robust CARbound static DHS) (Table_S5E). Interestingly, of 84 stringent CAR-free opening DHS that mapped to a TCPOBOP-induced gene (i.e., 25.1% of all 335 stringent CAR-free opening DHS), 70 mapped to the same TCPOBOP-inducible gene as did one or more CAR-bound opening DHS. Thus, a large fraction of the target genes of stringent CARfree opening DHS are also targets of a CAR-bound opening DHS. These findings show that TCPOBOP stimulates chromatin opening at many DHS that do not bind CAR, but are in the vicinity of, and map to the same TCPOBOP-inducible gene as DHS whose TCPOBOP-stimulated chromatin opening is associated with CAR binding (Fig. 2.7D).
2.4.8 TCPOBOP-induced binding of RXRA, CEBPA and CEBPB at TCPOBOPresponsive ΔDHS

We performed ChIP-seq analysis for RXRA, the heterodimeric partner of CAR, and for the basic leucine zipper transcription factors CEBPA and CEBPB to investigate their potential role as co-factors for TCPOBOP responses (Benet et al., 2010; Blazquez et al., 2013). ChIP-seq peaks induced or repressed following TCPOBOP exposure are summarized in Table_S1B. TCPOBOP-induced RXRA binding at 3-h was strongly enriched for occurring at DHS that open at 3-h (ES = 11.8, p = 5.5E-25) compared to the set of static RXRA binding sites. The enrichment of TCPOBOP-stimulated RXRA binding at opening DHS was even stronger at 27-h (ES = 32.6, p = 0) (Fig. 2.4C). Further, at 27-h, but not at 3-h, TCPOBOP-induced binding of CEBPA and CEBPB to Δ DHS was highly enriched (ES = 31.9, p = 1.6E-46; ES = 37.9, p = 7.4E-170, respectively) when compared to static binding for each transcription factor. Moreover, for all three transcription factors, a substantial fraction of induced binding sites at 27-h were found at DHS that open (50-56%) (Fig. 2.4C). TCPOBOP-repressed transcription factor binding at 27-h was strongly enriched at DHS that close at the 27-h time point (Fig. 2.4C, right). Transcription factor binding at DHS that open, particularly co-factor binding with activated CAR, may alter chromatin accessibility and thereby facilitate TCPOBOPinduced gene expression responses.

In complementary analyses, DHS that opened after TCPOBOP exposure were strongly enriched for induced RXRA binding: ES = 27.5 (p = 3.5E-34) at 3-h, and ES = 49.9 (p = 0) at 27-h, compared to a background set of static DHS. The opening DHS were highly enriched for TCPOBOP-induced binding of CEBPA (ES = 41.2, p = 3.78E-48) and CEBPB (ES = 56.7, p = 1.6E-186) at 27-h but not at 3-h (Fig. 2.4D). Further, DHS that close at 27-h showed strong enrichment (compared to static DHS) for TCPOBOPdependent repression of all three transcription factors at the 27-h time point (Fig. 2.4D).

2.4.9 Enriched motifs for associated co-factors

De novo motif analysis of Δ DHS regions was carried out to discover motifs for factors that may cooperate with CAR to regulate chromatin accessibility. Some of these motifs may be preferentially found in the 27-h Δ DHS regions, and could contribute to the secondary changes in chromatin and gene expression that follow the primary TCPOBOPstimulated responses. We discovered a total of 21 significant motifs (see Methods), which were grouped into 5 clusters based on sequence similarity (Fig. S2.11). Comparison to a set of 97 motif families defined earlier (Macisaac et al., 2006) and to an updated JASPAR motif database (579 annotated motifs) (Khan et al., 2018) revealed the following about each motif cluster: Cluster 1 encompasses motifs for nuclear receptors and related factors; cluster 2 includes motifs for hepatocyte-enriched nuclear factors, HNF1, HNF6/CUX and CEBP; cluster 3 includes FOX and IRF motifs; cluster 4 contains SP1 and related GC-rich motifs; and cluster 5 contains various CA-rich (zinc-finger) motifs. Enrichment analysis revealed three patterns of response (Fig. 2.8): group 1 motifs were enriched at DHS that open but not at DHS that close, group 2 motifs were enriched at both opening and closing DHS, and group 3 motifs showed greater enrichment in the sets of 27-h Δ DHS than in the 3-h DHS that open or close. The strictly opening motifs (group 1) were exclusively comprised of cluster 1 (NR half-site and DR-4 like) and cluster 2 (hepatocyte nuclear factor) motifs, which suggests a role for their cognate factors in binding directly to DHS that are opened by TCPOBOP. Cluster 2 motifs showed very modest enrichments (ES: 1.4-1.6), and 34-47% of DHS that open containing these motifs (Fig. S2.12A). The motifs enriched at both the early and the late ΔDHS (group 2 and group 3 motifs) were comprised of cluster 4 (GC-rich) and cluster 5 (GA-rich) motifs. Between 25-55% of DHS that open contained cluster 4 (GC-rich) motifs, while 60-87% of these Δ DHS contained cluster 5 (GA-rich) motifs (Fig. S2.12B). Notably, the stringent CAR-free opening DHS (Fig. 2.7A) showed stronger enrichment than the robust CARbound DHS for motifs for liver-enriched factors in cluster 2 and cluster 3, notably HNF1, HNF6/CUX, and FOX-like motifs (Fig. S2.13). This finding supports the proposal that CAR-free opening DHS preferentially bind liver-enriched transcription factors, which may facilitate cooperative interactions with nearby robust CAR-bound DHS and enhance liver-specific transcription of their TCPOBOP-inducible CAR target genes, as shown in the model in Fig. 2.7D. The FOX factor-like motifs showed the strongest enrichment in all sets of ΔDHS (ES = 8 to 13; Fig. 2.8), and are present in a large majority (71-77%) of DHS that open (Fig. S2.12A, Fig. S2.13A). Conceivably, FOX transcription factors,

which have pioneer activity (Iwafuchi-Doi et al., 2016), may cooperate with TCPOBOPactivated CAR to induce chromatin opening at these DHS.

2.5 Discussion

TCPOBOP induces widespread changes in gene expression that impact hepatic xenobiotic and lipid metabolism and are associated with dysregulation of cell cycle and DNA replication genes linked to liver tumorigenesis. Many gene responses occur within 3-h of TCPOBOP exposure, however, by 27-h there is a large expansion in the number of responding genes (Lodato et al., 2017) and their associated open chromatin regions (Lodato et al., 2018). Little is known about TCPOBOP effects on transcription factor binding or the epigenetic landscape at these putative regulatory regions, or their relationship to TCPOBOP-stimulated transcriptional responses. Here, we show that accessible chromatin regions responding to TCPOBOP exposure, identified as either opening or closing DNase-hypersensitive sites (collectively, ΔDHS) are strongly enriched for chromatin marks indicative of active enhancers and promoters. CAR binding was strongly enriched at opening DHS and their associated TCPOBOP-responsive target genes, which links CAR binding to the epigenetic and transcriptional changes induced by TCPOBOP exposure. However, CAR binding was also found at many DHS that are constitutively open in mouse liver. Moreover, a subset of opening DHS, and many closing DHS, do not bind CAR, and are deficient in CAR motifs, indicating an indirect mechanism for the TCPOBOP-induced changes in chromatin accessibility at those sites. We also saw strong enrichments between TCPOBOP-responsive ΔDHS and TCPOBOPinduced binding sites for the CAR heterodimerization partner RXRA, and in the case of

late (27-h) but not early (3-h) TCPOBOP-responsive Δ DHS, strong enrichment for TCPOBOP-induced CEBPA and CEBPB binding, whose motifs are associated with RXR-free binding of CAR (Tian et al., 2018). Δ DHS sequences were also enriched for motifs for several other liver transcription factors, which may facilitate CAR-induced chromatin opening and/or serve as co-factors for CAR-stimulated gene transcription.

We identified two major classes of DHS whose chromatin accessibility increases in TCPOBOP-exposed liver: sustained early ΔDHS , where chromatin accessibility increases within 3-h of TCPOBOP treatment and persists at 27-h; and late ΔDHS , where the change in accessibility is delayed until 27-h. At a third, less robust class of ΔDHS , chromatin opening seen at 3-h was not sustained at 27-h, despite the persistence of TCPOBOP exposure ($t_{1/2} \sim 16$ days) (Poland et al., 1980). CAR binding was highly enriched at both early and late ΔDHS ; however, some early ΔDHS , and many late ΔDHS , do not bind CAR, indicating that the change in chromatin accessibility at those sites is a secondary response to TCPOBOP-induced CAR activation. The responsiveness of both early and late ΔDHS to TCPOBOP, and their significant enrichment for CAR binding, CAR motifs and nearby TCPOBOP-responsive genes, lead us to propose these DHS include many functional, TCPOBOP-activated enhancers. Indeed, 80% of the ~2,660 TCPOBOP-responsive Δ DHS that we identified either have or acquire H3K27ac, a chromatin mark closely associated with active enhancers and promoters. However, only a small fraction of DHS with such marks are typically active, functional enhancers (Barakat et al., 2018; Grossman et al., 2017; Wang et al., 2018; Xie et al., 2017). Further

study will be required to determine which Δ DHS actually enhance CAR-regulated transcription in vivo, and to identify epigenetic marks and/or sequence-based features that distinguish them from what is anticipated to be a much larger set of functionally inactive Δ DHS sequences (Fu et al., 2018; Maricque et al., 2018). Such studies take on added importance in view of the major role that active enhancer DHS play as drivers of genetically-based human disease (Chatterjee and Ahituv, 2017).

TCPOBOP-activated CAR forms a CAR-RXR heterodimer that preferentially binds to DNA sequences comprised of a direct repeat of two AG(G/T)TCA hexamers with 4 spacing nucleotides (DR4 motif) (Frank et al., 2003; Niu et al., 2018; Tian et al., 2018). Consistent with our conclusion that CAR binds directly to TCPOBOP-responsive ΔDHS , we discovered a DR4 motif comprised of two well-defined hexamer half-sites by de novo motif analysis of the 3-h Δ DHS sequences. This motif was significantly enriched in TCPOBOP-opened chromatin regions, but not in TCPOBOP-closed regions, when compared to a background set of TCPOBOP-unresponsive DHS. The DR4 motif also yielded the strongest enrichments when compared to variants in orientation and spacing of the canonical NR half-site motif. Greater enrichment of the DR4 motif was seen at sustained early compared to late ΔDHS sequences, consistent with late DHS opening encompassing secondary regulatory events that do not involve direct CAR binding. CAR binding also showed much greater enrichment at sustained early than at late ΔDHS , as well as greater enrichment at genes induced at 3-h than after 27-h TCPOBOP exposure. Together, these findings support a model for the direct binding of TCPOBOP-activated

CAR in regulating chromatin accessibility, most prominently at sustained early ΔDHS regions.

Motif analysis identified three distinct groups of potential CAR cofactors enriched at Δ DHS compared to a background set of constitutively open DHS. One motif group was defined by its strong enrichment in opening but not closing DHS, and includes motifs for hepatocyte nuclear factors HNF1 and HNF6, CEBPs, and the DR4-like and NR half-site motifs, suggesting these factors cooperate with CAR to regulate downstream responses. Consistent with this, we found that TCPOBOP induces CEBPB binding at a subset of late opening DHS and suppresses CEBPB binding at a ~5-fold smaller subset of closing DHS. CEBP factors may play a role in chromatin remodeling or other processes associated with the large expansion of ΔDHS , and of gene responses, between 3-h and 27-h after TCPOBOP exposure. Further, CEBPA and HNF4A synergize with CAR to activate CYP2B6, an important drug-metabolizing enzyme in human liver (Benet et al., 2010). The other two motif groups were enriched at both opening and closing DHS; these include motifs for FOX factors, SP1 and other factors that bind GC-rich sequences, and zinc-finger proteins with GA-rich motifs. FOX transcription factors are pioneer factors, and are thought to actively displace nucleosomes at liver-specific enhancers, which enables cooperative binding of CEBPs and other hepatocyte nuclear factors (Iwafuchi-Doi et al., 2016).

CAR binding was frequently associated with chromatin opening at DHS that showed increases in activating histone marks, primarily H3K27ac, and much less frequently H3K4me1 and H3K4me3, following TCPOBOP exposure. TCPOBOP-induced chromatin opening was rarely associated with loss of the repressive histone mark H3K27me3. The specific mechanisms for TCPOBOP/CAR-induced chromatin opening and the associated epigenetic changes have not been determined, but could involve nucleosome redistribution (Sexton et al., 2014), DNA looping away from histone proteins (He et al., 2012), actions of the SWI/SNF chromatin remodeling complex (Kadoch et al., 2016), and recruitment of pioneer transcription factors (Swinstead et al., 2016). TCPOBOP-activated CAR was also bound at many DHS that are constitutively open prior to CAR activation (static DHS), of which 85% were already active enhancer DHS or active promoter DHS prior to TCPOBOP exposure, based on their constitutive histone marks (Table_S5D). Presumably, CAR binding to these open genomic regions occurs without the need for extensive chromatin remodeling and is determined by the level of active CAR in the nucleus, which increases strongly following TCPOBOP exposure. CAR binding outside of DHS regions was observed, but was significantly weaker than binding at DHS (Fig. S2.8A).

TCPOBOP stimulated chromatin opening at many sites that do not bind CAR. Examples include three of the six Δ DHS upstream of *Cyp2b10*, of which one is at the *Cyp2b10* promoter, and the other two are flanked by opening DHS that do bind CAR (Fig. 2.5A). All six opening DHS are sustained early Δ DHS showing strong increases in H3K27ac

marks within 3-h of TCPOBOP exposure; thus, they are early TCPOBOP-activated enhancers, consistent with the early, robust increase in Cyp2b10 transcription (Lodato et al., 2017). The presence of both CAR-free and CAR-bound opening DHS nearby the same TCPOBOP-inducible gene, as seen for Cyp2b10 (Fig. 2.5A), was common, as this pattern was also seen for 70 out of 84 stringently CAR-free opening DHS that map to a TCPOBOP-inducible gene. Conceivably, the CAR-free opening DHS may cooperate with nearby CAR-bound opening DHS to stimulate TCPOBOP-induced transcriptional responses (model, Fig. 2.7D). Motif analysis revealed that the CAR-free opening DHS were depleted of CAR DR4 and NR half-site motifs, as expected, but additionally, showed motif enrichment for several liver-enriched transcription factors, notably HNF1 and HNF6, which was not seen at CAR-bound opening DHS. Conceivably, factors such as HNF1 and HNF6 may preferentially interact with the CAR-free opening DHS to increase liver-specificity of transcription via interactions with the nearby CAR-bound opening DHS (Fig. 2.7D). The mechanism whereby TCPOBOP opens the CAR-free DHS in the absence of direct CAR binding is unknown.

TCPOBOP stimulated chromatin closing at many genomic sites. Chromatin closing was enriched for loss of H3K27ac activating marks, but was rarely accompanied by an increase in H3K27me3 repressive marks. While closing DHS show significant enrichment for mapping to genes repressed by TCPOBOP, closing DHS were not enriched for CAR binding. Given the high specificity of TCPOBOP for binding to CAR, it is likely that TCPOBOP-stimulated chromatin closing is dependent on CAR action, albeit by an indirect mechanism that is initiated as a secondary response to TCPOBOP exposure. This notion is consistent with the delay in TCPOBOP-induced chromatin closing as compared to chromatin opening described earlier (Lodato et al., 2018). However, we also observed CAR binding at many DHS that mapped to TCPOBOPrepressed genes; 85% of these DHS are static enhancer or static promoter DHS (Table_S5D). This association of static, i.e., constitutively open enhancer and promoter DHS with CAR binding sites that map to TCPOBOP-repressed genes is consistent with the notion that gene repression results from competitive displacement by CAR of more strongly activating transcription factor(s), such as the nuclear receptors HNF4 and PPARA, whose binding sites frequently overlap those of CAR (Tian et al., 2018). These findings suggest two independent mechanisms for gene repression by CAR: 1) TCPOBOP-stimulated chromatin closing, which does not involve direct CAR binding; and 2) competition between TCPOBOP-activated CAR and other nuclear receptors for binding to common binding sites at DHS that are constitutively open in mouse liver.

Although TCPOBOP-induced changes in enhancer marks were strongly enriched at both opening and closing DHS (Fig. 2.2D), only ~25-37% of the full sets of opening and closing DHS showed such TCPOBOP-induced changes in histone marks (Fig. S2.5). Furthermore, opening DHS without chromatin mark changes were significantly enriched for CAR binding and for target gene induction, albeit to a lesser extent than Δ DHS that did show chromatin mark changes. Similar patterns were seen for repressed enhancers mapping to repressed genes (Fig. S2.5DE). Thus, changes in the classical H3K27ac

enhancer mark are not an obligatory, mechanistic driver of chromatin opening or closing at TCPOBOP-responsive enhancer DHS. Furthermore, as noted, changes in chromatin opening and closing cannot be explained by a loss or gain in H3K27me3 (repressive) marks, respectively, as such changes were rarely found at TCPOBOP-responsive DHS.

Two distinct sets of TCPOBOP-responsive genes were recently identified based on a differential requirement of liver growth factor signaling for their TCPOBOP-stimulated gene responses (Bhushan et al., 2018). Specifically, induction or repression of ~40% of all TCPOBOP-responsive genes, including many genes active in DNA replication and cell cycle regulation, but not the major TCPOBOP-responsive Cyp or other drugmetabolizing enzyme genes, was blocked upon inactivation of the functionally redundant (Paranjpe et al., 2016) liver growth factor receptors MET and EGFR (Bhushan et al., 2018). Thus, MET and/or EGFR signaling is specifically required for CAR-activated transcription of genes such as TCPOBOP-responsive transcription factors implicated in hepatocyte proliferation (Delgado et al., 2011; Hu et al., 2014; Yang et al., 2015; Yu et al., 2016). Given the delayed response to TCPOBOP that is seen for many of the DNA replication/cell cycle genes required for induction of hepatocyte proliferation (Lodato et al., 2017), the findings of (Bhushan et al., 2018) had suggested that induction of the proliferation response is an indirect response to CAR activation, perhaps a secondary response downstream of CAR activation of MET and/or EGFR signaling (Bhushan et al., 2018). Our findings strongly suggest, however, that many TCPOBOP-induced proliferation genes are direct targets of CAR, as indicated by CAR ChIP-seq binding to

proximal DHS. Further study is needed to identify the precise role of CAR and how it interfaces with MET/EGFR signaling to increase TCPOBOP-stimulated transcription, which may give unique insights into the proliferative actions of CAR linked to hepatocellular carcinoma.

2.6 Conclusions

We used the model environmental chemical TCPOBOP, a highly specific agonist ligand of CAR, to investigate early effects of xenobiotic exposure on the epigenome and chromatin states in mouse liver. This study highlights the need to better understand the complexity of gene regulatory mechanisms in the context of chromatin accessibility, histone marks, and transcription factor binding. Global epigenetic maps were obtained for activating (H3K27ac, H3K4me1, H3K4me3) and repressive histone marks (H3K27me3), and TCPOBOP-responsive binding sites for several liver transcription factors were identified and used to characterize changes to the liver epigenome and chromatin states following TCPOBOP exposure. We found that exposure to TCPOBOP substantially alters the enhancer landscape of mouse liver, inducing time-dependent chromatin opening or closing at several thousand sites identified as differential DNase-hypersensitive sites (DHS) with enhancer marks, many of which map to TCPOBOP-responsive genes. DHS opened by TCPOBOP were significantly enriched for CAR binding and for CAR DR4 motifs, indicating CAR is likely to bind directly to these enhancer DHS sequences. We also found that a subset of opening DHS proximal to the CAR-bound enhancer DHS, as well as many closing DHS mapping to TCPOBOP-responsive genes, did not bind CAR, evidencing an indirect mechanism for those changes in chromatin accessibility. These

findings directly link TCPOBOP exposure to widespread changes in chromatin accessibility associated with CAR binding and altered gene transcription. The global maps of dysregulated sites of chromatin accessibility and other epigenetic changes described here constitute a rich resource for further research on environmental chemical effects on the epigenome.

Fig. 2.1 TCPOBOP-stimulated chromatin opening and closing.

A-Overlap between TCPOBOP-opened DHS (Δ DHS) at 3-h and 27-h in male (top) and in female (bottom) liver, defining Δ DHS as either transient, sustained early, or late responding (Table_S4, columns T, U). **B**-Enrichment of Δ DHS for mapping to TCPOBOP induced genes in the same TAD compared to static DHS. A single target gene was identified for each DHS as the closest RefSeq or multi-exonic lncRNA TSS in the same TAD as the DHS. **C**-Distribution of fold-change for chromatin opening (TCPOBOP-exposed/vehicle- treated) calculated from RiPPM-normalized DNase-seq read counts for the indicated Δ DHS sets, including sustained early Δ DHS at both time points. Wilcoxon rank-sum test with Benjamini–Hochberg p-value adjustment: **, P<1e-03; ***, P<1e-10 for sustained early, late Δ DHS vs transient Δ DHS; &&, P<1e-03 between Δ DHS time points. **D**-MA plots showing log2 fold-change in DHS signal intensity for TCPOBOP-exposed vs. control liver (y-axis) vs. peak width-normalized DHS signal (average RiPPM-normalized reads for TCPOBOP and control samples; xaxis) for each Δ DHS set. Colors distinguish data points for indicated Δ DHS sets.





Fig. 2.2 Enhancer and promoter DHS classes.

A-Number of significantly induced and repressed sites for activating histone marks, after 3-h or 27-h TCPOBOP, by ChIP-seq. B-Enrichment of TCPOBOP-induced histone marks at DHS that open or close compared to background of static histone marks. Bar height: ES score; bar color: Fisher Exact test p-value for enrichment; numbers above bars: percentage of ∆mark sites at each DHS set. C-Numbers of Enhancer DHS and Promoter DHS with the indicated histone mark responses. Last row: percentages of all Enhancer + Promoter DHS. D,E-DHS Enhancer (D) and promoter (E) subclasses shown in (C) separated into sets of Δ DHS that close, open, or were unchanged (static) by TCPOBOP. Above bars: numbers of DHS in the set, and their significant enrichment scores (ES), e.g., for 587 enhancer DHS with induced histone marks, ES = 22.3 (p=0) compared to corresponding DHS class with static histone mark (see Fig_S5B-S5E). F-Box plots showing log2 fold expression ratios for target genes associated with the indicated DHS subsets (based on average RNA-seq gene expression ratios for 3-h, 27-h TCPOBOP-exposed vs. control male livers). Wilcoxon rank-sum test with Benjamini-Hochberg p-value adjustment: *, P<0.05; **, P<1e-03; ***, P<1e-20.



Fig. 2.3 DR4 and half-site motifs in ΔDHS sequences.

A-De-novo motif discovery for Δ DHS that open in male (M) and female (F) liver identified DR4-like motifs (3-h TCPOBOP Δ DHS) and NR half-site motifs (27-h TCPOBOP Δ DHS), very similar to the canonical CAR motif (motif 65, top) (Macisaac et al., 2006). **B-E:** Enrichments of each Δ DHS set for containing motif 65 as compared to a background set of static DHS (sDHS). Above each bar: % Δ DHS with CAR motif; dashed horizontal line: CAR motif in static DHS (50% frequency). Fisher Exact test with Benjamini–Hochberg p-value adjustment for enrichments: *, P<0.01; **, P<1e-10; ***, P<1e-20. **F**-ES scores for six Δ DHS sets, indicated at top, for DR, ER and IR motifs constructed from de novo-discovered female liver NR half-site motif AGGTCA with spacing from 0-5 bp between half-sites (Fig_S6B). Shown are data for 7 constructed motifs with at least one ES≥2. Also see Fig_S6C. **G,H**-ROC curves for constructed motifs for opening DHS in 3-h, 27-h TCPOBOP-exposed male liver vs static DHS background set (Fig_S7).





Fig. 2.4Enrichment of transcription factor binding at TCPOBOP-responsiveDHS and genes.

Shown are ES (bar heights), Fisher Exact test significance (log p-values, color bar), and % responding DHS or genes (values above bar). Dashed line, ES=1. A-Enrichments of TCPOBOP-induced opening DHS for mouse CAR binding sites. Values above bars: % opening DHS with a CAR binding site, based on data from (Niu et al., 2018) for mouse CAR activated by TCPOBOP or phenobarbital. No enrichments were seen for closing DHS (not shown). B-Enrichments of genes induced by TCPOBOP in male liver for proximal mouse CAR binding (CAR binding sites mapped to closest RefSeq or multiexonic lncRNA in TAD (CAR target genes). Background for enrichments: TCPBOPunresponsive genes. Values above bars: % TCOPOBOP-induced genes that are CAR targets. C-Enrichments of TCPOBOP-responsive binding sites for each transcription factors for being at an opening or closing DHS. TCPOBOP-induced and repressed transcription factor binding sites (ΔTF) were discovered in male liver for each factor at both time points. Values inside bars: number of TCPOBOP-responsive binding sites; values above bars: % transcription factor binding sites at DHS that open or close. TCPOBOP induced binding was enriched at opening DHS; TCPOBOP-repressed binding was enriched at closing DHS. D-Enrichments of TCPOBOP-responsive DHS for containing TCPOBOP-responsive transcription factor binding sites. Values inside bars: number of TCPOBOP-responsive DHS; values above bars: % DHS with induced or repressed binding sites for DHS that open or close.









Fig. 2.5ChIP-seq binding sites for mCAR, hCAR, RXR, and active histonemarks at ΔDHS regions of TCPOBOP-inducible genes.

UCSC Brower tracks for male mouse liver showing (top to bottom): DHS peaks in unexposed liver (DHS) and 27-h after TCPOBOP (DHS/TCPO). Horizontal bars beneath DHS/TCPO track: DHS peak regions that were induced (red) or unchanged (black) by TCPOBOP; gene structures for RefSeq (purple) and liver-expressed lncRNAs (green); ChIP-seq binding sites for mCAR (TCPOBOP), hCAR (CITCO), based on (Niu et al., 2018); and ChIP-seq peaks for RXR binding and for histone marks H3K4me3 and H3K27ac, both without and with 27-h TCPOBOP exposure. Screen shots are shown for these TCPOBOP-inducible genes: A-Cyp2b10, lnc_5998 (2 isoforms). Only 3 of 6 Δ DHS upstream of Cyp2b10 bind CAR; B-lnc_13509, with 15 liver-expressed isoforms; C-Ces2a, anti-sense lnc_7359 (3 isoforms); D-lnc_5335 (7 isoforms). Antxr1 is antisense to this lncRNA, but is not TCPOBOP-induced; E-lnc_8460 (3 isoforms) is antisense to Ginm1, a modestly TCPOBOP-inducible gene; F-Por and the anti-sense lnc_4655 (2 isoforms).

Fig. 2.5

A26.07000 28.875,000 28.880,000 28.885,000 26.890,000 Chr7	50 ml B340.000 22345.000 23356.000 23365.000 23365.000 23365.000 233 Chr16	C 107,280,000 107,285,000 107,270,000
DHS		
DHS/TCPO		
Cyp2b10	1810053B23Rik	← Ces2a →
Curc 5998 Cyp2b10 →		0w28
mCAR		€ Lnc 7359
hCAR		
RXR		
RXR/TCPO		
K4me3		
K4mo2/TCDO		
K27ac		
K27ac/TCPO		
Date: 1	10 M	20 16
87,200.pc0 Chr6	7,425000 7,505,000 7,505,000 Chr10	138,160,000 138,170,000 Chr5
D DHS Chr6 E	7,80500 7,500,000 7,500,000 Chr10	F
D DHS Chr6 E	7-445001 7500,000 7500,000 Chr10	F And
D DHS DHS/TCPO	2-445001 (2000)	
D DHS DHS Chr6 E Chr6 E Chr6	2×≈5001 / 2×50000 / 2×50000 Chr10	
D DHS DHS/TCPO ← Antxr1 noRNA_as_chr6_5335 → (7 forms) mCAR	yaasool (200,000 chr10	138,140,000 38,170,000 Chr5 F 138,170,000 Max. ← Lnc 4655 Por → year ender rd594, sec.sch, 4800 2 Pointed Chr4
DHS DHS/TCPO Chr6 E DHS/TCPO Chr6 E M Chr6	2445001 (200,000 Chr10 Common Chr10 Common Chr10 Common Chr10 Common Chr10 Common Chr10	F (32, 100,000) (32, 100,000) Chr5 F (32, 100,000) Chr5 ← Lnc 46555 Por → (32, 100,000) (32, 100,
Chr6 D DHS Chr6 E DHS/TCPO Chr6 E Chr6 C Chr6 C Chr6 C C Chr6 C C C C C C C C C C C C C	2480.001 (200.000 Chr10 (200.000 Chr10 (200.	F (Sk, 100,000) Chr5 F (Sk, 100,000) Chr5 ← Lnc 46555 Por → Several Chr5 (Sk) (Sk, 100,000) 2 (Sk, 100,000) (Sk) (Sk) (Sk) (Sk) (Sk) (Sk) (Sk) (Sk
Chr6 D DHS Chr6 E DHS/TCPO ← Antxr1 noRNA_ss_chr0_5335 → (7 forms) mCAR hCAR RXR RXR RXR/TCPO	yeasool (200,000 chr10 c Ginm1 ncRNA, as chr10 8480(3 forms) Lnc 8460 →	(32,100,000 (32,1000
Chr6 D DHS DHS/TCPO ← Antxr1 icRNA_ss_chr0_5335 → (7 forms) mCAR hCAR RXR RXR RXR/TCPO K4me3	2×45001 (2×00.000 Chr10 ← Ginm1 noRNA, as chr10 8480(3 forms) Linc 8460 →	F (32, 100,000) (32, 100,000) Chr5 F (32, 100,000) (32,
Chr6	2×450 cl 2×450 cl 2×50 cl	F (36, 160, 200) (36, 170, 200) Chr5 F (10, 16, 10, 200) (10, 10, 200
D DHS Chr6 D DHS I DHS/TCPO ····································	2∞0.00 200.00 Chr10 ← Gim1 ncRNA as chr10 8400(3 forms) Linc 8460 →	
Chr6 D DHS DHS/TCPO ← Antxr1	2 495,001 (200,000 Chr10 ← Ginm1 ncRNA, as chr10 8480(3 forms) Linc 8460 →	F (20, 100,000) (20, 100,000) Chr5 F (20, 100,000) (20,

Fig. 2.6mCAR and hCAR binding sites and ΔDHS regions nearbyTCPOBOP-inducible genes involved in DNA synthesis and cell cycle regulation.

UCSC Browser screen shots, as in Fig. 5, showing five cell cycle/DNA replication genes that require MET and EGFR signaling for TCPOBOP induction, and also bind CAR, most often at TCPOBOP-induced opening DHS (red, orange bars beneath gene track). This binding indicates these genes are direct CAR targets. Inset beneath D: two TCPOBOP-inducible Δ DHS far downstream (~100 kb) of Gadd45a; bind both mCAR and hCAR.

Fig. 2.6



Fig. 2.7 CAR binding to opening and to static DHS.

A-Model showing four DHS sets: TCPOBOP-induced opening DHS and static DHS that are stringently CAR-free or show robust CAR binding (see Methods). B-Left: Boxplots of magnitude of TCPOBOP-induced DHS opening (fold-change DHS signal) for each DHS set, based on RiPPM-normalized DNase-seq reads for TCPOBOP/control male liver after 27-h TCPOBOP. Right: Boxplots showing normalized DNase-seq signal for the DHS subsets. The four DHS groups are designated A-D (see panel A). C-% DHS with NR half-site and DR4 motifs. Motifs scans by FIMO using female 3-h DR4 and NR half site motifs (Fig. 3A). X-axis: number of motif occurrences in each DHS (n=0-6). Y-axis: % DHS in each group with corresponding number of motifs. Stringent CAR-free static DHS showed highest percentage without a motif, followed by CAR-free opening DHS. Group B showed similar distribution of NR half-site motifs as group D, but fewer DR4 motifs: 51% 0 and 1 motif occurrences and 23% 3-5 motif occurrences for group B (not shown) vs. 36.5% and 34%, respectively, for group D. D-Model for TCPOBOP-activated CAR inducing chromatin opening, both at a DHS with a DR4 motif, where CAR binds strongly, and at nearby DHS which are stringently CAR-free but are enriched for binding sites for several liver-enriched transcription factors (see text), which may cooperate with the nearby CAR-bound sites to induce target gene expression, as shown at bottom.



Fig. 2.7

Fig. 2.8Enrichment of de novo-discovered motifs at TCPOBOP-responsiveΔDHS.

Shown are ES (values in each cell) and Fisher Exact test significance log p-value (color depth) for each Δ DHS set. 21 significant motifs were grouped into 5 distinct clusters based on sequence similarity (Fig. S2.11); the cluster number (1-5) is indicated for each motif. The motif name indicates the factor name and the corresponding dataset use for motif discovery; male 3-h (M30), 27-h (M270), female 3-h (F30) and 27-h (F270) DHS that open. Example: CAR_F270 refers to CAR motif discovered from DHS that open in female liver after 27-h TCPOBOP. Hierarchical clustering identified motifs that showed similar patterns of enrichment across Δ DHS sets and were grouped accordingly (Groups 1-3). Group 1 motifs, enriched at opening but not at closing DHS, are composed of NR half-site, DR4-like, and HNF motifs (cluster 1 and 2 motifs). Groups 2 and 3: motifs strongly enriched in DHS that open, close after 27-h TCPOBOP.

er		DHS that open			DHS that close					
	Clust	Motif	Male 3 h	Male 27 h	Female 3 h	Female 27 h	Male 3 h	Male 27 h	Female 3 h	Female 27 h
Group 1	2	CEBP	1.5	1.4	1.5	1.4	0.8	1	0.9	1.1
	1	CAR	2.3	2.4	3.5	2	0.8	1	0.8	0.9
	2	HNF1A	1.1	1.4	1.3	1.5	1	1.1	0.7	0.8
	2	HNF6	1.4	1.5	1.2	1.6	1.1	1.1	1.6	0.8
Group 2	5	RXRA	3.7	2.3	2.9	2.4	2.4	2.7	3.6	3.6
	5	NR4A2	5.4	3.3	4.9	3.6	2.8	4.2	3.4	4.7
	3	RREB1	3	3.8	3.1	3.7	3	2.8	2.3	2.8
	5	GATA1	3.6	2.8	2.5	2.7	3.1	3.8	4.7	3.6
	4	TFAP2A	3.3	2.2	2.2	2.3	3	3.4	4	3.4
Group3	4	SP1	5	3	3.5	3.4	3.1	4.5	6.8	4.5
	5	KLF4	8.1	5.5	5.8	5.5	7.7	6.2	9.4	6
	5	AZF1	7.2	8.7	7.3	8.2	7.2	6.4	4.7	5.8
	5	ELF-1	7	4.3	4.3	4.5	7.4	6.4	8.2	6.3
	3	FOX	13	12.8	12.2	12.1	8.6	10.4	8.1	10.4

Adj p-val < E-100 E-50-100 E-5-50 > E-05 79

Fig. 2.8

Fig. S2.1 TCPOBOP opened DHS and induced genes in male and female liver

A, Enrichment of transient, sustained early, and late Δ DHS for mapping to TCPOBOP induced genes in the same TAD as compared to static DHS in female liver, as shown for male liver in Fig. 1B. A single target gene was identified for each DHS as the closest RefSeq or multi-exonic lncRNA TSS in the same TAD as the DHS. Up to 47% of DHS that open mapped to a TCPOBOP-induced gene, with an enrichment score of 17.8 compared to static DHS.

B, C, Transient, sustained early, and late Δ DHS that open sub-grouped by whether the DHS showed a standard or robust response (see Methods) to TCPOBOP exposure in male (B) or female (C) liver. A larger percentage of sustained early and late Δ DHS showed a robust response, indicating they are generally of better quality than the transient Δ DHS.

D, E, Sex-specific and sex-independent (common) Δ DHS that open sub-grouped by whether they show a standard or robust response after 3 h (D) or 27 h (E) TCPOBOP exposure. Although similar numbers of Δ DHS were responsive to the early exposure (300 Δ DHS in male and 340 Δ DHS in female, Fig. S2.3B), the vast majority of male-specific sites showed a standard response compared to the majority of female-specific Δ DHS showing a robust response (Fig. S2.1C). This indicates that female mice show a stronger response to TCPOBOP exposure at an early time point as compared to male mice. Many more sex-specific sites were found in the later time point exposure with less disparity between standard and robust responding Δ DHS (Fig. S2.1D)

Fig.	S2.1
------	------

A. ∆DHS and Differentially expressed genes (DEGs) in female liver	% ∆ DHS mapped	Enrichm ent Score	p-Value
Transient ΔDHS mapping to 3 h up genes	17	16.7	3.72E-21
Early, sustained ΔDHS mapping to 3 h up genes	29	28.5	1.35E-75
Early, sustained Δ DHS mapping to 27 h up genes	47	17.8	9.90E-84
Late ∆DHS mapping to 27 h up genes	25	9.6	2.86E-209

Percent (and count) of ΔDHS that show either "standard" or "robust" response









Fig. S2.2 Comparisons of DHS width between subsets of opening ΔDHS regions

A, Boxplots showing the distribution of DHS width for the transient, sustained early, and late Δ DHS that open in male (left) and female liver (right). Transient and sustained early Δ DHS were significantly wider than the late opening Δ DHS.

B, Boxplots showing the distribution of DHS width for the sex-specific and sexindependent (common) Δ DHS that open after 3 h (left) or 27 h (right) TCPOBOP exposure. At 27 h TCPOBOP exposure, sex-independent and female-specific Δ DHS were significantly wider than the male-specific Δ DHS.





Fig. S2.2

Fig. S2.3 TCPOBOP exposure induces sex-differences in chromatin accessibility

A, Venn diagrams showing the overlap between TCPOBOP-opened DHS (Δ DHS) in male compared to female mouse liver after 3 h (left) and 27 h TCPOBOP exposure (right), defining Δ DHS as either male-specific, sex-independent, or female-specific DHS. Also see Fig. 2.1A.

B, Boxplots showing the distribution of fold-change for chromatin opening (TCPOBOPexposed/vehicle- treated) calculated from RiPPM normalized DNase-seq read counts for Δ DHS shown in (A). A Wilcoxon rank-sum test with Benjamini–Hochberg p-value adjustment was used to determine significant differences between distributions for sexindependent (common Δ DHS) and female-specific Δ DHS vs male-specific Δ DHS (*P < 0.05, **P < 1e-03, ***P < 1e-10). Comparisons were also made between sexindependent Δ DHS at 3h versus 27 h, as indicated. Δ DHS common to both sexes showed greater increases in DHS intensity in female than in male liver at both TCPOBOP time points.

C, MA plots (as shown in Fig. 2.1D) showing the fold-change in DHS signal intensity vs. normalized DHS signal (average RiPPM read counts for TCPOBOP-exposed and vehicle-treated liver DHS) for Δ DHS after 3 h (left) and 27 h (right) TCPOBOP exposure in male (top) and female (bottom) liver. Shown are the Δ DHS that open (green) or close (red) in common in both sexes, and also Δ DHS that are sex-specific (i.e. TCPOBOP-respondence) responsive DHS in either male or female liver) (black). Δ DHS that respond in a sex-

specific manner (black dots) showed a wide-range of responsiveness and degree of openness at both time points.

86



Log2 (normalized DHS Signal)
Fig. S2.4Enrichment of ΔDHS vs. static DHS for containing activating
chromatin marks

A, Number of genomic sites in male liver showing a significant increase (opening DHS) or decrease (closing DHS) in DNase-seq signal intensity (Δ DHS), after 3 h or 27 h TCPOBOP exposure.

B, Enrichment of DHS that open (top) or DHS that close (bottom) relative to static DHS for containing induced or repressed activating chromatin marks (H3K27ac, H3K4me1, or H3K4me3) after 3 or 27 h TCPOBOP exposure. Bar height indicates magnitude of enrichment, and bar color indicates the Fisher Exact test p-value for the enrichment. Listed above each bar is the percentage of Δ DHS that contain the associated Δ mark sites. The enrichments of Δ DHS for sites where TCPOBOP increased H3K27ac marks was seen at 3 h (ES = 15.1, p = 2.79E-42) and at 27 h (ES = 70.8, p = 0). DHS that open at 27 h were also enriched for induced H3K4me1 marks (ES = 34.1, p = 1.37E-30) and for induced H3K4me3 marks (ES = 135.7, p = 1.39E-22). Further, DHS that close at 27 h showed strong enrichment for loss of H3K27ac marks (ES = 18, p = 3.48E-52).

88

Α	TCPOBOP treatment	Number of DHS that open	Number of DHS that close	
	Male 3 h	472	72	
	Male 27 h	1767	564	



Enrichment of ΔDHS vs. sDHS for containing activating chromatin marks

Fig. S2.5 Identification and enrichment of enhancer and promoter DHS

A, Flowchart showing the classification of the 60,739 DHS used in this study into either enhancer or promoter DHS and respective subclasses as shown in Fig. 2.2C. Several sets of DHS were filtered out; such as those that lack active chromatin marks and those that showed a conflicting response to TCPOBOP exposure (see Methods).

B, C, Enrichments of induced, poised, and repressed enhancer DHS (B) or promoter DHS (C) relative to the respective set of static DHS for being at DHS that open or DHS that close following TCPOBOP exposure. Shown are the enrichment scores (ES), Fisher Exact test p-value, and the number and percentage of enhancer or promoter DHS that overlap a Δ DHS. Induced and repressed enhancer and promoter DHS were enriched for DHS that open and close respectively.

D, E, Enrichments of induced enhancers at DHS that open (D) and repressed enhancers at DHS that close (E) for mapping to TCPOBOP responsive genes. Enhancer DHS were mapped to the closest RefSeq or multi-exonic lncRNA gene TSS in the same TAD (enhancer target genes). Enrichments were calculated to evaluate the effect of either DHS opening (top), acquisition of the H3K27ac mark (middle), or both relative to the background sites indicated. Red is used to highlight the differences in how each set of enrichment scores was calculated. The combined effect of DHS opening and acquisition of the H3K27ac mark yielded the strongest enrichment (ES: 11.7, p = 5.47E-111). Similarly, the highest enrichment for TCPOBOP-repressed genes was obtained for closing DHS with a loss of activating enhancer marks (ES: 6.6, p = 9.68E-06).



Fig.	S2.5	BC

S5B. Enrichment of enhancer classes at ΔDHS vs. static enhancer DHS	Enrichment compared to static enhancers (30,7 sites) at the same set of dDHS:		
Enhancer class (total number of sites)	∆DHS that open ES (p-value)	∆DHS that close ES (p-value)	
Induced (1,603 sites)	22.3 (p = 0) 587 (37%)	NS 15 (1%)	
Poised (10,475 sites)	NS 314 (3%)	0.5 (1.87E-08) 64 (1%)	
Repressed (739 sites)	NS 20 (3%)	7.8 (p = 4.34E-33) 66 (9%)	

S5C. Enrichment of promoter classes at ΔDHS vs. static promoter DHS	Enrichment compared to static promoters (12,2 sites) at the same set of dDHS:		
Promoter class (total number of sites)	∆DHS that open ES (p-value)	∆DHS that close ES (p-value)	
Induced (215 sites)	32.6 (1.74E-52) 53 (25%)	7.2 (3.34E-04) 6 (3%)	
Poised (307 sites)	0 sites	NS 2 (1%)	
Repressed (156 sites)	NS 1 (1%)	7.5 (2.64E-04) 6 (4%)	

Fig. S2.5DE

S5D. Enrichment of enhancers with induced histone marks at opening DHS $(\Delta DHS / \Delta EnhMark)$: [enrichments # 1, 2, 4, below, based on 587 Δ DHS; Fig. 2D], or for enhancers with static marks at DHS that open $(\Delta DHS / EnhMark;$ enrichment #2 below, based on 779 Δ DHS; Fig. 2D) for mapping to TCPOBOP induced genes. Results indicate that TCPOBOP-induced chromatin opening and TCPOBOP-induced chromatin opening and TCPOBOP-induced contribute to the enrichment for TCPOBOP-induced genes.

Opening DHS, induced marks, genes:

- **1.** $\frac{\Delta DHS / \Delta EnhMark}{DHS / EnhMark} = ES: 11.7$ (p = 5.47E-111)
- 2. $\frac{\Delta DHS/EnhMark}{DHS/EnhMark} = ES: 3.6$ (p = 1.15E-24)
- **3.** $\frac{\Delta DHS / \Delta EnhMark}{\Delta DHS / EnhMark} = ES: 3.2 (p = 6.25E-18)$
- 4. $\frac{\Delta DHS / \Delta EnhMark}{DHS / \Delta EnhMark} = ES: 3.3$ (p = 2.81E-20)

S5E. Enrichment of repressed enhancers at Δ DHS that close (66 sites; Fig. 2D) for mapping to TCPOBOP repressed genes [enrichments#1, 2, 4, below]. Enrichment#3 is based on 382 closing DHS with static enhancer marks (Δ DHS/EnhMark; Fig. 2D). Results indicate that TCPOBOP-induced chromatin closing and TCPOBOPrepressed enhancer histone marks both contribute to the enrichment for TCPOBOP-repressed genes.

Closing DHS, repressed marks, genes:

- **1.** $\frac{\Delta DHS / \Delta EnhMark}{DHS / EnhMark} = ES: 6.6 (p = 9.68E-06)$
- 2. $\frac{\Delta DHS/EnhMark}{DHS/EnhMark} = ES: 3.6$ (p = 1.16E-09)
- **3.** $\frac{\Delta DHS / \Delta EnhMark}{\Delta DHS / EnhMark} = NS$
- 4. $\frac{\Delta DHS / \Delta EnhMark}{DHS / \Delta EnhMark} = ES: 2.3$ (p = 0.0291)

Fig S2.6 De novo and constructed motifs from TCPOBOP opened DHS sequences

A, Enrichments of each indicated Δ DHS set for containing de novo motifs (shown in Fig. 2.3A) relative to a background set of static DHS. Shown in each cell is the percentage, ratio of Δ DHS with the motif, enrichment score (ES), and Fisher Exact test p-value. Cells highlighted in red indicate the strongest enrichment calculated for each Δ DHS set. B, Diagram showing the set of 18 constructed motifs composed of direct repeat (DR), everted repeat (ER), and inverted repeat (IR) with 0 to 5 bp spacer nucleotides between hexamers. The nuclear receptor half-site motif discovered from female liver (shown in A; bottom row) was used to construct these motifs. The first position of this 7 bp motif was omitted, due to low information content, and the remaining AG(T/G)TCA sequence was used to construct the 18 motifs as shown.

C, Enrichments of each indicated Δ DHS set for containing constructed motifs (shown in B) relative to a background set of static DHS. Each cell indicates the enrichment score and Fisher Exact test p-value (cell color). The constructed DR4 motif showed the strongest enrichment in the sets of DHS that open (C1) but not DHS that close (C2). Similar patterns of results were seen for the set of transient (T), sustained early (S), and late Δ DHS (C3) and sex-independent DHS (C4). Closing DHS did not show significant enrichment (ES > 1.3 at p < 1E-03) for any of the 18 constructed motifs or for the de novo NR half-site motif.

∆DHS that open →	Male 3hr	Male 27 hr	Female 3 hr	Female 27 hr
Male DR4	61% (286/472) ES: 1.7 p =1.60E-08	57% (999/1767) ES: 1.5 p = 4.03E-15	63% (323/512) ES: 1.9 p = 1.62E-12	55% (1301/2357) ES: 1.4 p = 6.77E-15
Male half-site	74%	71%	76%	67%
	(350/472) ES: 2.5 p = 7.18E-20	(1251/1767) ES: 2.1 p = 4.55E-50	(389/512) ES: 2.8 p = 1.74E-25	(1569/2357) ES: 1.8 p = 3.95E-38
Female DR4	65%	66%	74%	61%
	(305/472) ES: 2.3 p = 5.91E-17	(1171/1767) ES: 2.4 p = 9.51E-71	(379/512) ES: 3.5 p = 1.35E-39	(1446/2357) ES: 2.0 p = 2.02E-55
Female half-site	72% (340/472) ES: 2.7 p = 3.64E-23	67% (1189/1767) ES: 2.2 p = 1.06E-52	74% (377/512) ES: 2.9 p = 6.65E-29	63% (1475/2357) ES: 1.8 p = 1.63E-38

Fig. S2.6A: Motif scan for de-novo discovered motifs







Fig. S2.6C. Heat map of enrichment scores (based on FIMO motif scan). Cells are colored by adjusted p-value.

	M (T)	M (S)	M (L)	F (T)	F (S)	F (L)
AGGTCA	1.7	2.4	1.4	2.3	2.1	1.3
DR0	1.4	1.7	1.2	1.4	1.5	1.2
DR1	1.2	1.4	1.2	1.8	1.5	1.3
DR2	1.2	1.3	1.2	1.3	1.4	1.2
DR3	1.6	1.9	1.3	1.6	1.8	1.3
DR4	1.7	4	2.1	2.3	3.8	1.5
DR5	1.3	2.1	1.2	1.5	1.7	1.3
ER0	1.4	1.4	1.2	1.4	1.4	1.2
ER1	1.3	2.1	1.5	2.3	2.2	1.3
ER2	1.3	1.3	1.2	1.1	1.3	1.1
ER3	1.3	1.5	1.2	1.6	1.4	1.3
ER4	1.5	1.7	1.1	1.4	1.4	1.1
ER5	1.2	1.7	1.3	1.8	1.7	1.2
IR0	1.5	1.2	1.2	1.5	1.3	1.1
IR1	1.3	1.7	1.2	1.9	1.5	1.2
IR2	1.3	1.7	1.3	1.5	1.7	1.3
IR3	1.1	1.6	1.3	1.8	1.3	1.3
IR4	1.3	1.7	1.3	1.7	1.5	1.2
IR5	1.4	1.7	1.3	1.8	1.4	1.2

		Sex-			Sex-indep	
	M3hr	indep 3h	F3hr	M27hr	27h	F27hr
AGGTCA	1.6	2.9	1.9	1.5	1.5	1.3
DR0	1.3	1.9	1.3	1.2	1.2	1.2
DR1	1.1	1.8	1.5	1.3	1.2	1.4
DR2	1.2	1.4	1.3	1.3	1.2	1.3
DR3	1.5	2.2	1.5	1.3	1.5	1.3
DR4	1.7	5.5	2.5	2	2.4	1.4
DR5	1.3	2.3	1.3	1.2	1.4	1.3
ER0	1.3	1.5	1.3	1.2	1.2	1.2
ER1	1.2	2.4	2.1	1.7	1.5	1.3
ER2	1.3	1.2	1.2	1.1	1.3	1.1
ER3	1.1	1.9	1.3	1.3	1.2	1.3
ER4	1.3	2.3	1.1	1.1	1.1	1.2
ER5	1.3	1.6	1.7	1.3	1.4	1.1
IR0	1.2	1.5	1.3	1.2	1.1	1.2
IR1	1.3	1.7	1.6	1.3	1.2	1.2
IR2	1.2	2	1.5	1.3	1.3	1.3
IR3	1.1	1.6	1.4	1.3	1.4	1.3
IR4	1.2	2	1.3	1.3	1.4	1.1
IR5	1.4	1.8	1.4	1.4	1.3	1.1



Fig. S2.7 Receiver operating characteristic (ROC) curve analysis of constructed motifs in TCPOBOP responsive DHS

A, B, C, D, Area under the curve (AUC) values calculated from ROC curves with the indicated Δ DHS set used as the foreground and a background set of static DHS. The constructed DR4 motif yielded a significant AUC in the DHS that open (A) but not the DHS that close (B). Similar patterns of results were seen for the set of transient (T), sustained early (S), and late Δ DHS (C) and sex-independent DHS (D). Legend at the bottom shows Y-axis labels in enlarged format.

E, F, ROC curves showing the false positive rate (x-axis) and true positive rate (y-axis) for the set of direct repeat motifs (DR0 – DR5; see legend at the bottom) and the nuclear receptor half-site motif discovered from female liver for the set of DHS that open (E) and sustained DHS (F). ROC curves for sex-independent DHS are shown in Fig. 2.3G, 2.3H. The constructed DR4 motif shows the strongest enrichment for being in the sets Δ DHS relative to a background set of static DHS.







Fig. S2.7EF: ROC curves using de novo F27HS_AGGTCA_hex

Fig. S2.8 Comparison of MACS14 score distributions for mouse (mCAR) and human (hCAR) CAR binding sites

A, Boxplots showing the MACS14 score distribution for CAR binding site data from (Niu et al., 2018) for human CAR activated by CITCO (hCAR_CIT), phenobarbital (hCAR_PB) or mouse CAR activated by TCPOBOP (mCAR_TCP) or phenobarbital (mCAR_PB) separated into those that are overlapping a DHS (salmon) or outside a DHS (blue). Activated CAR binding sites inside a DHS showed significantly greater MACS14 score than those outside of a DHS. Median values for each distribution are shown. B, The same CAR binding sites shown in (A) were separated into those that are at a DHS that opens (green), a DHS that closes (red), or a static DHS (blue) following TCPOBOP exposure of male liver at either the 3 h or the 27 h time point. Significantly greater MACS14 score was seen for CAR binding at DHS that open compared to those at sDHS or DHS that close.





Comparison of MACS14 score between CAR binding sites at $\ensuremath{\textbf{B}}$ ΔDHS vs. sDHS



Fig. S2.9Enrichment of ΔDHS vs. sDHS for containing bound CAR (analysisbased on CAR ChIP-exo data from Niu et al, 2018)

A, B, Enrichments of DHS that open (A) or DHS that close (B) relative to sDHS (static DHS) for containing an activated CAR binding site from (Niu et al., 2018). Shown is the enrichment score (ES), Fisher Exact test p-value, and the number and percentage of Δ DHS with a CAR binding site. In male and female liver, DHS that open after 3 h TCPOBOP exposure showed the strongest enrichment for containing activated CAR binding sites. The number and percent of sDHS that overlap each of the four sets of CAR binding sites is shown in the last row of panel A.

C, Enrichments of CAR binding, based on data from (Niu et al., 2018), as in A and B, except that the opening DHS sets analyzed were the transient, sustained early, and late Δ DHS for male liver defined in Fig. 2.1A.

D, E, Enrichments of DHS that open (A) or DHS that close (B) relative to sDHS for containing a TCPOBOP-induced CAR binding site, based on our analysis of the raw ChIP-seq data for CAR binding from (Tian *et al.*, 2018) (set of 2,901 induced CAR binding sites; see Methods and Table S5A, column I). Additional enrichments were calculated to examine induced CAR binding sites which contained either a nuclear receptor half-site or a de novo-discovered DR4-like motif. CAR binding sites with a de novo motif showed the strongest enrichments for being at DHS that open. Fig. S2.9AB: Enrichment of ΔDHS vs. static DHS for containing bound CAR. Analysis is based on CAR ChIP-exo data from Niu et al, 2018.

A. Enrichment at DHS that open

	∆DHS set		mCAR + TCPOBOP	mCAR + PB	hCAR + CITCO	hCAR + PB	
Sex	TCPOBOP Time point	Total number of ΔDHS	Nu	Enrichment Score (ES) p-value Number of ΔDHS with CAR binding sites (%)			
Mala	3 h	472	10.6 1.61E-129 283 (60%)	8.4 1.16E-109 336 (71%)	8.9 6.5E-92 405 (86%)	7.9 2.62E-78 410 (87%)	
Male	27 h	1767	7 8.91E-304 880 (50%)	4.7 4.27E-215 1024 (58%)	4.3 4.49E-178 1313 (74%)	4.3 5.01E-170 1383 (78%)	
Female	3 h	512	18 2.61E-208 368 (72%)	16.2 5.86E-180 423 (83%)	18.4 5.02E-137 474 (93%)	22.4 1.97E-129 486 (95%)	
	27 h	2357	6.3 p = 0 1115 (47%)	4.9 1.33E-299 1392 (59%)	4.7 3.82E-261 1794 (76%)	4.7 1.4E-245 1879 (80%)	
Number of static DHS (55,866 sites) with CAR binding sites (%)		6,943 (12%)	12,656 (23%)	22,573 (40%)	25,390 (46%)		

B. Enrichment at DHS that close

	∆DHS set	ΔDHS set mCAR + TCPOBOP mCAR + PB hCAR + CITCO				hCAR + PB
Sex	TCPOBOP Time point	Total number of ΔDHS	Enrichment Score (ES) p-value Number of DHS with CAR binding sites (%)			
Mala	3 h 72		1.0 p = NS	0.9 p = NS	1.1 p = NS	1.9 p = NS
Male	27 h	564	0.6 p = NS	0.9 p = NS	1.3 p = NS	1.2 p = NS
	3 h	101	0.9 p = NS	1.1 p = NS	1.6 p = NS	1.4 p = NS
Female	27 h	879	0.7 p = NS	1.4 5.32E-06 258 (29%)	2.3 1.04E-33 535 (61%)	2.3 6.51E-34 580 (66%)

Fig. S2.9C: Enrichment of Δ DHS vs. static DHS for containing bound CAR. Analysis is based on CAR ChIP-exo data from Niu et al, 2018. Also see Table_S5A, column H vs columns R, S.

Opening DHS		mCAR + TCPOBOP	mCAR + PB	hCAR + CITCO	hCAR + PB	
ΔDHS set (see Fig. 1A)	Total number of ΔDHS	Enrichment Score (ES) p-value Number of ΔDHS with CAR binding sites (%)				
Transient	263	6.4 8.87E-44 125 (48%)	5.6 4.19E-42 163 (62%)	5.8 8.3E-39 210 (80%)	5.1 5.06E-32 213 (81%)	
Sustained early	208	22.3 2.03E-97 158 (76%)	16.9 5.18E-76 173 (83%)	22.1 5.33E-60 195 (94%)	21.5 5.14E-53 197 (95%)	
Late 1558		6.1 1.18E-227 722 (46%)	6.1 4.1 1.18E-227 1.71E-158 722 (46%) 851 (55%)		3.8 3.55E-130 1185 (76%)	
		6,943 (12%)	12,656 (23%)	22,573 (40%)	25,390 (46%)	

D. Enrichment at DHS that open					CAR sites with a de-novo motif (NR half-site and/or DR4-like motif)	
Sex	TCPOBOP Time point	Total number of opening DHS	Number of DHS with induced CAR binding sites (%)	ES (p-value)	Number of DHS with induced CAR binding sites (%)	ES (p-value)
Mala	3 h	472	93 (20%)	5.4 (4.44E-33)	74 (16%)	8.6 (3.18E-39)
Male	27 h	1767	199 (11%)	2.8 (2.43E-31)	145 (8%)	4.1 (8.97E-40)
Female	3 h	512	118 (23%)	6.6 (4.72E-49)	101 (20%)	11.3 (9.14E-63)
	27 h	2357	252 (11%)	2.6 (3.13E-35)	197 (8%)	4.2 (1.08E-53)

Fig. S2.9DE: Enrichment of Δ DHS vs. sDHS for CAR binding. Analysis is based on CAR ChIP-seq data from Tian et al, 2018, processed as described in Methods.

E. Enrichment at DHS that close					CAR sites with a de-novo motif (NR half-site and/or DR4-like motif	
Sex	TCPOBOP Time point	Total number of closing DHS	Number of DHS with repressed CAR binding sites (%)		Number of DHS with <mark>repressed</mark> CAR binding sites (%)	ES (p-value)
Male	3 h	72	0 (0%)	0 (1)	0 (0%)	0 (1)
IVIAIC	27 h	564	21 (4%)	42.3 (9.87E-25)	21 (4%)	93.9 (3.3E-30)
Female	3 h	101	1 (1%)	10.9 (0.1346)	1 (1%)	24.3 (0.127)
Female	27 h	879	1 (0.1%)	1.2 (0.556)	1 (0.1%)	2.8 (0.333)

Fig. S2.10Enrichment for TCPOBOP-responsive genes (differentially expressedgenes, DEGs) for containing a CAR binding site (ChIP-exo) in the same TAD

A, B, C, Enrichments of TCPOBOP-induced genes (A), TCPOBOP-repressed genes (B), and MET/EGFR-dependent TCPOBOP-inducible genes (see Methods) (C) for proximal mouse CAR binding, based on CAR binding sites mapped to the closest RefSeq or multiexonic lncRNA gene TSS in the same TAD (CAR target genes). TCPOBOPinduced/repressed genes, identified in male liver after 3 h or 27 h TCPOBOP exposure, were compared to TCPBOP-unresponsive genes for being a CAR target. TCPOBOPinduced genes at both time points and repressed genes at 27 h showed enrichment for containing proximal CAR binding sites. Similar patterns of enrichment were seen between the gene sets that were independent or dependent on MET/EGFR signaling. The percentage of TCPOBOP-unresponsive (background) genes targeted by each set of CAR binding sites was 24% and 35% for mCAR activated by TCPOBOP and phenobarbital (PB), respectively; and was 46% and 51% for hCAR activated by CITCO and PB, respectively.

D, Enrichments of TCPOBOP-induced genes for containing proximal mouse CAR binding (as defined in A, B, C) for various subsets of DHS with the set of 2,901 induced CAR binding sites based on the ChIP-seq data of (Tian *et al.*, 2018) and shown in Table S5A, column I. Either the full list of 60,739 DHS (rows 1, 2) or the set of DHS that open in male liver (rows 3-5) were used for these enrichment calculations. Enrichments indicate the effect of any DHS (rows 1, 2) with an induced CAR binding site with a motif. TCPOBOP induced genes showed stronger enrichments for containing a proximal induced CAR binding site with a motif compared to those without a motif. DHS that open with an induced CAR binding site with a motif showed the strongest enrichment for being proximal to TCPOBOP induced genes (Enrichments numbered 3-5). **Fig. S2.10ABC**: Enrichment for TCPOBOP-responsive genes (DEGs) for containing a CAR binding site (ChIP-exo) in the same TAD

Α.	TCPOBOP-induced	genes
- 1		

		hCAR + CITCO	hCAR + PB	mCAR + PB	mCAR + TCPOBOP
TCPOBOPTotalEnrichment Score (ES)Timenumberp-valuepointof DEGsNumber of DEGs with CAR binding sites (%)		Enrichment S p-valu Number of DEGs with C		s (%)	
Male 3 h	138	6.6 1.4E-20 118 (86%)	6.2 3.78E-18 120 (87%)	6.5 3.67E-24 108 (78%)	7.4 9.81E-30 97 (70%)
Male 27 h	708	3.3 2.97E-49 525 (74%)	3.2 5.51E-44 544 (77%)	3.1 9E-50 445 (63%)	3.7 4.61E-62 375 (53%)

B. TCPOBOP-repressed genes

		hCAR + CITCO	hCAR + PB	mCAR + PB	mCAR + TCPOBOP
TCPOBOP Time point	Total number of DEGs	Enrichment Score (ES) p-value Number of DEGs with CAR binding sites (%)			
Male 3 h	35	3.2 p = NS	2.7 p = NS	1.9 p = NS	1.6 p = NS
Male 27 h	356	3.9 5.57E-32 275 (77%)	3.5 1.99E-26 280 (79%)	3 6.82E-24 219 (62%)	2.5 1.35E-15 153 (43%)
			1		

С.		hCAR + CITCO	hCAR + PB	mCAR + PB	mCAR + TCPOBOP
Gene Set	Total number of DEGs	Enrichment Score (ES) p-value Number of DEGs with CAR binding sites (%)			
TCPOBOP-induced,		3.4	3.4	2.6	3.1
Not dependent on	320	7.19E-24	3.32E-23	1.31E-16	6.85E-23
ME1/EGFR		238 (75%)	250 (78%)	185 (58%)	157 (49%)
TCPOBOP-induced.		3	2.6	2.7	3.5
Dependent on	187	1.54E-12	2.45E-09	1.46E-11	6.06E-17
MET/EGFR		135 (72%)	136 (73%)	111 (59%)	97 (52%)

1

Figure S2.10D: Enrichment of CAR binding sites proximal to TCPOBOP responsive genes (DEGs)

	Enrichment for mapping	F3 h up	M3 h up	F27 h up	M27 h up
	to induced DEGs	genes	genes	genes	genes
1	Ratio A: DHS + induced	3.2%	2.4%	5.6%	7.4%
	CAR TFBS	(90/2790)	(68/2825)	(149/2671)	(194/2626)
	Ratio B: DHS without CAR	3.2	3.5	2.1	1.9
	TFBS	p = 1.9E-18	p = 5.18E-16	p = 1.1E-14	p = 1.58E-14
2	Ratio A: DHS + induced	4.3%	3.7%	6.9%	9.3%
	CAR TFBS + motif	(61/1414)	(53/1432)	(93/1357)	(124/1328)
	Ratio B: DHS without CAR	4.2	5.3	2.6	2.4
	TFBS	p = 1.12E-18	p = 2.75E-20	p = 1.9E-14	p = 2.46E-16
3	Ratio A: dDHS that open Ratio B: sDHS	10.7% (331/3096) 10.4 p = 1.17E- 174	9.5% (298/3136) 13.4 p = 2.43E- 180	22.8% (637/2789) 8.5 p = 1.53E- 288	26.6% (717/2701) 6.7 p = 3.73E- 271
4	Ratio A: dDHS that open +	20.3%	13.9%	26.7%	33.1%
	induced CAR TFBS	(61/300)	(44/317)	(75/281)	(89/269)
	Ratio B: sDHS without	20	20	10.1	8.5
	CAR TFBS	p = 1.92E-52	p = 9.73E-39	p = 5.21E-44	p = 1.13E-45
5	Ratio A: dDHS that open +	23.3%	16.7%	30.5%	35.8%
	induced CAR TFBS + motif	(49/210)	(37/222)	(60/197)	(68/190)
	Ratio B: sDHS without	22.9	24	11.5	9.2
	CAR TFBS	p = 3.55E-45	p = 1.46E-35	p = 2.28E-38	p = 3.67E-37

Fig. S2.11 21 de novo motifs, discovered from sets of Δ DHS that open, clustered by sequence similarity

A, De novo motifs discovered from TCPOBOP opened DHS sequences were clustered by motif sequence similarity, as shown. MEME-ChIP and cisGenome were used for motif discovery and motifStack was used to cluster motifs by sequence similarity; a total of 5 clusters were obtained. The header above each motif logo indicates the data set used to discover that motif. For instance, "F3hr_dDHS_open" and "F27hr_dDHS_open" indicate that those motifs were discovered in TCPOBOP opened DHS in female liver at 3 h and 27 h exposure time points, respectively.

B, C, De novo motifs shown in (A) are listed along with the corresponding best matching motif, which was determined as the motif with the most significant Tomtom q-value. Shown is the de novo motif cluster designation (column 1), the de novo motif logo (column 2), the motif logo for the 1st best matching motif (column 3), and the corresponding motif name and q-value (column 4) with additional matching motifs listed.



Figure S2.11A: Clusters of de novo motifs (using motifStack)



Figure S2.11B: Identifying de novo motifs based on sequence similarity to known motifs (using Tomtom and Meme-Chip)

Motif Cluster	De novo motif	1 st best match	Motif name (q-value)
Cluster 1: Nuclear receptor	CAANT	¹ ACGTCAAAGGTCAAA	 NR1H2-RXRA (motif2) (4.1e-04) NR2F1 (motif71) (8.1e-04)
motifs	aAGTCA	ACGTCA	 RORA_1 (motif18) (4.5e-03) Nr2f6(var.2) (MA0728.1) (5.5e-03) motif65 (VDR,CAR,PXR) (8.97e-03)
	AGGTCA AUT-CA		 VDR,CAR,PXR (motif65) (4.7e-07) Rarg(var.2) (MA0860.1) (8.8e-06) Nr1h3::Rxra (MA0494.1) (1.0e-05)
	CAACT_I _ CASCT		 Rarg(var.2) (MA0860.1) (9.2e-05) VDR,CAR,PXR (motif65) (1.5e-04) Nr1h3::Rxra (MA0494.1) (7.7e-03)
Cluster 2: Nuclear factor motifs		s the AttAc	1. HNF1 (motif33) (3.3e-02)
	AAT AT TAASTA	s. the AttAc	 HNF1 (motif33) (6.1e-05) HNF1A (MA0046.2) (3.8e-03) HNF1B (MA0153.2) (1.0 e-02)
		T. sfA.	 CEBP (motif54) (2.1e-03) NFIL3 (MA0025.1) (5.1e-03) CEBPA (MA0102.3) (5.1e-03)
			1. Clox (motif24) (1.1e-04) 2. ONECUT1 (MA0679.1) (2.3e-04) 3. ONECUT2 (MA0756.1) (2.3e-04)
Cluster 3: FOX factor motifs	AA SEAAA & AA & AA		1. IRF1 (motif5) (1.4e-01) 2. IRF2 (motif72) (1.4e-01) 3. FOXI1 (motif45) (4.7e-01)
	Asa asa Asasa asasa		 FOXI1 (motif45) (7.1e-02) I. IRF2 (motif72) (7.79e-02)

Figure S2.11C: Identifying de novo motifs based on sequence similarity to known motifs (using Tomtom)

Motif Cluster	De novo motif	1 st best match	Motif name (q-value)
Cluster 4: GC-rich motifs	GCCCG CCCC		1. SP1 (MA0079.3) (3.6e-03) 2. SP1 (motif92) (4.5e-03)
			1. KLF9 (MA1107.1) (2.2e-03) 2. RREB1 (MA0073.1) (8.2e-03) 3. SP3 (MA0746.1) (7.5e-03)
	<mark>ççç_eççççççççç</mark>		 Egr (motif87) (2.4e-02) RREB1 (motif85) (1.2e-01)
		N/A	No significant motif matches reported
		, T , , , , , , , , , , , , , , , , , , ,	1. Zfx (MA0146.2) (6.9e-03)
Cluster 5: GA-rich motifs		N/A	No significant motif matches reported
	6.61. <u></u>	<u></u>	1. ZNF263 (MA0528.1) (2.5e-05) 2. Elf-1 (motif88) (1.27e-03)
	CaGeGes Vaga	SCCTCASAUGTCA	1. NR2C2 (MA0504.1) (8.3e-02)
	AAGUACACA	tt CACCACC ,c <mark>GGaCAG</mark> ec	1. NRSF (motif51) (3.9e-02)
		cGAGGAcca	1. ZNF263 (MA0528.1) (2.4e-03)
	eeeeAAAeeAAAe		1. ZNF384 (MA1125.1) (9.6e-03)

Fig. S2.12 Enrichment of DHS that open for containing de-novo discovered motifs

A, B, Enrichments of DHS that open for containing motifs belonging to clusters 2-5 (shown in Fig. S2.11) relative to a background set of 55,866 static DHS (a subset of these motifs is presented in Fig. 2.8). The strongest enrichment was seen for a FOX-like motif, discovered from DHS that open in female liver after 27 h TCPOBOP exposure, with up to 73% of Δ DHS containing a match to this motif (ES: 13, p = 6.24E-153).

dDHS that open	Male 3hr	Male 27 hr	Female 3 hr	Female 27 hr
	Cluster 2:	Nuclear factor mo	otifs	
	38%(177/472) ES: 1.2 (BG:34%)0.131	42%(743/1767) ES: 1.4 (BG:34%)5.89e-12	44%(224/512) ES: 1.5 (BG:34%)8.09e- 06	43%(1008/2357) ES: 1.5 (BG:34%)8.41e-18
HNF1A_F27o	40%(188/472) ES: 1.1 (BG:37%)0.318	45%(799/1767) ES: 1.4 (BG:37%)5.96e-11	45%(228/512) ES: 1.3 (BG:37%)0.0014	47%(1100/2357) ES: 1.5 (BG:37%)6.26e-19
	45%(211/472) ES: 1.5 (BG:35%)3.77e-05	43%(766/1767) ES: 1.4 (BG:35%)5.24e-12	46%(233/512) ES: 1.5 (BG:35%)3.18e-06	43%(1015/2357) ES: 1.4 (BG:35%)1.56e-14
HNF6_F270	38%(178/472) ES: 1.4 (BG:30%)0.001	39%(696/1767) ES: 1.5 (BG:30%)4.67e-15	34%(172/512) ES: 1.2 (BG:30%)0.126	41%(976/2357) ES: 1.6 (BG:30%)3.62e-28
	Cluster 3	: FOX factor moti	fs	
RREB1_F3o	59%(280/472) ES: 3 (BG:33%)2.57e- 31	65%(1145/1767) ES: 3.8 (BG:33%)5.72e- 160	60%(309/512) ES: 3.1 (BG:33%)2.91e- 36	64%(1518/2357) ES: 3.7 (BG:33%)1.28e- 205
FOX_F270	73%(343/472) ES:13 (BG:17%) 6.24e-153	72%(1279/1767) ES: 12.8 (BG:17%)0	71%(366/512) ES:12.2 (BG:17%) 4.26e-159	71%(1681/2357) ES: 12.1 (BG:17%)0
	Cluster	4: GC-rich motifs		
SP1_M27o	73%(343/472) ES: 5 (BG:35%)9.6e-62	62%(1092/1767) ES: 3 (BG:35%)3.67e- 114	65%(334/512) ES: 3.5 (BG:35%)2.75e- 43	65%(1521/2357) ES: 3.4 (BG:35%)1.16e- 181
SP1_M3o [*] <mark></mark>	66%(313/472) ES: 5.9 (BG:25%) 3.04e-77	54%(963/1767) ES: 3.6 (BG:25%) 8.71e-146	56%(287/512) ES: 3.8 (BG:25%) 1.13e-48	56%(1328/2357) ES: 3.8 (BG:25%) 4.54e-215

Fig. S2.12A: Enrichments for set of 21 motifs in opening DHS

dDHS that open	Male 3hr	Male 27 hr	Female 3 hr	Female 27 hr	
	Cluster 4: GC	-rich motifs (cont	inued)		
RSC30_M3o	38%(178/472) ES: 2.4 (BG:20%)1.2e-18	25%(446/1767) ES: 1.4 (BG:20%)8.19e- 08	26%(133/512) ES: 1.4 (BG:20%)0.0011 2	27%(632/2357) ES: 1.5 (BG:20%)2.57e- 15	
abi4_F3o	55%(258/472) ES: 3.1 (BG:28%)2.07e-32	46%(806/1767) ES: 2.1 (BG:28%)2.68e-52	47%(242/512) ES: 2.3 (BG:28%)3.4e-19	47%(1111/2357) ES: 2.3 (BG:28%)2.96e-80	
	53%(252/472) ES: 3.3 (BG:26%)3.33e-36	43%(767/1767) ES: 2.2 (BG:26%)1.32e-55	44%(224/512) ES: 2.2 (BG:26%)5.34e- 18	44%(1049/2357) ES: 2.3 (BG:26%)1.63e-81	
	Cluster 5: GA-rich motifs				
Gata1_M27o	79%(373/472) ES: 3.6 (BG:51%)8.51e- 35	75%(1324/1767) ES: 2.8 (BG:51%)2.6e-89	73%(372/512) ES: 2.5 (BG:51%)2.28e- 22	74%(1747/2357) ES: 2.7 (BG:51%)1.98e- 109	
	80%(376/472) ES: 7 (BG:36%) 7.67e-83	71%(1248/1767) ES: 4.3 (BG:36%) 3.73e-186	71%(362/512) ES: 4.3 (BG:36%) 4.55e-56	72%(1686/2357) ES: 4.5 (BG:36%) 2.98e-258	
RXRA_F27o	84%(398/472) ES: 3.7 (BG:59%)6.33e- 31	77%(1362/1767) ES: 2.3 (BG:59%)1.72e- 53	81%(415/512) ES: 2.9 (BG:59%)6.4e-25	78%(1834/2357) ES: 2.4 (BG:59%)1.84e- 76	
	87%(412/472) ES: 5.4 (BG:56%)3.36e-48	80%(1422/1767) ES: 3.3 (BG:56%)2.24e- 102	86%(441/512) ES: 4.9 (BG:56%)1.03e- 47	82%(1931/2357) ES: 3.6 (BG:56%)1.64e- 152	
Klf4_M27o	85%(403/472) ES: 8.1 (BG:42%)4.04e-83	80%(1413/1767) ES: 5.5 (BG:42%)1.48e- 226	81%(413/512) ES: 5.8 (BG:42%)4.96e- 70	80%(1882/2357) ES: 5.5 (BG:42%)4.18e- 297	
	60%(283/472) ES: 7.2 (BG:17%)1.68e-93	64%(1138/1767) ES: 8.7 (BG:17%)0	60%(308/512) ES: 7.3 (BG:17%)3.65e- 102	63%(1482/2357) ES: 8.2 (BG:17%)0	

Fig. S2.12B: Enrichments for set of 21 motifs in opening DHS

Fig. S2.13 Enrichment of CAR bound opening and static DHS for containing denovo discovered motifs

A, B, C, Enrichments of DHS groups B, C, and D (defined above panel A, and see Fig. 2.7A) for containing the same set of 21 de novo motifs shown in Fig. S2.11, relative to a background set of static DHS (the same set of 55,866 static DHS used in Fig. S2.12). The robust CAR-bound opening DHS (Group D) showed the strongest enrichment for the DR4-like motif with 84% of Δ DHS containing the motif (ES: 4.4, p = 5.64E-40) (S2.13A). Corresponding enrichments for motifs belonging to cluster 2, 3 (S2.13B) and cluster 4, 5 (S2.13C) are shown. DHS group C was depleted of NR half-site and DR4 motifs (cluster 1 motifs) compared to Groups B and D (S2.13A). DHS group C also showed significant enrichments for motifs for liver-enriched transcription factors HNF1 and HNF6/CUX, which were not seen with Groups B and D; while DHS in Group B, and to a lesser extent DHS in Group D, showed stronger enrichment for most Cluster 4 and Cluster 5 motifs than DHS in Group C.

Fig. S2.13A. Motif scan summary for de-novo motifs matching the known CAR motif (motif 65)



	DHS subsets	Group_C_335_ DHS	Group_D_455_ DHS	Group_B_2535_ DHS
	NR2F1_F270	67%(225/335) ES: 1.2 (BG:63%) 1.12e-01	80%(365/455) ES: 2.4 (BG:63%) 2.56e-15	79%(1999/2535). ES: 2.2 (BG:63%) 1.25e-64
	CAR_F270	55%(185/335) ES: 1.3 (BG:49%) 2.62e-02	72%(327/455) ES: 2.7 (BG:49%) 5.02e-23	66%(1684/2535)ES: 2.1 (BG:49%) 2.79e-68
	HNF4A F30	60%(202/335) ES: 1.3 (BG:55%) 4.60e-02	84%(383/455) ES: 4.4 (BG:55%) 5.64e-40	70%(1781/2535) ES: 2 (BG:55%) 1.04e-55
GA	CAR_F30	51%(171/335) ES: 1.3 (BG:45%) 2.77e-02	77%(349/455) ES: 4.1 (BG:45%) 8.40e-43	59%(1489/2535). ES: 1.8 (BG:45%) 2.94e-43

DHS subsets	Group_C_335_	Group_D_455_	Group_B_2535_
	DHS	DHS	DHS
	Cluster 2: Nuclear f	actor motifs	
	48%(160/335)	32%(147/455)	29%(724/2535)
	ES: 1.8(BG:34%)	ES: 0.9(BG:34%)	ES: 0.8(BG:34%)
	4.01e-07	4.80e-01	1.13e-08
HNF1A_F27o	51%(170/335)	36%(164/455)	31%(776/2535)
	ES: 1.7(BG:37%)	ES: 0.9(BG:37%)	ES: 0.7(BG:37%)
	1.58e-06	5.60e-01	2.42e-12
	45%(150/335)	42%(190/455)	41%(1043/2535)ES:
	ES: 1.5(BG:35%)	ES: 1.3(BG:35%)	1.3(BG:35%)
	4.75e-04	4.94e-03	1.77e-09
HNF6_F27o	50%(167/335)	33%(151/455)	28%(718/2535)
	ES: 2.3(BG:30%)	ES: 1.1(BG:30%)	ES: 0.9(BG:30%)
	4.38e-13	2.22e-01	2.71e-02
	Cluster 3: FOX fac	tor motifs	
RREB1_F30	72%(241/335)	55%(250/455)	56%(1415/2535)ES:
	ES: 5.3(BG:33%)	ES: 2.5(BG:33%)	2.6(BG:33%)
	2.63e-47	7.20e-22	2.59e-119
FOX_F270	77%(257/335)	68%(310/455)	65%(1660/2535)ES:
	ES: 16.1(BG:17%)	ES: 10.4(BG:17%)	9.3(BG:17%)
	1.28e-124	8.38e-126	0.00e+00
	Cluster 4: GC-ric	h motifs	
SP1_M27o	54%(182/335)	69%(312/455)	74%(1879/2535)ES:
	ES: 2.2(BG:35%)	ES: 4.1(BG:35%)	5.4(BG:35%)
	7.19e-13	8.19e-48	0.00e+00
SP1_M3o	48%(160/335)	62%(282/455)	68%(1720/2535)ES:
	ES: 2.7(BG:25%)	ES: 4.8(BG:25%)	6.3(BG:25%)
	2.92e-18	2.41e-60	0.00e+00

Fig. S2.13B. Enrichments for set of 21 motifs in DHS from Groups C and D

DHS subsets	Group_C_335_	Group_D_455_	Group_B_2535_		
	DHS	DHS	DHS		
Cluster 4: GC-rich motifs					
RSC30_M3o	24%(79/335)	31%(143/455)	40%(1017/2535)		
	ES: 1.2(BG:20%)	ES: 1.8(BG:20%)	ES: 2.7(BG:20%)		
	1.04e-01	7.60e-09	4.75e-114		
abi4_F3o	40%(133/335)	53%(243/455)	59%(1502/2535)		
	ES: 1.7(BG:28%)	ES: 2.9(BG:28%)	ES: 3.7(BG:28%)		
	9.13e-06	5.91e-29	6.16e-221		
	39%(130/335)	52%(237/455)	59%(1484/2535)		
	ES: 1.8(BG:26%)	ES: 3.1(BG:26%)	ES: 4.1(BG:26%)		
	4.01e-07	3.21e-32	1.39e-251		
	Cluster 5: GA-ric	h motifs			
Gata1_M27o	70%(233/335)	76%(347/455)	82%(2070/2535)		
	ES: 2.2(BG:51%)	ES: 3.1(BG:51%)	ES: 4.2(BG:51%)		
	2.94e-11	2.33e-27	3.09e-214		
Elf-1_F270	63%(212/335)	81%(367/455)	83%(2108/2535)		
	ES: 3.1(BG:36%)	ES: 7.4(BG:36%)	ES: 8.8(BG:36%)		
	2.50e-23	1.84e-83	0.00e+00		
	70%(233/335)	85%(385/455)	86%(2187/2535)		
	ES: 1.6(BG:59%)	ES: 3.8(BG:59%)	ES: 4.3(BG:59%)		
	1.99e-04	7.03e-31	4.41e-185		
	76%(253/335)	85%(386/455)	86%(2181/2535)		
	ES: 2.4(BG:56%)	ES: 4.4(BG:56%)	ES: 4.9(BG:56%)		
	3.03e-13	2.57e-39	4.83e-226		
	77%(258/335)	83%(379/455)	86%(2190/2535)		
	ES: 4.6(BG:42%)	ES: 6.9(BG:42%)	ES: 8.8(BG:42%)		
	5.41e-38	1.54e-72	0.00e+00		
	70%(235/335)	57%(261/455)	52%(1311/2535)		
	ES: 11.3(BG:17%)	ES: 6.5(BG:17%)	ES: 5.2(BG:17%)		
	1.56e-99	3.49e-81	1.98e-323		

Fig. S2.13C. Enrichments for set of 21 motifs in DHS from Groups C and D

CHAPTER 3 – Chromatin Accessibility Alterations Due to Plasma Growth Hormone Pulses in Adult Male Mouse Liver

3.1 Abstract

Background: Growth hormone (GH) secretory patterns in mouse liver regulate the transcription of $\sim 1,000$ sex-biased genes either through positive regulation (class I) or negative regulation (class II) sex-biased genes. GH stimulation of hepatocytes is either persistent, near continuous (female liver) or intermittent and pulsatile (male liver) in mouse models. Mechanisms for GH activation of STAT5 regulating chromatin structure of sex-biased DHS are unknown. Further, little is known about how pulsatile GH secretion contributes to male-biased chromatin accessibility in male liver. Approach: We used DNase hypersensitivity site (DHS) analysis to identify genomic regions that open and close in response to endogenous GH/STAT5 pulses in male liver. We also investigated the impact of a single pulse of GH given to hypophysectomized (hypox) male mice determined 30, 90, or 240 min later. Results: We identified two classes of male biased DHS (Male-DHS): dynamic male-biased DHS (834 sites), almost all bound by STAT5; these DHS undergo repeated cycles of chromatin opening and closing induced by each GH pulse as the primary mechanism for the male bias in chromatin accessibility; and static male-biased DHS (1,895 sites), whose accessibility in male liver does not vary with GH-stimulated STAT5 pulses, and whose male bias results from these sites being more closed in female liver, as demonstrated by DNase-I cutting frequency analysis. Strikingly, a single physiological dose of GH in hypophysectomized male mice

was sufficient for rapid and temporary restoration of chromatin accessibility at ~60% of DHS that are opened by endogenous GH/STAT5 pulses. Our findings also indicate that direct binding of STAT5 is associated with increased chromatin opening at these GH-responsive DHS. <u>Conclusions:</u> GH secretion patterns play a direct role in regulating sexspecific differences in chromatin accessibility in mouse liver. We identify two distinct mechanisms for explaining male-biased DHS: those that are dynamic vs. static. Chromatin state and transcription factor binding data suggest that dynamic male-biased DHS are active enhancer regions and are preferentially bound by STAT5. Together, these studies reveal that endogenous rhythms of male plasma GH pulsation dynamically open and close chromatin at localized regulatory regions and provide a better understanding of male-biased chromatin accessibility.

3.2 Introduction

Growth hormone (GH) regulates liver metabolizing enzymes and transporters that play critical roles in detoxification in mammalian liver following exposure to many lipophilic foreign chemicals (Vijayakumar et al., 2011). These enzymes are also intricately involved in regulating liver physiology by metabolizing cholesterol, steroid hormones, and other endogenous lipophiles. Dysregulation of GH-regulated enzymes can lead to major disorders (Ayuk and Sheppard, 2006)and development of disease, such as obesity (Scacchi et al., 1999). Cytochromes P450 (Cyps) and other steroid-metabolizing gene families show striking sex-biased gene expression in mouse and rat liver, thereby enabling each sex to meet their respective metabolic requirements (Waxman and Holloway, 2009). Adult liver expression of sex-biased genes is regulated by two sexspecific, temporally distinct patterns of adult pituitary GH secretion (O'Connor and Waxman, 2006). Approximately 1,000 sex-biased genes are regulated either through positive regulation (class I sex-biased genes) or negative regulation (class II sex-biased genes) (Clodfelter et al., 2006). This sex-differential regulation is achieved when plasma GH levels are persistent (near continuous), as occurs in female liver or in cGH-treated male liver) and when it is intermittent (pulsatile), as occurs in intact male liver, as was shown in both mouse and rat liver models (O'Connor and Waxman, 2006). The sex-dependent actions of GH require the transcription factor STAT5 (primarily STAT5b), which is activated by GH-induced tyrosine phosphorylation of STAT5 stimulated by the Janus kinase-signal transducer and activator of transcription (JAK-STAT) signaling pathway (Aaronson and Horvath, 2002). Activated STAT5, phosphorylated by the activation of GH-receptor on a single tyrosine residue, is translocated to the nucleus where target gene expression is induced by proximal STAT5 binding (O'Connor and Waxman, 2006). In male mouse and rat liver, STAT5 activation and nuclear translocation is repeatedly stimulated by each successive plasma GH pulse. The activated STAT5 homodimer binds to specific DNA sequences (binding site motif: TTCNNNGGA) at sites that are strongly enriched nearby male-biased genes (Zhang et al., 2012). In female liver, STAT5 is persistently activated by the near-continuous presence of circulating GH and shows enriched binding nearby female-biased genes (Zhang et al., 2012). Other GH/STAT5-regulated transcription factors (TFs) facilitate sex differences in liver gene expression, namely BCL6, a male-biased repressor (Zhang et al.,
2012), and CUX2, a female-specific repressor (Conforto et al., 2012). STAT5 and other GH-regulated liver TF genome-wide binding sites, identified by ChIP-seq, show malebiased binding enriched at open chromatin regions (DNase-I hypersensitive sites, DHS) (Ling et al., 2010)enriched nearby male-biased genes, and female-biased binding enriched nearby female-biased genes (Conforto et al., 2012; Zhang et al., 2012). These robust liver sex differences are, in part, achieved by a complex interplay of multiple GHregulated TFs. A subset of sex-biased genes respond rapidly to pulses of GH and liver STAT5 activity(Connerney et al., 2017); however, most sex-biased genes respond slowly (over the course of days) (Lau-Corona et al., 2017) to a change in plasma GH status, in part due to a requirement for secondary changes, which include changes in histone modifications and the underlying chromatin state (Sugathan and Waxman, 2013). Changes in chromatin accessibility, a hallmark of epigenetic regulation, can be measured by global DHS analysis using DNase-seq, which we used to identify $\sim 70,000$ in male and female liver (Ling et al. 2010) that were classified as enhancer (e), promoter (p), or insulator (i) regions (Matthews and Waxman, 2018). We identified ~4,000 GH regulated and sex-biased DHS, ~95% of which are enhancer DHS (eDHS) based on data for activating and repressive chromatin marks and ChIP-seq binding profiles of insulator complex proteins CTCF and cohesin (Matthews and Waxman, 2018). These sets of sexbiased DHS are significantly depleted of promoter and insulator DHS, show strong enrichment for correspondingly sex-biased binding sites for STAT5 and other GHregulated TFs, and are strongly enriched for proximity to sex-biased gene promoters (Ling et al., 2010; Sugathan and Waxman, 2013). Thus, GH-regulated sex differences in

chromatin structure are a key feature of sex-differential liver gene expression. However, little is known about the mechanisms by which sex-dependent temporal GH secretion patterns dictate robust sex differences in chromatin accessibility to regulate liver gene expression.

Here, we used DNase-seq to identify genomic regions that open and/or close in response to endogenous pulses of STAT5 activity in male liver. These sites were used to identify 834 dynamic male-biased eDHS, where direct STAT5 binding induced by each plasma GH pulse repeatedly induces liver chromatin to open and then close in direct response to each plasma GH pulse. Further, we show that this rhythmic chromatin opening induces pulsatile transcription of a subset of STAT5-dependent male-biased genes. Pulsatile opening of these dynamic eDHS in male liver accounts for the male bias of 31% of all male-biased eDHS. Thus, we show that male plasma GH stimulation directly establishes the sex difference in accessibility at these DHS, which underlies sex-dependent binding of STAT5 and other TFs linked to sex-biased liver gene expression. Further, we use a hypophysectomized (hypox) male mouse model to investigate the chromatin accessibility changes due a single pulse of GH and evaluate the relative levels of chromatin opening at these GH responsive sites at biologically relevant time points. Finally, we investigate the ability of a single GH pulse given to hypox mice to recapitulate the changes in chromatin accessibility seen in pituitary-intact male mice liver.

3.3 Materials and Methods

3.3.1 Animal treatments and DNase-seq analysis

All animal work was carried out in accordance with the accepted standards of humane animal care, in compliance with the procedures and protocols approved by the Boston University Institutional Animal Care and Use Committee. Animal handling and treatments and other wet laboratory work was performed as described previously (Connerney et al., 2017) and was carried out by Dr. Jeannette Connerney of this laboratory. Briefly, male and female CD1 mice (ICR strain), 7-weeks old, were purchased from Charles River Laboratories (Wilmington, MA) and kept on a 12-hour light/dark cycle with food and water without restriction. Mice were intact or were hypophysectomized by the supplier with the absence of weight gain and the lack of Mup protein in mouse urine as indicators of complete hypophysectomy (hypox). Hypophysectomized male mice were given a single intraperitoneal injection of vehicle (control) with or without recombinant rat GH at 125 ng of growth hormone (GH) per gram body weight and were euthanized by cervical dislocation 30, 90, or 240 min after GH injection (Connerney et al., 2017). In addition, fresh livers were collected from individual untreated, intact male mice (between 8 to 10 weeks of age). Portions of each liver were used to prepare a protein extract for analysis of liver STAT5 activity by electrophoretic mobility shift assay (EMSA) (Connerney et al., 2017) to identify individual mouse livers as STAT5-high activity (n = 10) and STAT5-low activity DNaseseq samples (n = 11) at the time the mice were euthanized. DNase-I treatment to release

DNA fragments from DHS regions to identify liver DNase hypersensitive sites (DHS) were carried out as described elsewhere (Connerney et al., 2017).

3.3.2 DNase-seq libraries, sequencing, and data analysis

DNase-seq was performed for intact or hypophysectomized (hypox) male and female mouse liver, and livers of hypox male mice given a single injection of GH and euthanized either 30, 90, or 240 min later. DNase-seq libraries were prepared for each treatment group by Dr. Jeannette Connerney of this laboratory, with each library representing a pool of independent liver DNase-seq samples, as follows. For each treatment or control group, DNase-I-released fragments were purified from nuclei obtained from each individual liver, and then combined to give a single pooled sample, which was used to prepare a single sequencing library using the NEBNext Ultra Library Prep Kit for Illumina (New England Biolabs). Sequencing was performed on an Illumina HiSeq instrument, and single-end sequence reads, 40 to 50 bp in length, were obtained at a sequencing depth ranging from 12 to 33 million total mapped reads for each sample. More detailed sequencing statistics are shown in Supplementary Table_S1. Raw and processed data are available at www.ncbi.nlm.nih.gov/gds under accession GSE131848 and GSE131852 (SuperSeries GSE131853).

Sequencing data was analyzed using a custom DNase-seq pipeline, as described elsewhere (Lodato et al., 2018). Briefly, the pipeline processes raw FASTQ files and outputs various control metrics, including FASTQC reports (FastQC v0.11.7), confirmation of read length, verification of the absence of read strand bias, and quantification of contaminating adaptor sequence (Trim_galore v0.4.2). Reads were

mapped to the mouse genome (mm9 release) using Bowtie2 (v2.2.6) (Langmead and Salzberg, 2012). Regions of DNase hypersensitivity (DHS) were discovered as peaks identified by MACS2 (v2.1.0.20150731) (Feng et al., 2012)using the options (-nomodel shift -100 -extsize 200), to inhibit read shifting, and (-keep-dup), to retain all reads that contribute to peak signal. Peaks were discovered for each DNase-seq sample per treatment condition and were filtered to remove those that overlap ENCODE blacklisted regions (Landt et al., 2012) as well as peaks comprised of > 4 identical reads that do not overlap any other read ("straight peaks"). Additional intact male liver DNase-seq samples generated from the ENCODE consortium (GEO accession: GSM1014195, replicates #5 to #13, n = 9) were analyzed using the custom DNase-seq pipeline described above. Raw data FASTQ files were downloaded from the UCSC Genome Browser (https://genome.ucsc.edu/: "UW DNaseI HS Downloadable Files") with the following parameters: cell line: liver, strain: C57BL/6, age: adult 8 weeks, sex: male, view: raw data, and tissueSourceType: Individual. Single end sequencing reads, 36 bp in length, were obtained at a sequencing depth ranging from 19 to 37 million total mapped reads for each sample.

3.3.3 Liver DHS classification

We previously identified a set of 72,862 DHS regions identified in male and female mouse liver (Ling et al., 2010). A subset of these DHS (515 peaks) overlapped ENCODE blacklisted regions and were filtered out and excluded from all downstream analyses. All but 2,136 of the remaining DHS could be classified according to a five class DHS model defined by promoter, weak promoter, enhancer, weak enhancer, and insulator DHS based on ChIP-seq signals for H3K4me1, H3K4me3, CTCF binding, and proximity to RefSeq gene transcription start sites (TSS) (Matthews and Waxman, 2018). This resulted in a set of 70,211 DHS regions, which were used for all downstream analysis.

3.3.4 Mapping of DHS to gene targets

Each of the 70,211 liver DHS was assigned a single putative gene target (RefSeq gene or multi-exonic lncRNA gene) that corresponds to the closest TSS within the same topologically associating domain (TAD). Supplementary Table_S2 presents the putative gene target of each of the 70,211 DHS and its gene expression data in male and female liver, as well as the response of each DHS to hypox and to hypox+GH treatment at each time point mentioned above.

3.3.5 STAT-high and STA5-low activity livers and ΔDHS regions

Individual STAT5-high activity and STAT5 low-activity mouse livers, identified by EMSA analysis, as described above, were analyzed by DNase-seq to discover DHS that respond dynamically to endogenous plasma GH pulses and liver STAT5 activity. diffReps analysis (Shen et al., 2013) was used to discover genomic sites that were more open or were more closed (DNase-seq |fold-change| > 2 and FDR < 0.05 [Benjamini-Hochberg adjusted p-value]) in the EMSA-identified STAT5-high activity livers compared to the EMSA-identified STAT5-low activity livers, using the diffReps nucleosome option (200 bp window size) and setting (-frag) to zero for all comparisons. The significant differential sites identified by diffReps were then used to evaluate the distributions of normalized read counts (reads per kilobase per million mapped reads, RPKM) for each individual DNase-seq sample by two methods: principal component analysis (PCA) and boxplot analysis. Two biological replicates, G74A_M1 and G74A_M2, which EMSA identified as STAT5-high liver samples, showed DNase-seq read count distributions for the top 200 or the top 600 diffReps sites that were more similar to those of STAT5-low livers. Furthermore, one biological replicate, G92_M5, which EMSA identified as a STAT5-low liver sample, showed a read count distribution in the set of differential sites that was more similar to that of the STAT5-high livers. These three outlier liver samples were omitted from all downstream analysis. diffReps was then used, with the same cutoffs described above, to compare the remaining EMSAidentified STAT5-high activity (n = 8) and STAT5-low activity (n = 10) DNase-seq samples. This step yielded a set of diffReps differential sites, which was overlapped with a single MACS2 DHS peak union list, generated by merging the MACS2 DHS peak calls from all n = 18 male mouse liver DNase-seq samples (n = 8 STAT5-high and n = 10STAT5-low livers) using the BEDTools "merge" command. The output of this overlap was a final set of MACS2 DHS peak union list-filtered differential sites, comprised of 2,831 ADHS that open and 123 ADHS that close in STAT5-high compared to STAT5low livers. DHS in the DHS peak union list that did not overlap a diffReps-identified ΔDHS were annotated as static DHS (67,812 static DHS regions). Concatenation of the 67,812 static DHS with the diffReps-identified Δ DHS yielded a combined list of 70,767 DHS peak union sites. This list of 70,767 DHS peak union sites was then used to label the 70,211 mouse liver DHS, described above, based on whether an individual DHS overlapped a STAT5-high ΔDHS ("STAT5 hi dDHS", n = 2,373 liver DHS), a STAT5low ΔDHS ("STAT5 low dDHS", n=98 liver DHS), a STAT5 static DHS

("STAT5_sDHS", n = 45,754), or "none" (n = 21,986) if there was no overlap between the liver DHS and the set of 70,767 DHS peak union list. Further, we applied the previously determined sex bias designation for each of the 70,211 liver DHS used in this study (Ling et al., 2010), namely, male-specific, female-specific, or sex-independent DHS. Members of the set of 70,211 liver DHS were further designated "dynamic" if they overlapped a STAT5-high Δ DHS (i.e. a GH/STAT5 pulse-opened DHS); otherwise, they were designated as "static" with respect to chromatin accessibility changes in response to endogenous GH/STAT5 pulses. Thus, sex-specific DHS were designated dynamic malebiased DHS, static male-biased DHS, dynamic female-biased DHS, or static femalebiased DHS with respect to GH/STAT5-induced chromatin opening. Sex-independent DHS were similarly designated dynamic or static sex-independent DHS.

3.3.6 Impact of hypox and GH treatment on chromatin opening and closing

MACS2 was used to discover differential DHS peaks in DNase-seq datasets obtained for intact male and female mouse liver, for hypox male (MHx) and hypox female (FHx) mouse liver, and for livers of hypox male mice given a single injection of GH (MHx+GH) and euthanized either 30, 90, or 240 min later. The effect of hypox was determined by comparing hypox liver samples to the corresponding intact liver (control) samples in male and female liver. The effect of a single injection of GH was determined by comparing hypox male samples treated with GH to the hypox male samples for each time point. For each comparison, a DHS peak union list was generated by merging the MACS2 DHS peak calls from all individual biological replicates for the corresponding

treatment group. For example, a single peak union list for the male hypox compared to intact male liver samples was generated by merging all the MACS2 peaks from each of the n = 8 individual male mouse liver DNase-seq samples (n = 6 intact male control and n = 2 MHx livers). Genomic regions that showed significantly differential DNase-seq signal between hypox and intact control livers, or between MHx+GH and MHx (control) livers were discovered separately for each comparison using diffReps, using the nucleosome option (200 bp window). Significance was based on |fold-change| > 2 and FDR < 0.05 (Benjamini-Hochberg adjusted p-value) for diffReps-normalized signal intensity values. The diffReps-identified differential sites were then filtered by their overlap with the peak union list for all 8 samples to obtain the sets of ΔDHS shown in Fig. S3.4 (e.g., 2,142 DHS that open and 4,856 DHS that close in male liver in response to hypox). MACS2 DHS peak union sites that did not overlap a diffReps region were annotated as static DHS peaks (50,055 sites). Concatenation of the diffReps-identified Δ DHS with this set of static DHS yielded the full set of 57,053 DHS peak union sites for this data set. Each of the 70,211 liver DHS (see above) was then labeled based on whether it overlapped with a diffReps-identified ΔDHS that opens following hypox ("dDHS open"), a diffReps Δ DHS that closes following hypox ("dDHS close"), a static DHS ("static"), or "none", for DHS on the list of 70,211 that do not overlap the set of 57,053 DHS peak union sites. Corresponding analysis were performed for the intact and hypox female liver samples, and for the MHx GH time-course comparisons mentioned above. Figure S3.4 presents a summary of the numbers of hypox and hypox+GH

responsive ΔDHS for each of these comparisons, and the subsets that overlap the set of 70,211 liver DHS used in this study.

The overlap between STAT5-high Δ DHS that were closed by hypophysectomy and/or opened by a single exogenous GH pulse was determined using a variation of the total flag sum (TFS) procedure (Wauthier et al., 2010). The following data sets: MHx/Male control, MHx+GH/MHx at 30, 90, and 240 minutes were labeled as 1-4 respectively. For each of the 70,211 DHS, a four-digit TFS number was generated where the first digit corresponds to the DHS response in data set 1, the second digit corresponds to the response in data set 2, etc... A binary definition was used to summarize the DHS response in each data set. For data set 1: a value of 1 was assigned if the DHS was closed due to hypophysectomy or 0 otherwise, for data sets 2-4: a value of 1 was assigned if the DHS was opened by a single exogenous GH pulse or 0 otherwise. For each DHS, the binary values were concatenated to generate a four-digit TFS value which concisely summarizes the DHS response in all four data sets. The collection of all TFS values for the set of STAT5-high Δ DHS was used to construct a 4-oval Venn diagram which summarizes the overlap between these data sets.

3.3.7 Sex-specific genes and hypox response classification

Sex-specific genes were identified from RNA-seq data from intact male and female mouse liver based on a gene list comprised of 24,197 RefSeq genes (Connerney et al., 2017). An expanded list of sex-specific genes was generated using both RefSeq and multi-exonic lncRNA genes for this study. RefSeq sex-specific genes were identified from total polyA RNA-seq samples from intact male and female liver using the

"genebody" method of read counting (Connerney et al., 2017) with |fold-change| > 1.5, adjusted p-value < 0.05 and a FPKM > 0.25 for the sex with greater signal intensity, which yielded 387 male-specific and 517 female-specific genes. Sex-specific multiexonic lncRNA genes were identified from nuclear polyA RNA-seq samples in male and female mouse liver using the "ExonCollapsed" read counting method (Melia and Waxman, 2019) with |fold-change| > 2.0, adjusted p-value < 0.05 and a FPKM > 0.25 for the sex with greater signal intensity, which yielded 121 male-specific and 102 femalespecific genes. A stringent set of sex-independent genes was defined by FPKM > 1 in both male and female liver, |fold-change| < 1.2, and adjusted p-value > 0.1, which yielded a total of 7,253 stringently sex-independent genes. Sex-biased genes responsive to hypox were defined by |fold-change| > 2 and an adjusted p-value < 0.05 for hypox versus intact mouse control in male liver, or separately, in female liver. Using these thresholds, RNAseq data from intact and hypox male and female liver samples were used to classify sexspecific genes into classes I and II and their corresponding subclasses (IA, IB, IC, IIA, and IIB) (Connerney et al., 2017). Class I sex-biased genes are those that are downregulated by hypox in the sex where they show higher expression in intact mice. Class II sex-biased genes are those that are upregulated by hypox in the sex where they show lower expression in intact mice. Subclasses A, B, and C indicate the response to hypox in the dominant sex (class II genes) or in the opposite sex (class I genes). See Fig. S3.3 for a summary of gene counts in each class, and see Supplemental Table_S3 for a complete listing of genes, their sex bias, and their hypox response class.

3.3.8 STAT5 binding and motif analysis

ChIP-seq analysis of STAT5 binding in male and female mouse liver previously identified 15,094 merged peaks comprised of male and female enriched and male-female common STAT5 binding sites (Zhang et al., 2012). A small subset of these STAT5 binding sites, 75 peaks, overlapped ENCODE blacklisted regions and were filtered out from downstream analysis. BEDTools was used to determine the overlap between the set of liver 70,211 DHS and these STAT5 binding sites. Each of the 70,211 DHS was designated as either "STAT5 bound" if a STAT5 binding site was present; otherwise the DHS was designated "not bound" (see Supplementary Table_S2, which includes information about the STAT5 binding site, the sex-specificity of STAT5 binding (male_enriched, female_enriched, or common), and the normalized ChIP-seq read counts for STAT5 binding in male STAT5-high activity (MH) and in female STAT5-high (FH) livers, and the calculated average of these read counts. The STAT5B motif M00459 from the TRANSFAC motif database (Release 2011.1) (Matys et al., 2006) was used to determine the frequency of STAT5 motif occurrence in DHS sequences. Motifs present in DHS sequences were identified using FIMO (v4.12.0) (Grant et al., 2011) with the option (--thres 0.0005) to improve detection of short length motifs. The number of STAT5B motif occurrences in each of the 70,211 DHS sequences is shown in Supplemental Table S2.

3.3.9 Enrichment Analysis

For all enrichment calculations described below, the significance of the enrichment was determined by a Benjamini and Hochberg adjusted Fisher's Exact test p-value as implemented in R.

Enrichments of sex-biased DHS for being enhancer, promoter, or insulator regions were calculated (e.g. male-biased DHS for being enhancers compared to the background set of sex-independent DHS) as follows: Enrichment score = ratio A/ ratio B, where: ratio A = the number of sex- biased DHS that are enhancers, divided by the number of sex-biased DHS that are enhancers, divided by the number of sex-biased DHS that are enhancers, divided by the number of sex-biased DHS that are not enhancers; and ratio B = the number of sex-independent DHS that are enhancers, divided by the number of sex-independent DHS that are enhancers, divided by the number of sex-biased DHS that are not enhancers. For example, 2,551 male-biased DHS were classified as enhancers, and 178 other male-biased DHS are not enhancers (2,551/178 = 14.3), whereas 43,591 sex-independent DHS were classified as enhancers (43,591/22,525 = 1.93), which gives an enrichment score A/B = 14.3/1.93 = 7.41. The set of 66,116 sex-independent DHS were used as the background for these enrichment calculations.

Enrichments of sex-biased DHS subsets (enhancer (e), insulator (i), or promoter (p)) for mapping to sex-specific genes were calculated (e.g. male-biased eDHS mapping to malespecific genes) as follows: Enrichment score = ratio A/ ratio B, where: ratio A = the number of male-biased eDHS that map to male-specific genes, divided by the number of male-biased eDHS that map to sex-independent genes; and ratio B = the number sexindependent eDHS that map to male-specific genes, divided by the number of sexindependent eDHS that map to sex-independent genes. For example, among the malebiased DHS classified as enhancers, 404 male-biased eDHS map to male-specific genes, and 525 male-biased eDHS map to sex-independent genes (404/525 = 0.77), whereas 1,495 sex-independent eDHS that map to male-specific genes, and 15,457 sexindependent eDHS map to sex-independent genes (1,495/15,457 = 0.10), which gives an enrichment score A/B = 0.77/0.10 = 7.7. The set of 66,116 sex-independent DHS were used as the background for these enrichment calculations.

Enrichments of hypox-responsive DHS for mapping to class I or class II sex-specific genes were calculated (e.g. DHS that close in response to hypox in male liver and mapping to male class I genes) for male and female liver as follows: Enrichment score = ratio A/ ratio B, where: ratio A = the number of DHS that open (or that close) following hypox and that map to class I (or to class II) sex-specific genes, divided by the number of DHS that open (or that close) following hypox and that map to sex-independent genes; and ratio B = the number of hypox-unresponsive DHS (static DHS) that map to class I or class II sex-specific genes, divided by the number of static DHS that map to sexindependent genes. For example, in male liver, 217 DHS that close following hypox map to a male class I sex-specific gene, and 1,203 other DHS that close map to sexindependent genes (217/1, 203 = 0.1803), whereas 444 static DHS map to a male class I sex-specific gene, and 14,053 static DHS map to a sex-independent gene (444/14,053 =0.0316), which gives an enrichment score A/B = 0.1803/0.0316 = 5.7. The static DHS used for these enrichment calculations correspond to the set of 35,562 static DHS in male liver and 30,394 static DHS in female liver.

Enrichment for overlap with genomic regions containing biologically relevant sets of transcription factor binding sites, chromatin marks and combinations of epigenetic features (chromatin states) was calculated for each of the following four sets of DHS: (1) dynamic male-biased DHS (834 sites), (2) static male-biased DHS (1,895 sites), (3) dynamic sex-independent DHS (1,532 sites), and (4) static female-biased DHS (1,359 sites) relative to a background set of static sex-independent DHS (64,584 sites). Briefly, these sets of DHS were identified based on their sex-specificity (Ling et al., 2010) and were classified as "dynamic" if they overlapped a STAT5-high Δ DHS otherwise they were classified as "static" (see section 3.3.5). Characterization of the 70,211 DHS used in this study as static or dynamic with corresponding sex-specificity designations are listed (Table S2). Biologically relevant regions evaluated include sex-biased transcription factor binding of STAT5 and BCL6 (Zhang et al., 2012), CUX2 (Conforto et al., 2012), and FOXA1/2 (Li et al., 2012). We also examined DHS enrichment for sex-biased chromatin marks and chromatin states defined for male and female liver (Sugathan and Waxman, 2013). FOXA1 and FOXA2 ChIP-seq data (Li et al., 2012) was processed and reanalyzed with MAnorm (Shao et al., 2012), carried out by Gracia Bonilla of this laboratory, to identify sex-biased FOXA1 and FOXA2 binding sites in mouse liver. Data from the MAnorm analysis of FOXA1 and FOXA2 ChIP-seq peaks (Table S8) was filtered by the log2(fold-change) (i.e. M-value) to define male-biased and female-biased binding sites. Coordinates for each of the sex-biased transcription factor binding sites, chromatin marks, and chromatin states are provided (Table S5 and Table S7). Enrichments of DHS for overlapping biologically relevant regions were calculated (e.g. dynamic male-biased

DHS for containing STAT5-High (Male) binding sites compared to the background set of static sex-independent DHS) as follows: Enrichment score = ratio A/ ratio B, where: ratio A = the percentage of dynamic male-biased DHS with a STAT5 binding site; and ratio B = the percentage of static sex-independent DHS with the same type of binding site. For example, 235 dynamic male-biased DHS contain a STAT5-High (Male) binding site (235/834 = 0.2818), whereas 311 static sex-independent DHS contain a binding site (311/64,584 = 0.0048), which gives an enrichment score A/B = 0.2818/0.0048 = 58.7. The set of 64,584 static sex-independent DHS were used as the background for these enrichment calculations. The significance of the enrichment was determined by a Benjamini and Hochberg adjusted Fisher's Exact test p-value as implemented in R. Calculations with adjusted p-value < 1E-03 were considered statistically significant.

3.3.10 Chromatin state map analysis

Chromatin state maps (14 state model) were previously developed for male mouse liver, and separately for female mouse liver, using a panel of six histone marks and DHS data, and used to identify sex differences in chromatin state and chromatin structure and their relationships to sex-biased gene expression (Sugathan and Waxman, 2013). BEDTools was used to determine the overlap between the 14 chromatin states identified in male liver and the set of 70,211 DHS used in this study. Each of the 70,211 DHS was assigned to one of the 14 chromatin states based on overlap and is shown in Supplemental Table_S2. DHS whose genomic coordinates span two or more different chromatin states were assigned to the state with the largest number of overlapping base pairs. Enrichments

of DHS for being in one of the 14 chromatin states in male liver were calculated (e.g. dynamic male-biased DHS for being in chromatin state E6, which is an enhancer-like state) for the set of static and dynamic male-biased DHS as follows: Enrichment score = ratio A/ ratio B, where: ratio A = the number of male-biased DHS not in that are in a particular chromatin state, divided by the number of male-biased DHS not in that chromatin state; and ratio B = the number of static sex-independent DHS in that chromatin state, divided by the number of static sex-independent DHS not in that chromatin state, divided by the number of static as chromatin state. For example, 604 dynamic male-biased DHS are classified as chromatin state E6 (604/230 = 2.626), whereas 24,082 static sex-independent DHS are classified as chromatin state E6 (24,082/40,502 = 0.5946), which gives an enrichment score A/B = 2.626/0.5946 = 4.4. The background used for these enrichment calculations correspond to the set of 64,584 static sex-independent DHS.

3.3.11 DNase-I cut site aggregate plots

Normalized DNase-I cut site aggregate plots were generated using the relevant input DNase-seq dataset and a set of input genomic regions (DHS sequences) used for read counting. First, FASTQ files from DNase-seq biological replicates were concatenated to obtain a single combined replicates file for each treatment group. For example, for male liver, we generated a single STAT5-high combined DNase-seq FASTQ file by merging the n = 8 biological replicate FASTQ files, and separately, we generated a single STAT5-low combined sample by merging the n = 10 biological replicates. Combined replicate

FASTQ files were generated for intact female, hypox female, hypox male, and each of the hypox male + GH treatment time point samples in the same manner. Second, for each combined replicate file, a BED file comprised of positive and negative strand reads was processed to determine the DNase-I cut site, which corresponds to the 5'-end of each sequence read. Third, the set of input genomic regions was processed to generate a listing of 2 kb regions centered at the midpoint of each DHS. BEDTools "coverage" using the (d) option was used to count the number of DNase-I cut sites at each nt position of the 2 kb midpoint-centered regions, thus producing a read count matrix composed of 2,000 read counts for each input genomic region. Fourth, a custom R script was used to load the read count matrix, calculate the sum of read counts at each nt position, normalize the raw read counts by the number of reads in the 70,211 DHS, normalize by the number of input genomic sites, and generate a plot of the DNase-I signal across the input genomic region for the input DNase-seq data set. Normalized DNase-I cutting profiles were then smoothed using a LOWESS smoother as implemented in R (package: gplots v3.0.1.1). An offset was applied to the profile by subtracting an average read count (calculated from the first 200 nt positions) from the normalized read count intensity to standardize each profile baseline. Profiles for other, related DNase-seq data sets were processed in the same manner and were plotted on the same set of axes for direct comparison.

3.3.12 DHS peak normalization

(<u>https://genome.ucsc.edu/</u>) were normalized using the number of sequence reads in each DHS peak region per million mapped sequence reads (reads-in-peaks-per-million,

DNase-seq data to be visualized in the UCSC Genome browser

RiPPM) as a scaling-factor (Lodato et al., 2018). Normalization was carried out using a comprehensive list of DHS peak regions (i.e. the peak union) for each dataset, obtained by concatenating FASTQ files for biological replicates of each control and treatment group, as described above. DHS peak regions identified in both the individual and the combined samples were concatenated into a single list, and then BEDTools "merge" command was used to combine overlapping features to generate a single list of non-overlapping DHS peaks. The fraction of reads in peaks for each sample was then calculated to obtain a scaling factor. Raw read counts were divided by this per-million scaling factor to obtain RiPPM normalized read counts.

3.4 Results

3.4.1 Sex-biased DHS are enhancer regions that target sex-specific genes

Genomic regions associated with set of 70,211 liver DHS, including male-specific, female-specific, and sex-independent DHS (Ling et al., 2010), were classified as enhancers, insulators, and promoters based on their chromatin mark patterns and CTCF binding (Matthews and Waxman, 2018). Sex-biased DHS show significant enrichment for being enhancer DHS (eDHS), with >85% classified as enhancers as compared to 66% for sex-independent DHS. Further, sex-biased DHS were significantly depleted of insulator and promoter regions compared to sex-independent DHS (Fig. 3.1A). To investigate whether sex-biased DHS target sex-specific genes (Supplemental Table_S3), we calculated their enrichment for mapping to male-specific and female-specific genes, as compared to sex-independent DHS, with the target gene of each DHS defined as the closest RefSeq or multi-exonic TSS in the same TAD as the DHS. Significant enrichments were seen for all sets of sex-biased enhancer DHS, with 15 - 24% of eDHS mapping to a sex-specific gene, suggesting distal regulation of gene expression (Fig. 3.1B). The strongest enrichments were obtained for sex-biased DHS classified as promoter DHS; however, only 1-2% of those DHS mapped to a sex-biased gene. Examination of distance between each DHS and its target gene revealed that promoter DHS are the closest in proximity to their target gene TSS, as expected, followed by sexbiased insulator DHS and enhancer DHS, which likely regulate gene expression through distal mechanisms (Fig. 3.1C).

3.4.2 DHS responsive to endogenous pulses of GH and STAT5 in male liver

Sex differences in pituitary GH secretion (pulsatile in males vs. near continuous in females) regulate the sex-dependent expression of hundreds of genes in adult mouse liver. We hypothesized that the pulsatile activation of STAT5 by GH in male liver dynamically alters the chromatin accessibility landscape in male mouse liver. More specifically, pulsatile activation of STAT5 is expected to induce differential chromatin opening and closing at a subset of liver DHS in male mice euthanized at a peak vs. at a trough of liver STAT5 activity. To investigate this hypothesis, we collected DNase-seq data from n = 21 individual male mouse liver samples whose STAT5 DNA-binding activity was determined by EMSA analysis of whole liver extracts. EMSA identified n =10 individual mouse liver samples that have strong STAT5 activity (STAT5-high activity livers) and n = 11 livers with little to no detectable STAT5 activity (STAT5-low activity livers) (Connerney et al., 2017). Principal component analysis of the DNase-seq data using either the top 200 or the top 600 diffReps-discovered differential sites,

determined from a comparison of EMSA-identified STAT5-high versus STAT5-low samples, revealed that two of the STAT5-high livers samples and one of the STAT5-low liver samples were outliers (Fig. 3.2A). These samples were confirmed as outliers when the DNase-seq read count distributions were viewed in box plot format (Fig. 3.2B), and were excluded from all downstream analyses. The remaining n = 8 STAT5-high and n =10 STAT5-low DNase-seq liver samples were used to identify a robust set of n = 2,831opening ΔDHS and n= 123 closing ΔDHS . These sites represent genomic regions that significantly open or close in male mouse liver in response to each endogenous pulse of GH-activated STAT5 signaling. Fig. 3.2C (*left*) visualizes the extent to which chromatin opens at the 2,831 sites, as seen in aggregate plots of DNase-seq cuts in the set of STAT5-high liver samples compared to STAT5-low liver samples. An intermediate level of chromatin opening is seen in female mouse liver. In contrast, at the 123 closing genomic sites, STAT5-low livers show greater chromatin accessibility than the STAT5high livers, consistent with chromatin closing at the latter set of sites (Fig. 3.2C, *right*). Finally, we confirmed the binary pattern of high vs. low DNase-seq read count distributions using an independent set of nine individual male liver DNase-seq samples generated by the ENCODE consortium. Five of the nine livers showed distributions similar to the STAT5-low livers and four livers were similar to the STAT5-high group.

3.4.3 Dynamic and static male-biased DHS

Given the large number sites showing significant differential chromatin opening between STAT5-high and STAT5-low livers, we hypothesized that STAT5 binding is an important driver of the changes in chromatin accessibility seen at male-biased DHS. To

investigate this hypothesis, normalized DNase-I cut site aggregate plots were generated for male-biased DHS that bind STAT5 (1,307 male-biased DHS). These analyses revealed that STAT5-high livers have the highest level of chromatin opening, with STAT5-low males and female liver samples both having ~3-fold lower level of chromatin opening at the set of 1,307 STAT5-bound male-biased DHS (Fig. 3.2D, plot 1). In contrast, the set of 1,422 male-biased DHS that do not bind STAT5 showed near equal chromatin accessibility in STAT5-high compared to STAT5-low male livers, but gave ~2-fold lower accessibility in female liver (Fig. 3.2D, plot 2). Thus, the male-bias of the STAT5-bound male-biased DHS can be explained by the increase in chromatin opening that occurs when STAT5 has been activated by a recent plasma GH pulse. In contrast, the male-biased DHS that do not bind STAT5 are unresponsive to the temporal changes in liver STAT5 activity. Rather, their male bias is apparently due to other factors that lead to relative closing of chromatin of those sites in female liver. This closing of chromatin in female liver for the STAT5-unbound male-biased DHS becomes apparent when the normalized DNase cutting frequency is compared to that of the large number of STAT5unbound sex-independent DHS (n = 53,404 DHS; Fig. 3.2D, plot 4). Sex-independent DHS that bind STAT5 showed much greater chromatin accessibility than sexindependent DHS without STAT5 bound, as seen in both female and male liver samples, independent of the STAT5-high vs. STAT5-low liver status of individual males (Fig. 3.2D, plots 3 vs plot 4).

These findings suggests two distinct subsets of male-biased DHS can be distinguished by their STAT5 binding status: one subset displays male-biased chromatin accessibility due

to the increase in accessibility that occurs upon binding of male GH pulse-activated STAT5; and a second subset of male-biased DHS is constitutively open in male liver (i.e., independent of GH-induced STAT5 pulses) but is comparatively closed in female liver.

3.4.4 Dynamic and static male-biased DHS and the presence of the STAT5 motif Given the above effects of STAT5 binding on chromatin accessibility, we used the differential chromatin accessibility of STAT5-high vs. STAT5-low livers (2,831 opening Δ DHS) to classify the full set of liver DHS as dynamic or static with respect to their responsiveness to the GH-stimulated pulses of STAT5 activity seen male liver (Fig. 3.2E). Strikingly, 31% of male-biased DHS (834 DHS) responded dynamically to endogenous male plasma GH pulses of liver STAT5, whereas only 1 – 2% of femalebiased DHS and sex-independent DHS are dynamic (Fig. 3.2E). Thus, STAT5 pulses are an important feature for a substantial subset of male-biased DHS.

At dynamic male-biased DHS (834 sites), STAT5-high livers showed nearly 4-fold greater chromatin opening than male STAT5-low or female liver samples (Fig. 3.3A, plot 1). In contrast, chromatin accessibility at static male-biased DHS (1,895 sites) was greater in male liver than in female liver, which accounts for the male bias of these genomic regions; however, the overall magnitude of chromatin opening is much lower than that of dynamic male-biased DHS in the STAT5-high samples (Fig. 3.3A, plot 2 vs. plot 1). We also examined chromatin accessibility at the set of static female-biased DHS (1,359 sites), which was >2-fold higher in female than in male liver, both with and without male GH pulses, and can account for the female bias of these genomic regions

(Fig. 3.3A, plot 3). Of note, the aggregate chromatin accessibility of the 1,359 static female-biased DHS in female liver was very similar to that of the bulk of 64,584 sexindependent static DHS (Fig. 3.3A, plot 4 vs. plot 3). This strongly suggests that the female bias of these 1,359 sites is due to chromatin closing in male liver, rather than chromatin opening. Chromatin closing at these sites in male liver is not directly affected by STAT5 binding (Fig. 3.3A, plot 3, STAT5-high vs. STAT5-low livers). We hypothesize that STAT5 binding to its motif, TTC-NNN-GAA, is a key feature that distinguishes dynamic from static male-biased DHS. To investigate this hypothesis, we used FIMO to scan the dynamic and static DHS sequences for the occurrence of the TRANSFAC STAT5 motif (see Methods). A larger percentage of the dynamic DHS (73-86%) contain the STAT5 motif when compared to static DHS (19-38%) (Fig. 3.3B). Closer examination of the motif occurrences revealed that a majority (60-80%) of static DHS do not have a STAT5 motif, whereas, the majority of dynamic DHS have at least one motif occurrence (Fig. 3.3C). MEME-ChIP (using the MEME and DREME algorithms) was used to discover de-novo motifs in the set of dynamic and static DHS. Dynamic male-biased (Fig. S3.2A) and dynamic sex-independent DHS (Fig. S3.2D) each show examples of de-novo motifs with strong sequence similarity to the STAT5B motif. In contrast, none of the motifs discovered from static DHS (male-biased, female-biased, or sex-independent) (Fig. S3.2B, C, E) showed matches for a STAT5B motif.

3.4.5 Dynamic DHS are associated with stronger STAT5 binding and associated chromatin opening

We hypothesize that direct STAT5 binding at dynamic DHS distinguishes them from static DHS and is associated with greater levels of chromatin accessibility. Genome-wide STAT5 binding sites identified in male and female mouse liver (Zhang et al., 2012) were overlapped with the set of 70,211 DHS used in this study. A large majority (81- 86%) of the dynamic DHS contain a STAT5 binding site versus only 18-32% for static DHS (Fig. 3.4A). Further, for male-specific, female-specific, and sex-independent DHS, the subsets of dynamic DHS showed significantly greater STAT5 binding (normalized STAT5 ChIP-seq read counts) than static DHS. These results support the conclusion that direct binding of STAT5 is an important distinguishing feature of dynamic vs static DHS (Fig. 3.4B). Male-enriched STAT5 binding intensity was greater in the set of dynamic than the set of static female-biased DHS. Similarly, female enriched STAT5 binding intensity is greater in the set of dynamic than the set of static female-biased DHS. Dynamic sex-biased and dynamic sex-independent DHS showed greater STAT5 binding intensity than those that are static (Fig. S3.1).

Further examination of normalized DNase-I cut site aggregate plots revealed that STAT5-high livers show the highest degree of chromatin opening in the set of dynamic male-biased DHS with STAT5-low and female samples showing the same baseline level of openness. Dynamic male-biased DHS with STAT5 bound (710 sites) showed greater chromatin opening than the subset of DHS without STAT5 bound (124 sites) (Fig. 3.4C, plot 1A vs. 1B). Male liver samples showed greater chromatin accessibility than female liver samples in the set of static male-biased DHS. Static male-biased DHS with STAT5 bound (597 sites) (Fig. 3.4C, plot 2A) showed STAT5-high samples with the highest level of chromatin opening, whereas DHS without STAT5 binding (1,298 sites) (Fig. 3.4C, plot 2A vs. 2B) showed no difference between STAT5-high and STAT5-low male liver samples. Female liver samples showed the largest degree of chromatin opening compared to male liver samples in the set of female-biased DHS, as expected. Static female-biased DHS with STAT5 bound (258 sites) (Fig. 3.4C, plot 3A) showed greater levels of chromatin opening than the subset of DHS without STAT5 bound (1,101 sites) (Fig. 3.4C, plot 3B). STAT5-high samples show the greatest chromatin opening, followed by female and STAT5-low samples in the set of dynamic sex-independent DHS. Dynamic sex-independent DHS with STAT5 bound (1,239 sites) (Fig. 3.4C, plot 4A) showed greater levels of chromatin opening than the subset of DHS without STAT5 bound (293 sites) (Fig. 3.4C, plot 4B). Female liver samples show slightly greater levels of chromatin accessibility compared to the male liver samples in the set static sexindependent DHS with STAT5 bound (11,473 sites) (Fig. 3.4C, plot 5A); however, male and female samples show the same level of baseline chromatin openness in the set of static sex-independent DHS without STAT5 bound (53,111 sites) (Fig. 3.4C, plot 5B). These results show that dynamic male-biased sites are more open than static male biased sites and the relative levels of chromatin accessibility for female liver distinguishes sexindependent dynamic sites from those that are male-biased.

3.4.6 Chromatin accessibility changes following hypox and impact of a single replacement pulse of GH in male liver

We hypothesize that hypox will lead to closure of many of the open chromatin regions (enhancer DHS) that regulate sex-specific gene expression, and that a single pulse of GH will be sufficient to open a subset of male-biased DHS in hypox male liver. We investigated this hypothesis by performing DNase-seq on liver nuclei from intact female (F) and male (M) mice, hypox female (FHx), and hypox male mice (MHx), and hypox male mice given a single pulse of GH and euthanized 30, 90, or 240 min later. diffReps was used to identify Δ DHS that were responsive to hypox or to GH replacement. We found that similar numbers of liver DHS open as close following hypox of female mice. In contrast, in livers of male hypox mice, many more DHS close than open following hypox (Fig. 3.5A). Furthermore, a single pulse of GH induced chromatin opening at ~ 3,550 sites in hypox male liver after 30 and 90 min; fewer sites respond after 240 min, when the levels of active liver STAT5 substantially returns to baseline (Connerney et al., 2017). The number of diffReps regions and those that overlap the set of 70,211 DHS are shown in Fig. S3.4, with full listings shown in Supplemental Table_S4.

Hypox ablates pituitary GH secretion and consequently the activation of liver STAT5. Since hypox largely abolishes liver sex differences, we hypothesize that DHS that close following hypox include many enhancers that map to sex-specific genes dependent on GH-activated STAT5. Supporting this hypothesis, DHS that close in male liver following hypox are most highly enriched for mapping to class I male-specific genes, and DHS that close following hypox in female liver most highly enriched for class I female-specific genes (Fig. 3.5B). Furthermore, DHS that open in hypox male liver are enriched for class II female-specific genes, whereas DHS that open in hypox female liver are enriched for class II male-specific genes. The pattern of gene responses, shown in dark shades of red (Fig. 3.5B, *bottom*), correspond to the pattern of strong enrichments calculated for the hypox-responsive DHS.

3.4.7 Many endogenous GH pulse-opened DHS are also opened by a single GH pulse in hypophysectomized male mouse liver

Next, we investigated the relationship between Δ DHS responsive to endogenous STAT5 pulses, i.e., STAT5-high and STAT5-low Δ DHS (Fig. 3.6A), and Δ DHS responsive to hypox and GH treatment in the time course study. We hypothesize that many of DHS that respond dynamically to endogenous GH pulses in intact male mouse liver close following hypox and are re-opened by a replacement dose of GH. Supporting this proposal, we found that 1,487 (63%) of the 2,373 DHS that are activated by endogenous GH pulses (i.e., are more open in STAT5-high than in STAT5-low liver), close following hypox. In contrast, only 262 (11%) of these sites close in hypox female liver. Further, 59-62% of the endogenous GH pulse-responsive DHS are re-opened by a single GH pulse after 30 or 90 min; this percentage declines to 23% after 240 min (Fig. 3.6B), when liver STAT5 signaling terminates (Connerney et al., 2017). In contrast, a majority (77%) of the DHS that close in STAT5-high compared to STAT5-low liver are static in their response to hypox in male liver. Furthermore, these sites tend to close rather than open with GH pulse treatment (Fig. 3.6C).

To further investigate the relationship between chromatin accessibility changes due to endogenous GH/STAT5 pulses and a single exogenous GH pulse, we determined the overlap among the subsets of STAT5-high Δ DHS that close due to hypophysectomy (n = 1,487 sites) and/or are opened by a single injection GH in hypox male mice either 30 (n = 1,475 sites), 90 (n = 1,393 sites), or 240 (n = 547 sites) minutes later (Fig. 3.6B). Of the STAT5-high Δ DHS (n = 2,373 sites), the vast majority (83%, 1,974 sites) are responsive to hypox and/or a single exogenous GH pulse while 17% (399 sites) are not responsive (Fig. 3.10). Of the responsive DHS (n = 1,974 sites), 82% (1,620 sites) are opened by an exogenous GH pulse while 18% (354 sites) are closed by hypox but are not opened by a single exogenous GH pulse (set 4 and 5, Fig. 3.10B). Finally, of the GH pulse opened DHS (n = 1,620 sites), 70% are also closed by hypophysectomy in male liver (set 7, Fig. 3.10B).

Of the STAT5-high Δ DHS (n = 2,373 sites), the subset that are responsive to hypox and/or a single exogenous GH pulse (n = 1,974) show greater chromatin opening due to an endogenous GH/STAT5 pulse than the non-responsive subset (n = 399) among pituitary-intact male livers (Fig.3.11A, set 2 vs. set 3). Aggregate profiles of DNase-I cutting at these DHS further validate the large increase in chromatin accessibility in the set of responsive DHS compared to non-responsive DHS (Fig. 3.11B, set 2 vs. set 3). Of the responsive DHS (n = 1,974 sites), the subset that are opened by an exogenous GH pulse (n = 1,620 sites) show greater chromatin opening due to an endogenous GH/STAT5 pulse than the subset not opened by a single exogenous GH pulse (n = 354) (Fig. 3.11B, set 4 vs. set 5), also confirmed by DNase-I aggregate cutting profiles (Fig. 3.11B, set 4 vs. set 5). Of the GH pulse opened DHS (n = 1,620 sites), the subset that are closed by hypophysectomy (n = 1,133 sites) show greater chromatin opening due to an endogenous GH/STAT5 pulse than the subset that do not close by hypophysectomy (n = 487 sites) (Fig.3.11A, set 6 vs. set 7), also confirmed by DNase-I aggregate cutting profiles (Fig. 3.11B, set 6 vs. set 7). Further, the DHS that are closed by hypophysectomy and are opened by an exogenous GH pulse at any of the three time points (30, 90, or 240 minutes) (n = 1,133 sites) show the largest magnitude (Fig. 3.11A) and relative levels of chromatin opening (Fig. 3.11B) due an endogenous GH/STAT5 pulse when compared to the other DHS subsets (sets 1-6 vs. set 7).

3.4.8 A single GH pulse is sufficient for opening chromatin in hypophysectomized male mice

Given the majority of DHS that open by an endogenous GH pulse are also opened by a single pulse of GH in hypox male liver, we examined the relative intensity of chromatin opening for intact female, hypox female (FHx), hypox male (MHx), and hypox male liver following GH pulse treatment. For the DHS that open (2,831 sites), male hypox samples treated with GH either 30 or 90 min show the greatest intensity of chromatin opening, followed by baseline levels of chromatin opening for MHx+GH 240 min, and lower levels for hypox female and hypox male livers, respectively (Fig. 3.7A, *left*). For DHS that close (123 sites), male hypox liver samples treated with GH either 30 or 90 min shows the greatest mit of H either 30 or 90 min shows the streated with GH either 30 or 90 min shows the livers, respectively (Fig. 3.7A, *left*). For DHS that close (123 sites), male hypox liver samples treated with GH either 30 or 90 min showed the lowest levels of chromatin accessibility, while all other samples showed baseline levels of chromatin opening (Fig. 3.7A, *right*). Thus, an exogenous GH pulse

given to hypox male mice is sufficient to induce chromatin opening within 30 min to a level similar to that of intact male mouse liver experiencing a natural, endogenous GH pulse (Fig. 3.2C). The effects of GH decrease after 90 min, leading to the lower level of chromatin accessibility seen in MHx+GH 240 treatment.

Relative levels of chromatin accessibility were also examined in the set of male-biased and sex-independent DHS separated by those with or without STAT5 bound. For malebiased DHS with STAT5 bound (1,307 sites), male hypox samples treated with GH either 30 or 90 min show the most chromatin opening intensity followed by baseline levels of chromatin opening for MHx+GH 240 min later, intact-female, and lower levels for hypox female and hypox male livers respectively (Fig. 3.7B, *top-left*). These samples indicate the chromatin to be essentially closed at the set of male-biased DHS without STAT5 bound (1,422 sites) (Fig. 3.7B, *bottom-left*). Sex-independent DHS with STAT5 bound (12,712 sites) show high levels of chromatin opening for the intact female and hypox male samples treated with GH with decreased accessibility seen for hypox female and hypox male samples respectively (Fig. 3.7B, *top-right*). For sex-independent DHS without STAT5 bound, baseline levels of chromatin opening was seen for all samples (Fig. 3.7B, *bottom-right*).

3.4.9 A single GH pulse induces robust chromatin at dynamic DHS in hypophysectomized male mouse liver

The relative levels of chromatin accessibility were examined in the set of dynamic and static DHS for the GH time course study dataset. In the set of dynamic male-biased DHS (834 sites): male hypox mice treated with GH either 30 or 90 min show the largest degree

of chromatin opening, followed by a decrease back toward baseline levels of chromatin opening after 240 min (Fig. 3.8A, plot 1). Importantly, the level of chromatin opening achieved in the GH pulse-treated MHx livers was indistinguishable from that seen for the same set of 834 DHS in STAT5-high male liver (see Fig. 3.3, plot 1). This indicates that a single exogenous pulse of GH recapitulates the effects of an endogenous GH pulse on chromatin opening in male mouse liver, and further, establishes that liver chromatin at these 834 sites is primed, and can respond rapidly (within 30 min) to a GH pulse even after several weeks of GH deficiency conferred by hypox. In contrast to the dramatic effects of GH pulse stimulation seen at the set of 834 dynamic male-biased DHS, the set of 1,895 static male-biased DHS showed only small increase in chromatin opening (Fig. 3.8A, plot 2), while the set of 1,359 female-biased DHS showed small decreases in chromatin opening in GH-treated hypox male liver (Fig. 3.8A, plot 3). Further, the set of 64,584 static sex-independent DHS were unresponsive to plasma GH stimulation (Fig. 3.8A, plot 4).

Next, we investigated the hypothesis that STAT5 binding is directly linked to the increased chromatin accessibility at the dynamic male-biased DHS. Male hypox mice treated with GH either 30 or 90 min showed the largest degree of chromatin opening in the set of 710 dynamic male-biased DHS that bind STAT5. Moreover, chromatin opening decreased substantially after 240 min (Fig. 3.8B, plot 1A), at which time the activation of liver STAT5 DNA binding activity is fully reversed (Connerney et al., 2017). Smaller increases in chromatin opening were observed with GH pulse treatment at the subset of 124 dynamic male-biased DHS that did not bind STAT5 (Fig. 3.8B, plot 1B), consistent

with their dynamic responses to endogenous GH pulsation, and suggesting that chromatin opening at these sites proceeds by a distinct mechanism than at the STAT5-bound dynamic sites.

Static male-biased DHS with STAT5 bound (597 sites) showed a more modest increase in chromatin opening with GH pulse treatment (higher basal level in MHx control and lower induced level with GH pulse; Fig. 3.8B, plot 2A), while static male-biased DHS without STAT5 binding (1,298 sites) showed little or no GH pulse responsiveness (Fig. 3.8B, plot 2B).

Chromatin opening at female-biased DHS was decreased by hypox, both at the 258 sites bound by STAT5 and at the 1101 sites that did not show STAT5 binding (Fig. 3.8B, plots 3A and 3B). Further, GH pulse treatment of MHx mice stimulated a modest decrease in chromatin opening at both sets of female-biased DHS. Finally, the dynamic, but not the static sex-independent DHS showed large increases in chromatin opening with GH pulse treatment (Fig. 3.8B, plots 4A and 5A), similar to the dynamic male-biased DHS. Overall, DHS that bind STAT5 showed higher levels of chromatin opening than DHS that do not bind STAT5 (Fig. 3.8B, *top* row vs. *bottom* row). Notably, this difference in chromatin accessibility is preserved in hypox male and female liver, even though liver STAT5 is inactive and cannot bind DHS under these conditions due to the absence of GH stimulation.

3.4.10 Chromatin state differences between static and dynamic male-biased DHS

Next, we investigated the hypothesis that the underlying chromatin state, defined by activating and repressive chromatin marks and open chromatin regions, is different

between dynamic and static DHS, particularly for male-biased DHS. Using the chromatin mark and CTCF-based classification of liver DHS described previously (Matthews and Waxman, 2018), 92-96% of dynamic and static male-biased DHS were classified as enhancers, which a somewhat larger fraction of weak enhancers in the case of static male-biased DHS (Fig. 3.9A). Promoter DHS and insulator DHS comprise the balance of each male-biased DHS set (2-5% each). Similarly, 86% of female-biased DHS and 90% of dynamic sex-independent DHS were enhancers or weak enhancers. By contrast, a much larger set of static sex-independent DHS are insulator or promoter regions (15-20%) compared to dynamic sex-independent DHS (3-8%) (Fig. 3.9A).

Next, we examined the chromatin state distribution of the dynamic and static male-biased DHS using the 14-emission chromatin state map developed for male mouse liver (Sugathan and Waxman, 2013). The majority (59-72%) of dynamic and static malebiased DHS were classified as being in chromatin state E6, whose emission parameters indicate high frequency of DHS with the activating chromatin marks H3K27ac and H3K4me1 (Fig. 3.9B). Much smaller percentages of both male-biased DHS sets were in: state E5, an enhancer state largely marked by DHS without activating enhancer marks; state E7, a promoter state; E11, an enhancer state characterized by H3K4Me1 but not other marks; and E12, a bivalent state associated with both activating (H3K4me1) and repressive marks (H2K27me3). No major differences in chromatin state distributions were seen for dynamic vs. static male-biased DHS.

Next, we investigated potential differences between dynamic and static male-biased DHS, dynamic sex-independent DHS, and static female-biased DHS, with respect to sex-

biased transcription factor binding, chromatin marks and chromatin states. Male-biased DHS showed enrichment for male-biased STAT5 binding, and correspondingly, femalebiased DHS showed enrichment for female-biased STAT5 binding (Fig. 3.9D, left). The strongest enrichment was seen for the set of dynamic male DHS (ES: 58.51, p-value: 2.7E-302), which implicates STAT5 binding is an important aspect of chromatin opening/closing at these genomic regions. More modest enrichments were seen for two STAT5-regulated repressors: Bcl6 in the sets of dynamic male-biased and dynamic sexindependent DHS, and for Cux2, most significantly at static male-biased DHS (Fig. 3.9D, *middle*). In general, sex-biased FoxA1 and FoxA2 binding sites were more strongly associated with DHS of the same sex bias, with the strongest enrichments seen for FoxA2 binding (Fig. 3.9D, *right*). Male-biased DHS – both dynamic and static – showed strong enrichment for male-biased enhancer chromatin marks (H3K27ac and H3K4me1) in male liver, while female-biased DHS showed strong enrichment for a repressive chromatin mark (H3K27me3) in male liver (Fig. 3.9E, left). Female-biased DHS also showed strong enrichment for female-biased enhancer chromatin marks (H3K27ac and H3K4me1) in female liver, while male-biased DHS showed enrichment for a repressive chromatin mark (H3K9me3) in female liver (Fig. 3.9E, right). Analysis of chromatin states revealed that male-biased DHS were enriched for enhancer chromatin states found in male liver, with chromatin state E6 yielding the most significant enrichment among male-biased and sexindependent DHS but not female-biased DHS and chromatin state E7, a promoter-like state, being significantly depleted (Fig. 3.9F). Female-biased DHS showed strong enrichments for inactive chromatin states (E1 and E2) in male liver (Fig. 3.9F). All four

sets of DHS showed depletion of promoter-like states (E7 and E8) in both male and female liver, which further supports the notion of these sites being enhancer regions (Fig. 3.9F, Fig. 3.9G).

3.5 Discussion

GH secretion profiles and corresponding activation of STAT5 is an important driver of sex-specific gene expression and sex-biased chromatin accessibility. In male liver, pulsatile GH activation of STAT5 is responsible for male-biased gene expression, but little is known about mechanisms controlling male-biased DHS. Studies of individual genes (Connerney et al., 2017) provide some insight into chromatin accessibility changes at specific genomic loci in response to GH stimulation; however, little is known about genome-wide epigenetic changes in mouse liver. Here, we used a mouse liver model to identify several thousand genomic regions whose chromatin accessibility responds to either an endogenous or exogenous GH pulse, which was given by injection to hypox (i.e., GH-deficient) young adult male mice. We discovered two distinct mechanisms for GH regulation of male-biased DHS. Dynamic male-biased DHS (n=834) are almost all bound by STAT5; they were shown to undergo repeated cycles of chromatin opening and closing stimulated by each GH pulse as the primary determinant of their male-biased accessibility. In contrast, chromatin accessibility at static male-biased DHS (n=1,895) does not vary with each pulse of GH/STAT5; rather, the male bias of those DHS is due to their being more constitutively closed in female liver, as shown by DNase cutting frequency analysis. Both male-biased DHS classes are ~95% enhancer DHS. Preliminary
findings show that both male-biased DHS classes show strong enrichment for mapping to male-biased genes in the same TAD.

3.5.1 Sex-biased DHS are enhancer regions with distal gene targets

Sex differences in mouse liver gene expression are widespread (Waxman and Holloway, 2009)and are under the control of GH and its sexually dimorphic pattern of secretion by the pituitary gland. GH regulates liver chromatin states, including chromatin accessibility, which both show major sex differences in localized regions throughout the genome (Ling et al., 2010; Sugathan and Waxman, 2013). Data for chromatin marks, binding of CTCF and cohesin, and proximity to RefSeq TSS in the same TAD were used to functionally classify the set of ~70,000 DHS discovered in male and female mouse liver (Matthews and Waxman, 2018). Approximately 95% of the ~4,000 GH-regulated sex-biased DHS are enhancer DHS, which are distal to the nearest gene TSS on the same TAD. Sex-biased promoter DHS showed the strongest magnitude of enrichment for mapping to sex-specific genes and were closest in proximity to their putative gene targets; however, they represent a small fraction (1-3%) of the sex-biased DHS (Fig. 3.1A). These results show that sex-biased DHS function as enhancer regions which target genes from a distance rather than acting in close proximity to gene TSSs.

3.5.2 Endogenous pulses of GH are responsible for chromatin accessibility changes in male mouse liver

We used DNase-seq to identify genomic regions (Δ DHS) where high or low STAT5 activity – which directly reflects recent stimulation by a plasma GH pulses – induces a change (an increase or a decrease) in male mouse liver chromatin accessibility. Although EMSA analysis was used to initially identify intact-male liver samples as showing STAT5 high or low activity, 3 of the 21 liver samples analyzed showed DNase-seq profiles that conflicted with the other biological replicates and appeared to be more similar to replicates of the opposite treatment group. One possible explanation of this discrepancy is that these male mice may have been euthanized very soon after either the beginning or the end of a STAT5 pulse, with more time needed to re-establish the chromatin accessibility profile at STAT5-responsive sites. The remaining STAT5-high and STAT5-low livers were used to discover regions of differential chromatin accessibility, ΔDHS . Genomic regions that showed the strongest increase in chromatin opening (top 200 sites, Fig. 3.2B), showed clear separation between STAT5-high and STAT5-low livers, and also separated a panel of n = 9 male liver replicates from the ENCODE consortium. This examination of biological replicates and classification of ENCODE samples strongly suggest these Δ DHS regions are robust in responding to endogenous GH pulses, also seen from normalized DNase-I cutting profiles of these samples compared to female liver (Fig. 3.2C).

3.5.3 Male-biased DHS can be classified into GH-pulse dynamic and static sites STAT5 is essential for establishing and maintaining male liver sex-specific gene expression, with 90% of male-biased genes down regulated to female-like levels, and 61% of female-biased genes upregulated to normal female liver levels in adult male STAT5b-deficient mice (Clodfelter et al., 2006). We also showed that sex-biased STAT5 binding sites are strongly enriched at sex-biased DHS and in proximity to sex-biased genes (Zhang et al., 2012). This raises the question of whether STAT5 binding at sexbiased DHS, particularly at male-biased DHS, is associated with the changes in chromatin accessibility that we describe between STAT5-high and STAT5-low male livers, and also between male and female livers. Of the male-biased DHS that we investigated (2,729 sites), approximately half are bound by STAT5 in male mouse liver (1,307 sites, 48%) and half are unbound (1,422 sites, 52%) by STAT5 (Fig. 3.2D). Strikingly, the STAT5-high male liver samples showed the highest levels of chromatin accessibility among the male-biased DHS with STAT5 bound. Male-biased DHS that lack STAT5 binding show moderate levels of chromatin opening in male vs. closed chromatin in female liver samples, thus explaining the male-bias of these DHS. Similarly, among the sex-independent DHS, much higher levels of chromatin accessibility are seen in the subset of DHS with STAT5 bound as compared to those without STAT5 binding (Fig. 3.2D).

The striking association between induced chromatin opening and STAT5 binding at male-biased DHS raises the question of whether the pulsatile plasma GH pattern stimulation of male liver is responsible for chromatin remodeling of male-biased DHS. We investigated this question by classifying the sex-biased DHS based on overlap with STAT5-high opened Δ DHS (termed dynamic) or those that lack overlap (termed static). This classification revealed two distinct mechanisms for GH regulation of male-biased DHS: those that are "dynamic", which undergo repeated cycles of chromatin opening and closing following each plasma GH pulse; and those that are "static", where chromatin accessibility does not vary in response to a pulse of GH-activated STAT5, but whose

male bias in accessibility is apparently due to these sites being more closed in female liver (Fig. 3.3A).

3.5.4 STAT5 binds directly to dynamic DHS

Dynamic DHS are genomic regions which open/close in response to each GH pulse; however, it was not known whether GH activation of STAT5 contributes to the observed changes in chromatin structure, or whether those changes are due to STAT5-independent mechanisms. We hypothesized that STAT5 binding at these dynamic DHS, which undergo repeated cycles of chromatin opening and closing following each GH pulse, is a key characteristic of these sites. We investigated this hypothesis by scanning the DHS sequences for the TRANSFAC STAT5 motif. The majority of dynamic DHS contain one or more instances of the STAT5 motif, whereas, the majority of static DHS lack the STAT5 motif. We also performed de-novo motif analysis of these DHS sequences and found that dynamic DHS yielded motifs with strong sequence similarity to the STAT5 motif, while no motif matches for STAT5 were found among the static DHS sequences (Fig. S3.2). Given these differences between the dynamic and static DHS sequences, we anticipated finding preferential binding of STAT5 at the dynamically-responsive DHS. Indeed, using genome-wide binding site data for STAT5 (Zhang et al., 2012), we found that the majority of dynamic DHS are STAT5 bound , while the majority of static DHS lack STAT5 binding sites. Further, the overall intensity of STAT5 binding is significantly greater at the set of dynamic vs. static DHS (Fig. 3.4B and Fig. S3.1). Thus, STAT5 binding is a key distinguishing characteristic that separates dynamic from static DHS. Furthermore, the presence of STAT5 binding is associated with greater levels of

chromatin accessibility at dynamic vs. static DHS (Fig. 3.4C). Taken together, these findings identify direct binding of STAT5 as an important mechanism for chromatin opening at dynamic DHS, which may involve the recruitment of chromatin modifying proteins (e.g. FOX factors) or changes in the chromatin state of these regions.

3.5.5 Several thousand genomic regions respond to a single pulse of GH in hypophysectomized male mouse liver

We used DNase-seq to investigate the impact of a single pulse of GH given to hypox mice on local chromatin accessibility, which we determined 30, 90, 240 min after hormone treatment. Comparison of hypox male and hypox female to pituitary-intact controls identified ~5,000 DHS that open or close following hypox. Further, GH pulse treatment stimulated a rapid increase (within 30 min) in chromatin opening at ~3,500 sites and induced chromatin closing at 700 other sites. Hypox-responsive ΔDHS were mapped to the closest RefSeq or multi-exonic lncRNA to determine the relative proximity of these sites to sex-specific genes (Fig. S3.3). DHS that close in hypox male liver showed the strongest enrichment for class I male-specific genes (ES: 5.7, p = 1.3E-69), while DHS that close in hypox female liver showed the strongest enrichment for class I female-specific genes (ES: 7.8, p = 1.69E-21) (Fig. 3.5B). These enrichments for DHS that close are consistent with the definition of class I sex-biased genes: those genes are down regulated by hypox in the sex where they show higher expression intact mice. Further, ΔDHS that open in hypox male liver were enriched for class II female-specific genes (ES: 3.6, p = 1.80E-15) and ΔDHS that open in female liver were enriched for class II male-specific genes (ES: 1.9, p = 4.8E-04) (Fig. 3.5B). These enrichments are

consistent with the definition of class II sex-biased genes, i.e. genes that are up regulated by hypox in the sex where they show lower expression in intact mice. These results further substantiate the functional importance of chromatin accessibility as a positive regulator of sex-specific gene expression in mouse liver.

We also examined the overlap between DHS that open due to endogenous GH/STAT5 pulses and Δ DHS from the GH time-course dataset. We hypothesized that genomic regions that open due to endogenous GH/STAT5 pulses in male liver are also sites that close by hypophysectomy and that a single pulse of GH is sufficient to restore chromatin accessibility at these sites. Of the 2,373 DHS that open, due to endogenous GH/STAT5 pulses in male liver, ~60% of them close following hypox, ~60% are opened by GH at 30 and 90 min, and 20% remain open after 240 min. Together, these results demonstrate that a high proportion of DHS that open, due to natural GH/STAT5 pulses in male liver, are indeed closed by hypox and then re-opened by a single pulse of GH given to hypox male mice. The effect of a single pulse of GH are short-lived, as shown by the decreasing numbers of opening DHS seen after 30, 90 and 240 min (Fig. 3.6B). Additionally, DHS that close due to an endogenous GH/STAT5 pulse are mostly static (i.e., unresponsive) to hypox or GH treatment (Fig. 3.6C), indicating that GH/STAT5 pulses are primarily associated with chromatin opening rather than closing in mouse liver.

The overlap among the subset of the 2,373 DHS that open due to endogenous GH/STAT5 pulses in male liver that also close due to hypophysectomy and/or are opened by a single exogenous GH pulse was determined. Of the 2,373 DHS, a large majority (n = 1,974

sites, 83%) of them are responsive to hypox and/or a single exogenous GH pulse (Fig. 3.10). The subset of unresponsive DHS (n = 399 sites, 17%) represent DHS which pass thresholds for being STAT5-high/STAT5-low Δ DHS but are generally weaker than the other DHS subsets identified (Fig. 3.11). Interestingly, of the 1,974 responsive DHS; 354 sites are closed by hypophysectomy and are not opened by a single GH pulse in hypox male livers. Although these sites are opened by GH/STAT5 pulses in pituitary-intact male livers, they are unresponsive to a single GH pulse which indicates these genomic regions may require additional GH pulses or other epigenetic changes (e.g. transcription factor binding or histone marks) to reopen chromatin at these regions. Of the DHS that are opened by an exogenous GH pulse in hypox male liver (n = 1,620 sites), the majority of them (n = 1,133 sites, 70%) are closed by hypophysectomy. These 1,133 sites represent GH-responsive DHS since they are (1) opened by GH/STAT5 pulses in pituitary-intact male livers, (2) closed by hypophysectomy, and (3) reopened by a single exogenous pulse of GH in hypox male livers (Fig. 3.10). These 1,133 sites also show the largest increase in chromatin accessibility between STAT5-high and STAT5-low male livers (set 7, Fig. 3.11).

3.5.6 A single GH pulse is sufficient for short-term restoration of chromatin accessibility in hypox male mice

We used a hypox male mouse model to assess the relative levels of chromatin accessibility due a single pulse of GH, one that approximates a physiological replacement dose in mouse liver (Connerney et al., 2017). We found that DHS that open due to endogenous GH/STAT5 pulses were also opened by a GH pulse (Fig. 3.7A). A similar pattern was observed for the subset of male-biased DHS with STAT5 bound (1,307 DHS, Fig. 3.7B) which closely resembles the profile of chromatin opening of STAT5high/STAT5-low male livers (Fig. 3.2D). A single GH pulse was sufficient for restoring chromatin accessibility after 30 or 90 min in the set of dynamic male-biased DHS (834 DHS), with similar profiles seen among intact-male liver samples (Fig. 3.3A). A similar pattern was seen for the subset of dynamic male-biased DHS with STAT5 bound (Fig. 3.8B, plot 1A) compared to intact male liver samples (Fig. 3.4C, plot 1A). Interestingly, the highest level of chromatin opening was seen for the set of dynamic sex-independent DHS 30 or 90 min after a single GH pulse. As described for intact male liver samples, hypox male liver samples treated with GH showed greater chromatin accessibility in genomic regions with STAT5 bound compared to those without STAT5 bound (Fig. 3.8B top vs. bottom). These findings show that a single, physiological pulse of GH is sufficient to restore chromatin opening at GH-responsive DHS with an eventual loss of chromatin accessibility seen at 240 min later, which coincides with the deactivation of liver STAT5 seen in the same set of livers (Connerney et al., 2017), and mimics the closing of dynamic DHS seen in STAT-low compared to STAT5-high livers. Notably, a subset of DHS that open due to endogenous GH/STAT5 pulses are not opened by a single pulse of GH. This indicates that these genomic regions may require additional GH pulses and potentially other epigenetic changes to restore regulatory function.

3.5.7 Chromatin state differences between dynamic and static male-biased DHS

Sex-specific GH secretion profiles establish and maintain sex-specific chromatin accessibility; here, we described two distinct mechanisms for GH regulation of malebiased DHS. We found that the male bias of dynamic DHS (834 sites), which represent ~ 30% of all male-biased DHS, can be explained by their repeated opening induced by each pulse of GH-activated STAT5. Notably, GH-induced opening of these genomic sites does not occur in female liver, where STAT5 is more persistently activated by the near continuous presence of GH in circulation. The sex specificity of static male-biased DHS (1,895 sites), which represent ~70% of all male-biased DHS, is regulated by a different mechanism, insofar as those sites are unresponsive to plasma GH pulses. The male bias of those sites is apparently due to their actively being closed in female liver – which is supported by aggregate DNase cutting profiles showing those DHS to be substantially more closed in female liver than the large bulk of sex-independent static DHS (Fig. 3.3A, plot 2 vs. plot 4). It is unclear what features, other than STAT5 binding, distinguish dynamic and static male-biased DHS from each other and from female-biased DHS. Chromatin state profiles were very similar for dynamic and static male-biased DHS (Fig. 3.9).

We hypothesize that sex-biased transcription factor binding is an important distinguishing characteristic between dynamic and static male-biased DHS. Dynamic male-biased DHS showed > 2-fold greater enrichment for containing STAT5-high binding sites in male liver compared to static male-biased DHS (ES: 58.51 for dynamic DHS vs. ES: 23.78 for static DHS, Fig. 3.9D). Much weaker enrichments for sex-based STAT5 binding were

seen for the sets of dynamic sex-independent DHS and static female-biased DHS. Thus, male-biased STAT5 binding is closely associated with dynamic male-biased DHS, consistent with the greater prevalence of a STAT5 motif and STAT5 binding at dynamic compared to static DHS (Fig. 3.3 and Fig. 3.4). Interestingly, GH/STAT5-regulated repressors BCL6 and CUX2 showed varying levels of enrichment in the set of male and dynamic sex-independent DHS. STAT5 and BCL6 play counteractive roles in regulating sex-specific gene expression in mouse liver; BCL6 binding is preferentially associated with repression of female-biased STAT5 targets in male liver (Zhang et al., 2012), but here we did not see enrichment at female-biased DHS.

Dynamic male-biased and dynamic sex-independent DHS showed the largest magnitude of enrichment for containing male-specific BLC6 binding sites. Although these sites are opening/closing in response to GH/STAT5 pulses, BCL6 binding at these sites maybe a mechanism for repressing female-specific genes in male liver. Analogously, CUX2 is a highly female-specific liver transcription factor, further, its binding in female liver is enriched at male-biased DHS and nearby male-enriched STAT5 binding sites (Conforto et al., 2012). Dynamic and static male-biased DHS showed enrichment for female-specific CUX2 binding sites, consistent with CUX2 functioning as a repressor of male-biased genes in female liver. We also investigated the role of the pioneer factors FOXA1 and FOXA2 at these DHS. Both FOXA factors showed ~2-fold greater enrichment of their male-biased binding sites at static male-biased compared dynamic male-biased DHS (for FOXA1: ES = 6.09 vs. 3.74; and for FOXA2: ES = 44.3 vs. 21.2, for static compared to dynamic male-biased DHS). Although dynamic male-biased DHS may require FOX

factor binding for chromatin remodeling in response to GH/STAT5 pulses, static malebiased DHS may require FOX factor binding for maintaining open chromatin at these sites. Similarly, static female-biased DHS showed the largest enrichment for containing female-specific FOXA binding sites. In terms of sex-specific chromatin marks, dynamic and static male-biased DHS show similar patterns of enrichment. Static male-biased DHS showed greater enrichment for containing H3K27ac (ES: 25.3 for dynamic DHS vs. ES: 41.5 for static DHS) and H3K36me3 (ES: 3.87 for dynamic DHS vs. ES: 25.6 for static DHS) in male liver. These results may, in part, be a consequence of different chromatin structure at dynamic vs. static male-biased DHS or maybe due to differential levels of target gene expression of these DHS. Among chromatin marks in female liver, dynamic male-biased DHS showed a ~2-greater enrichment for containing H3K9me3 sites (ES: 21.68 for dynamic DHS vs. ES: 10.91 for static DHS). These results indicate that the H3K9me3 mark is preferentially used in female liver to repress male-biased DHS activity in female liver. This pattern is opposite to that of H3K27me3, whose male-biased binding is strongly enriched at female-biased DHS in male liver, and is used to repress femalebiased gene expression in male liver. Finally, in terms of chromatin states, dynamic and static male-biased DHS showed similar patterns of enrichments when compared to the set of static sex-independent DHS. This result is consistent with the chromatin state profiles described in male liver above (Fig. 3.9B).

Fig. 3.1 Classification of liver DHS and mapping to liver-expressed genes.

A. Shown are the percentages of each indicated liver DHS subset (x-axis) that are classified as weak enhancer, enhancer, insulator or promoter regions (Matthews and Waxman, 2018). Robust male-biased and robust female-biased DHS (Ling et al., 2010), are subsets of all male-biased and female-biased DHS, respectfully. Enrichments of sexbiased DHS for being enhancer, insulator or promoter regions were compared to a background set comprised of sex-independent DHS (66,116 sites). Significance of enrichments were determined by Fisher Exact Test with Benjamini-Hochberg p-value adjustment: *, P<0.01; **, P<1e-10; ***, P<1e-50. Asterisks in black indicate enrichment, and asterisks in red indicate depletion compared to sex-independent DHS. B. Enrichment of sex-biased DHS sets (as shown in A) for mapping to sex-specific genes in the same TAD, as compared to a background set of sex-independent DHS mapping to the same gene set. Bar height indicates enrichment score (ES); numbers above bars specify the percentage of each sex-biased DHS set that maps to a male-specific or a femalespecific gene, respectively. 15-24% of sex-biased enhancer DHS, but only 1-3% of sexbiased promoter and insulator DHS map to sex-biased genes. C. Cumulative frequency distribution of the distance to the nearest TSS on the same TAD for male-biased and female-biased enhancer (e), insulator (i), and promoter (p) DHS.



Figure 3.1. Classification of liver DHS and mapping to liver expressed genes

Fig. 3.2Discovery and characterization of dynamic and static male-biasedDHS.

A. Principal component analysis of the distributions of DNase-seq reads per kilobase per million (RPKM) for the set of diffReps differential sites that are more open in STAT5high compared to STAT5-low livers. Eigenvector values for principal component 1 (PC1) are shown for the individual STAT5-high (blue) and STAT5-low liver samples (green) calculated for the 600 (left) or the 200 (right) diffReps differential sites showing the largest fold-change in chromatin opening. The dotted red line indicates an empirical cutoff used to separate STAT5-high from STAT5-low liver samples, and the dotted black circles mark outlier samples in each dataset. B. Boxplots of the log2(RPKM) DNase-seq signal in the top 200 DHS that open (as in panel A) for the sets of STAT5-high and STAT5-low liver samples, and for individual male mouse liver samples generated by the ENCODE consortium (replicates 5 to 13), as marked on x-axis). The dashed black line indicates an arbitrary cutoff used to separate STAT5-high and STAT5-low liver samples. Samples from each group that do not pass the cutoff (first two red bars, and seventh blue bar, from the left) are the same outlier livers identified in panel A. C. Normalized DNase-I cut site aggregate plots for STAT5-high (red) and STAT5-low male livers (blue), and for female livers (black). Cut sites were aggregated across the sets of diffReps-identified DHS that open (left) or close (right) in response to endogenous STAT5 pulses in male liver, i.e., sites that show greater diffReps normalized DNase-seq signal intensity in STAT5-high compared to STAT5-low livers. **D.** Normalized DNase-I cut site aggregate plots for STAT5-high (red) and STAT5-low male livers (blue) and female livers (black)

across the genomic regions included in the sets of male-biased (plots 1, 2) and sexindependent DHS (plots 3, 4), separated into subsets of DHS either with (plots 1, 3) or without STAT5 bound (plots 2, 4). **E.** Overlap between endogenous STAT5 pulse-opened DHS identified by diffReps (2,831 sites) and the sets of sex-biased and sex-independent DHS, defining static and dynamic DHS subsets.



Figure 3.2. Classification of male biased DHS as dynamic and static sites

Fig. 3.3 Dynamic and static male-biased DHS.

A. Normalized DNase-I cut site aggregate plot for male livers at STAT5-high, STAT5low, and for female liver for the set of static and dynamic male-biased DHS, static female-biased, and static sex-independent DHS, as identified in Fig. 3.2E. **B.** Bar plot showing the number and percent of static and dynamic DHS with one or more occurrences of the STAT5 motif. FIMO was used to scan each of the 70,211 liver DHS sequences to determine the number of TRANSFAC STAT5 motif occurrences. **C.** The percentage of dynamic (*left*) and static (*right*) DHS with zero or more occurrences of the STAT5 motif. Shown on the x-axis: the number of motif occurrences in each DHS (n =0-9), y-axis: % DHS in each group with the corresponding number of motifs.



Figure 3.3. Dynamic and static male biased DHS

Fig. 3.4 STAT5 binding is associated with chromatin opening.

A. The number and percentage of dynamic and static DHS that have one or more STAT5 binding sites, based on STAT5 ChIP-seq data for mouse liver. BEDTools was used to determine the overlap between the merged list of 15,094 STAT5 binding sites (Zhang et al., 2012)and 70,211 liver DHS used in this study. **B.** Distribution of STAT5 ChIP-seq signal intensity for the sets of dynamic and static DHS. For each DHS that contains a STAT5 binding site (shown in A), the corresponding STAT5 normalized ChIP-seq read count was obtained for the STAT5 male-high samples and STAT5 female-high liver samples (Zhang et al., 2012). Shown are the distributions of STAT5 ChIP-seq signal intensities for the indicated sets of DHS and using average read counts from STAT5 male-high and female-high ChIP-seq samples. **C.** Normalized DNase-I cut site aggregate plots for male livers at STAT5-high, STAT5-low, and female liver in the set of male-biased, female-biased, and sex-independent DHS separated by dynamic and static subsets.



Figure 3.4. STAT5 binding is associated with chromatin opening

Fig. 3.5 DHS responsive to hyox and their enrichment for class I and II sexbiased genes.

A. The number of DHS that open or close in response to hypox or a single injection of GH given to hypox male mice and euthanized 30, 90, or 240 min later. diffReps analysis of DNase-seq data for each of the indicated comparisons (DNase-seq treated/control samples) identify responsive ΔDHS; the number of diffReps regions and those that overlap the set of 70,211 DHS are shown in Fig. S3.4. A full listing of these datasets is provided in Supplementary Table_S4. **B.** Enrichment of hypox-responsive DHS as compared to hypox-unresponsive (static) DHS for mapping to four different classes of sex-specific genes. Gene targets of each DHS were as described in methods. Class I and II sex-specific genes were identified from RNA-seq gene expression data collected from intact and hypox male and female liver samples. The table at the bottom shows the total number and percentage of sex-specific genes in each of the four indicated classes of class I and II sex-specific genes that respond to hypox, as indicated. A more detailed analysis is presented in Fig. S3.3.

Figure 3.5. DHS responsive to hypophysectomy and their enrichment for class I and II sex biased genes



B The enrichment of hypox responsive DHS for mapping to sexbiased genes

	dDH3 (Male Hypox/	_ciose Male control)	dDH 3 (Female Hypox/	iciose Female control)	dDHS (Male Hypox)	_open Male control)	dDHs (Female Hypox/	s_open Female control)	
Male_class_I	5.7	1.3E-69	3.9	1.69E-21	0.3	9.55E-04	0.3	6.03E-08	
Female_class_I	2.9	4.1E-11	7.8	6.64E-46	0.9	8.83E-01	0.2	3.38E-06	
Male_class_II	1.3	2.2E-01	0.7	4.35E-01	1.7	5.62E-02	1.9	4.83E-04	
Female_class_II	1.3	1.7E-01	1.3	1.45E-01	3.6	1.80E-15	0.6	2.03E-02	
	# genes down in Male Hypox	percentage	# genes down in Female Hypox	percentage	# genes up in Male Hypox	percentage	# genes up in Female Hypox	percentage	Total number of genes in set
Male_class_I	202	100.0%	89	44.1%	0	0.0%	3	1.5%	202
Female_class_I	35	26.1%	134	100.0%	9	6.7%	0	0.0%	134
Male_class_II	0	0.0%	0	0.0%	40	44.4%	90	100.0%	90
Eamsla class	0	0.0%	0	0.09/	179	100.09/	70	20.29/	179

Fig. 3.6 Overlap between ΔDHS responsive to a single injection of GH in male hypox liver and DHS responsive to endogenous plasma GH pulse.

A. Shown are the number of STAT5-high and STAT5-low diffReps regions and the numbers of overlapping DHS from the full set of 70,211 liver DHS. **B and C.** Bar graphs show the numbers of STAT5-high vs. STAT5-low opening DHS (B) and closing DHS (C) that respond to hypox in male or female mouse liver, and that respond to GH pulse replacement in hypox male liver at the three time points indicated. Data are expressed as a percentage of the full sets of 2,373 DHS that open and the 98 DHS that close in response to an endogenous GH pulse. Percentages do not add up to 100% because a subset of the 2,373 DHS and a subset of the 98 DHS were not being identified as DHS in the hypox liver DNase-seq datasets.

Figure 3.6. Overlap between ΔDHS responsive to natural GH pulses and ΔDHS responsive to a single injection of GH in male hypox samples

A DHS Comparison	Number of diff DHS (ΔDH	Reps differential IS) regions	Number of DHS that overlap the set of 70,211 DHS			
(DNase-seq in male liver)	DHS that open	DHS that close	DHS that open	DHS that close	Static DHS (sDHS)	
STAT5-high/ STAT5-low	2,831	123	2,373	98	45,754	
В	DHS that o (2,373 site	pen s)	dDHS_op dDHS_clo static	ben ose		
Percentage Composition	1,487 60 - 50 - 40 - 20 - 10 - 0 - 3 20 - 10 1 21 21 21 21 21 21 21 21 21 21 21 21 21	1,475 1,340 421 262 2	13 547 392 4 4	43		
	MHx/ FH Male Fen Control Cor	Hx/ 30 min 9 nale MHx ntrol MHx	90 min 240 m x+GH/MHx	in		
Percentage Composition	100- 90- 80- 75- 70- 60- 50- 40- 30-	lose (98 sites)	40 26	51		

0

90 min

MHx+GH/MHx

240 min

30 min

6 6

FHx/

Female

Control

3

MHx/

Male

Control

Fig. 3.7 DNase cut site aggregate plots for the GH time course DHS data.

A. Normalized DNase-I cut site aggregate plots for hypox male mouse liver (MHx), hypox males treated with GH (MHx+GH) after 30, 90, and 240 min, intact females, and female hypox (FHx) across the sets of STAT5-high vs. STAT5-low diffReps-discovered DHS that open (*left*) or close (*right*) in response to endogenous STAT5 pulses in male liver, as shown in Fig. 3.2C. **B.** Normalized DNase-I cut site aggregate plots for the set of DNase-seq samples shown in A and in the set of male-biased DHS (*left*) and sex-independent DHS (*right*), separated by those with (*top*) or without (*bottom*) STAT5 bound at the DHS regions, as in Fig. 3.2D.



Figure 3.7. DNase cut site aggregate plots for the GH time course DHS data

A and B. Shown are normalized DNase-I cut site aggregate plots for hypox male (MHx), hypox male treated with GH (MHx+GH) after 30, 90, and 240 min, intact female and female hypox (FHx) across various sets of static and dynamic male-biased DHS, static female-biased, and static sex-independent DHS, as shown in Fig. 3.3A (A) and in Fig. 3.4C (B).





Fig. 3.9 Chromatin state analysis of static and dynamic male-biased DHS.

A. Distribution of static and dynamic DHS as weak enhancers, enhancers, insulators and promoter DHS, based on (Matthews and Waxman, 2018). B. Chromatin state distributions in male mouse liver of dynamic and static male-biased DHS, based on the 14 chromatin state model developed from the combination of six active and repressive chromatin marks and DHS, which segment the mouse genome into inactive, bivalent, enhancer-like, promoter-like, or transcribed-like states (Sugathan and Waxman, 2013). Chromatin state data are shown in Supplementary Table_S5. C. Emission probabilities for the six histone marks and DHS for each of the 14 chromatin states, from (Sugathan and Waxman, 2013), and the grouping of the 14 states into the indicated superstates. **D-G.** Shown are significant enrichments, and significant depletions, of the indicated sets of genomic regions for each of four sets of DHS: (1) dynamic male-biased DHS (834 sites), (2) static male-biased DHS (1,895 sites), (3) dynamic sex-independent DHS (1,532 sites), and (4) static female-biased DHS (1,359 sites). These enrichments and depletions were determined for overlap with sex-biased transcription factor binding sites (D), sexbiased chromatin marks (E), chromatin states in male liver (F), and chromatin states in female liver (G). The set of 64,584 static sex-independent DHS was used as the background for these enrichment calculations (see Methods). Data graphed are enrichment scores (and depletion scores), indicated by bar height, and their Fisher Exact test significance values (log p-values, indicated by bar color) for all values that are significant at p<E-03. Values that did not meet this significance threshold are graphed as ES = 1 (horizontal dashed green line), which indicates the absence of enrichment.

Enrichment scores < 1 at p<E-03 indicate significant depletion. All significant depletions with ES values <0.1 (i.e., >10-fold depletion) are graphed at ES = -10. Thus, all bars shown, except those at ES = 1.0, indicate significant enrichment or depletion. Enrichment calculations are presented as sets of four bars corresponding to the order of DHS listed above. Full details on the numbers of sites, the source publications used to identify these genomic regions, and corresponding BED files are shown in Table S7.





С **ChromHMM Emission Probabilities** В Composition of male biased DHS with 1. Inactive respect to chromatin states in male liver 2. Inactive 3. Inactive 100 -4. Low signal 90 male_bias_dynamic 5. Enhancer 6. Enhancer % of DHS in state 30 male_bias_static 7. Promoter 8. Promoter 10 -9. Enhancer 30 -10. Enhancer 11. Enhancer 50 -Bivalent 10 -13. Transcribed 14. Transcribed 30 me3 me3 me3 ຮ e] ne 20 -10 -0 19 0-06 1 15 Chromatin super states E4 E5 E6 E7 E8 E9 E10 E11 E12 E13 E14 É1 E2 É3 4 No marks Chromatin state 3 2 1 Inactive **Bivalent** 5 6 9 10 11 Enhancer-like

> 8 7 Promoter-like 13 14 Transcribed-like





Chromatin marks in male liver

Chromatin marks in female liver

Figure 3.9FG. Enrichments of male biased, sex-independent, and female biased DHS



Fig. 3.10 Venn diagram and flow chart of the DHS that open due to an endogenous GH pulse (2,373 sites)

A. Four-oval Venn diagram showing the number of STAT5-high Δ DHS that were closed by hypophysectomy and/or opened by a single exogenous GH pulse. The set of STAT5high Δ DHS (n = 2,373 sites) (Fig. 3.6A) was analyzed to determine their overlap with the set of DHS that close due to hypox (n = 1,487 sites) or DHS that open due a single exogenous GH pulse in hypophysectomized male liver after 30 (n = 1, 475 sites), 90 (n = 1,393 sites), or 240 (n = 547 sites) minutes (Fig. 3.6 BC).

B. Flowchart of STAT5-high Δ DHS (n = 2,373 sites) identifying hypox and exogenous GH pulse responsive DHS in male liver. Subsections of the 4-oval Venn diagram (shown in A) are used to illustrate the DHS subsets defined by the flowchart (7 subsets). Shown in grey are the set of 399 DHS that lacked chromatin closing in response to hypox and lacked chromatin opening in any of the GH time-course data sets.

Figure 3.10. Venn diagram and flow chart of the DHS that open due to an endogenous GH pulse (2,373 sites)



Note: there were 399 DHS that lacked chromatin closing in response to hypox (salmon oval) and lacked chromatin opening in any of the GH time-course data sets (green and blue ovals). The sum of all DHS counts equals 2,373 sites (shown above).



Fig. 3.11 Boxplot analysis and DNase-I cut site aggregate plots of the 7 DHS subsets

A. Relative magnitude of chromatin opening between STAT5 high and STAT5 low male livers for the 7 DHS subsets. The x-axis shows the DHS subsets numbered 1-7 with the corresponding number of DHS sites (defined in Fig. 3.10B). RiPPM normalized DNaseseq read counts were obtained from the STAT5-high and STAT5-low male liver samples for the set of 70,211 DHS used in this study. The relative magnitude of chromatin opening (i.e. fold-change) was calculated from the ratio of STAT5-high/STAT5-low RiPPM normalized DNase-seq read counts in the DHS region. Shown are the distributions of fold-change values for DHS subsets numbered 1-7 defined from the flowchart in Fig. 3.10B. A Wilcoxon rank-sum test with Benjamini–Hochberg p-value adjustment was used to determine significant differences between distributions of foldchange values (*P < 0.05, **P < 1e-03, ***P < 1e-10).

B. Normalized DNase-I cut site aggregate plots for male livers at STAT5-high, STAT5-low, and female liver in the 7 DHS subsets shown in A.




Fig. S3.1 STAT5 ChIP-seq signal intensity distributions for the indicated sets of dynamic and static DHS.

Also see Fig. 3.4B. Data are shown for DHS that have STAT5 bound ("any STAT5 TFBS"), and for separated subsets of DHS that have either male-enriched, femaleenriched, or male-female common STAT5 binding sites, as determined in (Zhang et al., 2012). As described in Fig. 3.4B, the ChIP-seq signal intensities for the STAT5 malehigh samples were used for the male-biased DHS, and STAT5 female-high samples were used for the female-biased DHS. An average of the male-high and female-high read counts was used for sex-independent DHS.

Fig. S3.1 These sites are the subset of DHS that have STAT5 bound



Fig. S3.2 Discovered motifs from DHS sequences.

De-novo motif analysis was carried out on the sets of dynamic and static DHS sequences using MEME-ChIP (Ma et al., 2014). Listed are the de-novo discovered motifs from dynamic male-biased DHS (834 sites) (**A**), static male-biased DHS (1,895 sites) (**B**), static female-biased DHS (1,359) (**C**), dynamic sex-independent DHS (1,532 sites) (**D**), and the top 10 significant motifs from static sex-independent DHS (64,584 sites) (**E**). No motifs were discovered from the very small number of dynamic female-biased DHS (7 sites). Static and dynamic DHS were identified as shown in Fig. 3.2E. The presence of the STAT5B motif is a distinguishing characteristic that separates the dynamic from static DHS, which supports our hypothesis that direct binding of STAT5 plays a key role in chromatin opening in response to GH pulses at these dynamic DHS.

Fig. S3.2A Summary of de-novo motif analysis

Male biased dynamic DHS (834 sites)				
Motif Logo	E-value for motif discovery (MEME/DREME)	Closest matching motif (Tomtom p-value)		
^₄] _∞ ۅ ۅ _≈ <u></u> ۅ ۅ , , , , , , , , ,	2.1e-30 (MEME)	motif72 (IRF2) (5.44e-04) motif45 (FOXI1) (2.00e-03) motif78 (FOXD1) (2.08e-03)		
ĨŢŢŢĊŢ <u>ŗ</u> ĢĠĂ	4.1e-23 (MEME)	motif29 (STAT5B) (2.84e-06)		
TTCTGAGA	3.7e-12 (MEME)	motif29 (STAT5B) (6.63e-03)		
ġ<u>ġ</u>ġ<u>ġ</u><u>ġ</u>ġġġġġġġġġġġġġġġġġġġġġġġġġġġ	4.3e-12 (MEME)	motif15 (Pax4) (1.05e-05) motif39 (MZF1_5-13) (5.14e-04) motif47 (SP1:SP3) (6.03e-04)		
<u><u><u></u>^t c_A_SGcCA_GcC_†GG_st_Ces</u></u>	2.2e-12 (MEME)	motif48 (SF1) (1.39e-03) motif68 (Hand1-Tcfe2) (4.38e-03)		
	1.2e-15 (DREME)	motif29 (STAT5B) (1.20e-04) motif80 (c-Ets-2) (7.50e-03) motif79 (ELF5) (9.13e-03)		

Fig. S3.2B Summary of de-novo motif analysis

Male biased static DHS (1,895 sites)				
Motif Logo	E-value for motif discovery (MEME/DREME)	Closest matching motif (Tomtom p-value)		
^ŗ <mark>"ĢĢĢĢĢ</mark> ĘĄ <mark>Ģ</mark> ŗĢ <mark>Ģ</mark> ŗĢĢ	6.3e-21 (MEME)	motif92 (SP1) (1.28e-05) motif39 (MZF1_5-13) (1.72e-04) motif47 (SP1:SP3) (8.89e-04)		
[#] <mark>I_₹_₹ŢĢ</mark> Ţ _₹ Ţ <mark>Ŗ</mark> <mark>ĢļĢ<mark>ļ</mark>Ģ<mark>ļĢŢ</mark>Ģ<mark>ĢŢ</mark>Ŗ</mark>	7.2e-13 (MEME)	No motif match reported		
∄ <mark>」_╤ᢗᢩͳ┍Ţ<mark>∁</mark>_ÇҀ<mark>Ҭ</mark>╤</mark>	1.4e-12 (MEME)	motif2 (NR1H2-RXRA) (1.26e-06) motif71 (NR2F1) (3.25e-05) motif48 (SF1) (1.08e-03)		
	2.7e-11 (MEME)	motif81 (MEIS1B:HOXA9) (8.98e-03)		
ATIGAT I	9.4e-28 (DREME)	motif19 (PBX1) (7.33e-04) motif24 (Clox) (1.04e-03) motif31 (CDP) (3.37e-03)		
E CAAAG ŢŢ	3.9e-17 (DREME)	motif71 (NR2F1) (7.43e-06) motif2 (NR1H2-RXRA) (5.52e-05) motif48 (SF1) (8.75e-03)		
	4.0e-12 (DREME)	motif45 (FOXI1) (9.20e-06) motif78 (FOXD1) (3.28e-03) motif97 (TEF-1) (5.15e-03)		

200

Fig. S3.2C Summary of de-novo motif analysis

Female biased static DHS (1,359 sites)				
Motif Logo E-value for motif discovery (MEME/DREME) (Tomtom p-value)				
	4.3e-17 (MEME)	motif39 (MZF1_5-13) (1.28e-03) motif15 (Pax4) (2.79e-03) motif92 (SP1) (5.47e-03)		
	2.6e-10 (MEME)	No motif match reported		
	2.8e-013 (MEME)	motif13 (TP53) (7.26e-03) motif48 (SF1) (7.26e-03) motif51 (NRSF) (7.27e-03)		
	3.8e-017 (DREME)	No motif match reported		

Fig. S3.2D Summary of de-novo motif analysis

Sex independent dynamic DHS (1,532 sites)				
Motif Logo	E-value for motif discovery (MEME/DREME)	Closest matching motif (Tomtom p-value)		
	2.4e-32 (MEME)	motif29 (STAT5B) (2.76e-09) motif38 (DEAF1) (3.69e-03)		
	3.5e-10 (MEME)	No motif match reported		
₄ ŢŢŢŢŢŢŢŢŢŢŢŢŢŢŢŢŢŢŢŢŢŢ	7.8e-10 (MEME)	motif45 (FOXI1) (8.73e-04) motif78 (FOXD1) (4.69e-03) motif88 (Elf-1) (6.39e-03)		
	1.3e-12 (MEME)	motif2 (NR1H2-RXRA) (2.01e-06) motif71 (NR2F1) (3.10e-06)		
	1.1e-27 (DREME)	motif29 (STAT5B) (1.52e-04)		
	1.4e-14 (DREME)	motif45 (FOXI1) (3.34e-04) motif78 (FOXD1) (7.76e-03)		

Fig. S3.2E Summary of de-novo motif analysis

Sex independent static DHS (64,584 sites)				
Motif Logo	E-value for motif discovery (MEME/DREME)	Closest matching motif (Tomtom p-value)		
· AcaCacacACACACAcacaCaCACACACAc	1.8e-35 (MEME)	No motif match reported		
	1.4e-12 (MEME)	motif78 (FOXD1) (4.22e-05) motif45 (FOXI1) (2.20e-03) motif72 (IRF2) (2.81e-03)		
	3.3e-458 (DREME)	motif88 (Elf-1) (3.97e-05) motif79 (ELF5) (5.49e-05) motif80 (c-Ets-2) (3.61e-04)		
	4.7e-286 (DREME)	motif78 (FOXD1) (3.27e-05) motif45 (FOXI1) (3.72e-04) motif76 (Lhx3) (7.75e-03)		
	1.3e-281 (DREME)	motif48 (SF1) (2.15e-05) motif2 (NR1H2-RXRA) (2.90e-04) motif18 (RORA_1) (3.10e-04)		
	2.5e-212 (DREME)	motif92 (SP1) (3.95e-06) motif39 (MZF1_5-13) (3.18e-03) motif87 (Egr) (6.29e-03)		
	4.2e-172 (DREME)	motif62 (TLX1-NFIC) (1.24e-03) motif68 (Hand1-Tcfe2) (4.94e-03)		
	1.0e-142 (DREME)	motif72 (IRF2) (3.46e-03) motif45 (FOXI1) (4.29e-03) motif78 (FOXD1) (7.79e-03)		
	7.1e-108 (DREME)	motif17 (Myc) (1.09e-07) motif6 (MYC-MAX) (1.34e-07) motif42 (Myf) (2.73e-04)		
	2.7e-104 (DREME)	motif34 (LUN-1) (6.09e-04) motif29 (STAT5B) (3.564-03)		

Fig. S3.3 Summary of sex-specific genes used in this study.

Table 1: Classification of pituitary hormone-dependent sex-biased RefSeq and lncRNA genes. Sex-specific genes were identified from RNA-seq data from intact male and female mouse liver based on a gene list comprised of 24,197 RefSeq and 3,152 multi-exonic lncRNA genes. First, sex-specific genes were identified with |fold-change| > 1.5, adjusted p-value < 0.05 (for RefSeq genes), and |fold-change| > 2.0, adjusted p-value < 0.05 (for lncRNA genes), with FPKM > 0.25 for the sex with greater signal intensity for both RefSeq and lncRNA datasets (see Methods). Sex-specific genes that were responsive to hypox (|fold-change| > 2 and an adjusted p-value < 0.05) were further classified into class I, II and corresponding subclasses (IA, IB, IC, IIA, and IIB) based on RNA-seq data from intact and hypox male and female liver samples. Shown are the full list of RefSeq and lncRNA genes (Table 1), the subset that are RefSeq genes (Table 2) and the subset that are lncRNA genes (Table 3).

Table 1. Classification of Pituitary Hormone-Dependent Sex-Biased (RefSeq and IncRNA) Genes					
Class	Number of Genes (%)	Response to MHx	Response to FHx	Gene Subclass	Number of Genes
Male class I	202 (69)	Down	-	IA	110
		Down	Down	IB	89
		Down	Up	IC	3
Male class II	90 (31)	-	Up	IIA	50
		Up	Up	IIB	40
Female class I	134 (43)	-	Down	IA	90
		Down	Down	IB	35
		Up	Down	IC	9
Female class II	178 (57)	Up	-	IIA	108
		Up	Up	IIB	70

Fig. S3.3. Summary of sex specific genes used in this study

Table 2. The subset of genes shown in Table 1 that are RefSeq genes					
Class	Number of Genes (%)	Response to MHx	Response to FHx	Gene Subclass	Number of Genes
Male class I	163 (71)	Down	-	IA	88
		Down	Down	IB	72
		Down	Up	IC	3
Male class II	68 (29)	-	Up	IIA	38
		Up	Up	IIB	30
Female class I	101 (40)	-	Down	IA	67
		Down	Down	IB	25
		Up	Down	IC	9
Female class II	149 (60)	Up	-	IIA	96
		Up	Up	IIB	53

Table 3. The subset of genes shown in Table 1 that are IncRNA genes					
Class	Number of Genes (%)	Response to MHx	Response to FHx	Gene Subclass	Number of Genes
Male class I	39 (64)	Down	-	IA	22
		Down	Down	IB	17
		Down	Up	IC	0
Male class II	22 (36)	-	Up	IIA	12
		Up	Up	IIB	10
Female class I	33 (53)	-	Down	IA	23
		Down	Down	IB	10
		Up	Down	IC	0
Female class II	29 (47)	Up	-	IIA	12
		Up	Up	IIB	17

Fig. S3.4 Summary of DHS counts for the GH time course study.

Shown are the number of Δ DHS identified from the GH time course study, including the diffReps-identified differential DHS sites and the subset of those that overlap the set of 70,211 liver DHS used in this study. The latter set of values is graphed in Fig. 3.5A. Many more sites close than open following hypox in male liver. Further, the numbers of DHS that open following hypox decreases from 30, 90, 240 min after a single GH injection.

Fig. S3.4. Summary of DHS counts for the GH time course study

Male hypophysectomy = MHx Female hypophysectomy = FHx

DHS	Number of diffReps differential DHS		Number of DHS in the set of 70,211 liver DHS that overlap diffReps differential DHS		
Comparison	sites that open	sites that close	DHS that open	DHS that close	Static DHS
MHx/ Male Control	2,142	4,856	1,709	4,147	35,562
FHx/ Female Control	3,444	2,753	2,624	2,477	30,394
MHx +GH 30 min/ MHx	4,243	1,113	3,589	736	23,614
MHx +GH 90 min/ MHx	3,829	1,646	3,241	1,214	23,570
MHx +GH 240 min/ MHx	2,040	777	1,825	500	26,118

APPENDIX

Appendix 1 – List of supplemental tables related to Chapter 2.

Supplemental files are available upon request by email to <u>djw@bu.edu</u>

Table S1A – Histone marks and transcription factor ChIP-seq analysis - Shown are the total and mapped read counts for the ChIP-seq samples prepared and analyzed in this study, the number of ChIP-seq peaks discovered in each sample either by MACS2 (for activating histone marks) or SICER (for repressive marks), and the fraction of sequence reads found in the sample's respective peak list. The ratio genome covered (column K) was calculated for regions discovered by SICER.

Table S1B – CAR, RXRa, CEBPa, and CEBPb binding sites at DHS. Shown are the number of transcription factor binding sites discovered genome-wide for the indicated TCPOBOP exposure (column A) and transcription factor (column B). MACS2 was used to generate a peak union list of transcription factor binding sites by merging peak calls across biological replicates. Total number of MACS2 binding sites are shown (column C). ChIP-seq datasets for RXRA, CEBPA, and CEBPB were analyzed with diffReps to identify genomic regions with ChIP-seq signals that differed significantly between vehicle-treated and TCPOBOP-treated ChIP samples for each TCPOBOP exposure (diffReps regions with a |fold-change| > 2 in diffReps-normalized read counts and FDR < 0.05 (Benjamini-Hochberg adjusted p-value)). Numbers of TCPOBOP-induced (column E) and TCPOBOP-repressed sites (column F) are shown. CAR ChIP-seq data (Tian et al, 2018) was reanalyzed to obtain a total of 14,631 CAR binding sites which either

overlapped any one of the 60,739 DHS or were discovered from the peak union of all replicate sample MACS2 peak calls. CAR ChIP-exo data (Niu et al, 2018) was converted from the mm10 to mm9 mouse genome assembly for analysis in this study. The total number of binding sites, number and percentage of them at any one of the 60,739 DHS are shown.

Table S2 Histone Mark Summary – Sets of ChIP-seq peaks discovered for activating (H3K27ac, H3K4me1, H3K4me3) and repressive (H3K27me3) chromatin mark datasets at 3 h and 27 h TCPOBOP exposure time points in male liver. Each sheet contains the peak union list (see Methods) generated from the respective data set and indicates a single overlapping DHS for each peak union site.

Table S2A - H3K27ac sites overlap with DHS. Merged list of 68,608 H3K27ac regions was generated based on the MACS2 peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M3hr H3K27ac TCPOBOP-treated combined sample (Treatment.cnt, column K) and the vehicle-treated combined sample (Control.cnt, column L) for the chromatin mark peaks defined in columns A-C.

Table S2B – **H3K4me1 sites overlap with DHS**. Merged list of 124,822 H3K4me1 regions was generated based on the MACS2 peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M3hr H3K4me1 TCPOBOP-treated combined sample (Treatment.cnt, column K) and the vehicle-treated combined sample (Control.cnt, column L) for the chromatin mark peaks defined in columns A-C.

Table S2C – H3K4me3 sites overlap with DHS. Merged list of 41,638 H3K4me3 regions was generated based on the MACS2 peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M3hr H3K4me3 TCPOBOP-treated combined sample (Treatment.cnt, column K) and the vehicle-treated combined sample (Control.cnt, column L) for the chromatin mark peaks defined in columns A-C.

Table S2D - H3K27ac sites overlap with DHS. Merged list of 59,595 H3K27ac regions was generated based on the MACS2 peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M27hr H3K27ac TCPOBOP-treated combined sample (Treatment.cnt, column K) and the vehicle-treated combined sample (Control.cnt, column L) for the chromatin mark peaks defined in columns A-C.

Table S2E – H3K4me1 sites overlap with DHS. Merged list of 108,263 H3K4me1 regions was generated based on the MACS2 peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M27hr H3K4me1 TCPOBOP-treated combined sample (Treatment.cnt, column K) and the vehicle-treated combined sample (Control.cnt, column L) for the chromatin mark peaks defined in columns A-C.

Table S2F – H3K4me3 sites overlap with DHS. Merged list of 23,542 H3K4me3 regions was generated based on the MACS2 peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M27hr H3K4me1 TCPOBOP-treated combined sample (Treatment.cnt, column K) and the vehicle-treated combined sample (Control.cnt, column L) for the chromatin mark peaks defined in columns A-C.

Table S2G – H3K27me3 sites overlap with DHS. Merged list of 19730 H3K27me3 regions was generated based on the SICER peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M3hr H3K27me3_Longer TCPOBOP-treated combined sample (Treatment.cnt, column K) and the vehicle-treated combined sample (Control.cnt, column L) for the chromatin mark peaks defined in columns A-C. **Table S2H - H3K27me3 sites overlap with DHS**. Merged list of 16,359 H3K27me3 regions was generated based on the SICER peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M3hr H3K27me3_Long TCPOBOP-treated combined sample (Treatment.cnt, column K) and the vehicle-treated combined sample (Control.cnt, column L) for the chromatin mark peaks defined in columns A-C.

Table S2I – H3K27me3 sites overlap with DHS. Merged list of 21,467 H3K27me3 regions was generated based on the SICER peaks called in each ChIP-seq replicate sample (see Methods) and assigned a unique name (Chromatin Mark Peak Number, column E). The genomic position of chromatin mark peak is listed in columns A-C, its response to TCPOBOP treatment (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin mark peak. RiPPM normalized read counts from the M27hr H3K27me3 vehicle-treated combined sample (Control.cnt, column L) and the TCPOBOP-treated combined sample (Treatment.cnt, column K) for the chromatin mark peaks defined in columns A-C.

Table S3 – Chromatin mark read counts in the set of \triangle DHS and static DHS.

Columns A-D, List of 60,739 liver DHS regions, identified by DNase-seq, as reported in (Lodato, Rampersaud, and Waxman 2018) for male and female mouse liver after 3 h or 27 h TCPOBOP exposure, or vehicle controls. Merged list of 60,739 DHS regions was generated based on the MACS2 peaks called in each DNase-seq replicate sample and assigned a unique name (DHS Peak Number, column D). Genomic position of DHS is listed in columns A-C. Shown beginning in column E are the RiPPM-normalized read counts in each DHS region, for the TCPOBOP-treated combined sample (Treatment.cnt) and the corresponding vehicle-treated combined sample (Control.cnt) for each chromatin mark dataset: M3hr_H3K27ac (columns E-F), M3hr_H3K4me1 (columns G-H), M3hr_H3K4me3 (columns I-J), M27hr_H3K27ac (columns K-L), M27hr_H3K4me1 (columns M-N), M27hr_H3K4me3 (columns O-P), M3hr_H3K27me3_Longer (columns Q-R), M3hr_H3K27me3_Long (columns S-T), and M27hr_H3K27me3 (columns U-V). Read counts are shown for all 60,739 liver DHS regions, including those that do not overlap a ChIP-seq histone mark peak (as shown in Table S2).

Table S4 - Chromatin marks at △DHS and static DHS annotated as promoter or enhancer regions

Columns A-I, Set of 60,739 liver DHS regions and their responses to 3 h or 27 h TCPOBOP exposure in male and female mouse liver as determined by DNase-seq, as reported in (Lodato, Rampersaud, and Waxman 2018). Merged list of 60,739 DHS regions was generated based on the MACS2 peaks called in each DNase-seq replicate sample and assigned a unique name (DHS Peak Number, column D). Genomic position of DHS is listed in columns A-C and the mouse liver TAD where the DHS is found is listed in column E.

Columns F-I: DHS response for each DHS in each TCPOBOP-exposed group, where "standard_dDHS_open" and "robust_dDHS_open" indicate DHS opening, and "standard_dDHS_close" and "robust_dDHS_close" indicate DHS closing. "Static" indicates no significant change in DNase hypersensitivity for the indicated TCPOBOP exposure. The DHS response was summarized for male liver (column J) in the following way: dDHS_that_opens, if chromatin opening was detected at either or both time points (provided no conflicts), dDHS_that_closes, if chromatin closing was detected at either or both time points (provided no conflicts), and sDHS, if the DHS did not show a significant change in its DNase-seq intensity at both time points. Sites labeled as "conflict A" refer to DHS that open/close at two different time points in male liver. Overall, 2,029 DHS open, 630 DHS close, and 58,078 DHS are static, i.e., unchanged in their accessibility in male liver at either time point. When responses in feamle liver are also taken into account, the set of static DHS was reduced to 55,866; that set of DHS was used as background for enrichment calculations (see text). The nearest gene within the same TAD to each DHS is listed in column N, gene TSS coordinates in columns K-M, TAD assignment in column O, and the linear genomic distance is listed in column P.

Columns Q-S: summary results from RNA-seq, as reported in Lodato et al (Tox Sci, 2017) in male and female mouse liver after 3 h and 27 h TCPOBOP exposure. RNA-seq results include FPKM (fragments per kilobase per million reads) in vehicle-treated and TCPOBOP-treated mouse liver, ratio, and adjusted p-value (FDR). Corresponding TFS (total flag sum) number for the gene was generated as follows: 3 h male, 1.1= up regulated, 1.2=down regulated; 27 h male, 2.01=up regulated, 2.02=down regulated; 3 h female, 4.001=up regulated, 4.002=down regulated; 27 h female, 8.0001=up regulated, 8.0002 = down regulated. Vh, vehicle; TC, TCPOBOP. The sum of these TFS values is calculated in column Q and the pattern represented by the TFS summarizes the significance and directionality of the response to all four TCPOBOP exposures. Column R summarize the target gene response in the following way: induced_gene, if gene induction was detected at one or both time points, repressed_gene, if gene repression was detected at one or both time points, static_gene, if the gene response was static at both time points in male liver.

Characterization of DHS based on time point specific responses to TCPOBOP exposure are summarized for male liver (column T) and female liver (column U), as indicated at the top. Characterization of DHS based on sex specificity of response to TCPOBOP is summarized for the 3 h time point (column V) and for the 27 h time point (column W). The chromatin mark response to TCPOBOP is shown for for three activating marks (H3K27ac, H3K4me1, and H3K4me3) at 3 h (columns X-Z) and at 27 h (columns AA-AC) exposure time points. Induced_mark and repressed_mark sites are genomic regions that showed significantly differential ChIP-seq signal between TCPOBOP-exposed and vehicle treated control samples using diffReps (Shen et al. 2013), with significance based on |fold-change| > 2 and FDR < 0.05 (Benjamini-Hochberg adjusted p-value). Static_mark sites did not show a significant change in diffReps normalized signal intensity values, and "none" indicates lack of overlap between the DHS and chromatin mark data set. Activating chromatin mark response based on 3 h and 27 h TCPOBOP time points for H3K27ac (column AD), H3K4me1 (column AE), and H3K4me3 (column AF). Briefly, the DHS was labeled "Induced" if the associated chromatin mark signal intensity increased at one or both TCPOBOP time points, "Repressed", if the associated chromatin mark signal intensity decreased at one or both TCPOBOP time point, "if no significant change was detected at either time point, or "absent", if there was no overlapping or nearby (within 400 bp) MACS2 chromatin mark peak at either time point. Column AG provides an overall summary of the chromatin mark response for all three activating marks.

DHS were designated as "promoter" DHS if they had an associated H3K4me3 peak and the DHS region was < 3 kb from the nearest RefSeq or multi-exonic lncRNA gene transcription start site in the same TAD. All other DHS are designated as "non_promoter" DHS (column AH). DHS were further classified into promoter and enhancer subclasses, which are either induced, repressed, static, or poised (see Methods) (column AI). The chromatin mark response (induced_mark, repressed_mark, static_mark) to TCPOBOP for a repressive chromatin mark (H3K27me3) at 3 h (columns AJ-AK) and at 27 h (column AL) exposure time points.

Columns AM and AN show the overlap with 52,436 CTCF peaks defined in B Matthews and DJ Waxman, eLife 2018 (Suppl file 1_TableS1C_CTCFPeakList_with Rad21 data.xlsx). Overlap is indicated by the type of CTCF peak and lack of overlap is indicated with a 0. Column AN extends the DHS genomic region considered for overlap by 400 bp in each direction.

Table S5A - ChIP-seq peaks for CAR, RXR and CEBP at liver DHS. Columns A-H, see legend for Table S4 for detailed column descriptions. Column I-K: analysis of CAR ChIP-seq data from Tian et al (2018). diffReps was used to identify sites of differential CAR ChIP-seq signal intensity between control and TCPOBOP-treated livers. DHS that contained a CAR binding site (TFBS, transcription factor binding site) were labeled induced_TFBS, repressed_TFBS, or static_TFBS, as determined by diffReps, and DHS that lack CAR binding were labeled "none". For those CAR binding sites at a DHS, we counted the number of nuclear receptor half-site motifs (column J) and the number of DR4-like motifs (column K) in the CAR binding site sequence. Columns L-Q: Subset of filtered list of 2,525 CAR binding sites determined by Tian et al (2018), of which 2,290 overlap a total of 2,287 liver DHS. Also shown is whether the DHS overlapped a binding site for RXR (column Q). Columns R-U: analysis of CAR ChIP-exo data from Niu et al

(2018). The UCSC Genome Browser Utility "liftover" command was used to convert BED from mouse genome assembly mm10 to genome assembly mm9. A CAR non-zero MACS14 score is shown for those DHS that overlap a CAR binding site. CAR binding data is shown for transgenic mice with humanized CAR (hCAR) treated with CITCO or phenobarbital (PB), and for transgenic mice with mouse CAR (mCAR) treated with phenobarbital or TCPOBOP. Column V: Marked are those DHS that were assigned to one of the following four sets: Opening DHS and Static DHS (sDHS; static with respect to TCPOBOP-induced chromatin opening) that overlap with robust CAR binding sites, or that are stringent CAR-free. Robust indicates CAR binding based on the data of Tian et al (2018) (i.e., CAR binding site in column I or in column L), and under all four conditions shown in columns R-U. Stringent CAR-free indicates no CAR binding under any of the above conditions (columns I, L, R-U). Columns W-AB: Transcription factor binding site (ChIP-seq) response to 3 h or 27 h TCPOBOP exposure in male liver for RXRa, CEBPa, and CEBPb. The transcription factor binding response to TCPOBOP was determined from diffReps analysis, which identified genomic regions with significantly differential ChIP-seq signals between vehicle-treated and TCPOBOP-treated samples with a |foldchange > 2 and FDR < 0.05 (Benjamini-Hochberg adjusted p-value) (see Methods). Columns AC and AD show chromatin mark responses and chromatin state designations for each DHS, taken from Table S4, columns AG and AI, respectively.

 Table S5B - CAR binding sites that overlap a DHS region, based on our analysis of

 Tian et al (2018) data. Raw sequencing data was obtained and analyzed from a recently

published mouse liver CAR ChIP-seq study (Tian et al., 2018; Sequence Read Archive accession SRP125957), with two low quality samples excluded from our analysis, as described in Methods. A merged list of 14,631 CAR binding sites was generated based on the MACS2 peaks called in each ChIP-seq replicate sample and sites identified from diffReps analysis (see Methods) and assigned a unique name (CAR Binding Site Peak Number, column E). The genomic position of the CAR binding site peak is listed in columns A-C, its response to TCPOBOP treatment (column D; TFBS, transcription factor binding site), and its overlapping DHS region (if there is one) (column F-I). For those CAR binding sites that do not overlap a DHS, values of (., -1, -1, .) are entered for the DHS coordinates (Chr, Start, Stop) respectively (column F-I). The number of base pairs overlap between the CAR binding site and the DHS is shown in column J; values of zero indicate lack of DHS overlap for that CAR binding peak. 8,124 of the 14,631 CAR peaks overlap a total of 7,666 unique DHS regions. 441 DHS harbor two CAR binding sites, and 17 DHS harbor three CAR binding sites; further, 21 CAR binding sites overlap two adjacent DHS (also see note below). A total of 3,734 CAR binding sites were significantly induced by TCPOBOP ("induced", column D): 3,021 of the 8,124 CAR binding sites at a DHS (37.2%) + 713 of the 6,507 CAR binding sites not at a DHS (11.0%). The higher percentage of TCPOBOP-induced CSR binding sites at a DHS (37.2 vs. 11.0%) is an indication of their higher quality and greater biological relevance. NOTE: The 7,666 DHS that overlap a CAR binding site, included here, are fewer in number than the 7,687 unique DHS with a CAR binding site shown in Table S5A, where the DHS with the maximum bp overlap with a CAR binding site are listed as the target of

that DHS in case of CAR binding sites that overlap two DHS. There are 21 cases where a single CAR binding site overlaps two neighboring DHS -- both of which are labeled static, induced, or repressed with regard to the overlapping CAR binding site in Table S5B. For this same reason, the number of DHS with an induced CAR binding site (Table S5B, column D) is greater than the 2,901 such DHS shown in Table S5A, column I. Also, see summary in Table S5C.

Table S5C. Shown is the number of CAR binding sites (based on raw ChIP-seq data from Tian et al., 2018) inside and outside of DHS regions (60K DHS set used in this study) with a significant increase (induced) or decrease (repressed) in CAR ChIP-seq signal between TCPOBOP-exposed and control livers, as determined by diffReps analysis (see Methods). Based on data in Tables S5A and S5B.

Table S5D. Chromatin marks at stringent CAR-free DHS compared to robust CARbound DHS, and at Closing DHS. Four sets of DHS were considered: 455 robust CARbound opening DHS, 335 stringent CAR-free opening DHS, and the corresponding two sets of static (i.e., constitutively open) DHS, as shown in column A. The DHS in each set are listed in Table S5A, column V. Shown in rows 4 to 7 are the number of DHS with the indicated pattern of H3K27ac (enhancer) markers, with their responsiveness to TCPOBOP exposure (Induced, Repressed, etc.), based on data shown in Table S4, column AD. Rows 9-12 present this data in the form of percentages, which highlights the relative depletion of DHS with TCPOBOP-induced H3K27ac marks and the relative

increase in DHS without H3K27ac marks in the stringent CAR-free opening DHS compared to the robust CAR-bound opening DHS. Shown in rows 15 to 18 are the number of DHS assigned to the indicated DHS class based on its chromatin marks and their responsiveness to TCPOBOP exposure (Induced enhancer, Induced promoter, etc.), based on data shown in Table S4, column AI. Rows 20-23 present this data in the form of percentages, which highlights the shift from Induced enhancer to Poised enhancer and Lack of chromatin marks for the set of stringent CAR-free opening DHS compared to the robust CAR-bound opening DHS. Major findings: Although the stringently CAR-free opening DHS were strongly enriched for induced enhancer marks as compared to the corresponding set of static DHS, the CAR-free opening DHS frequently lacked H3K27ac marks (40%, as compared to only 9.5% for the CAR-bound DHS set), and correspondingly, fewer of the CAR-free opening DHS were induced enhancer or induced promoter DHS (18.8%, as compared to 48.2% for robust CAR-bound DHS). In contrast, a high fraction of robust CAR-bound static DHS were found at either enhancer DHS (66.1%) or promoter DHS (21.0%) that did not show changes in active chromatin marks (Table S5D). Rows 25-35: Shown in rows 26-27 are the number of chromatin states (as in rows 14-23) for closing DHS, either with and without CAR bound; with CAR binding based on mCAR binding data in Table S5A, rows R and S (i.e., mCAR binding with TCPOBOP or phenobarbital treatment). Results show that a majority of closing DHS are in an enhancer state, with strong enrichment for the Repressed Enhancer state (bold). Rows 38-45: Shown in rows 40-41 are the number of chromatin states (as in rows 14-23) for DHS that map to TCPOBOP-repressed genes (Table S5A, column G), both with and

without CAR bound; with CAR binding based on mCAR binding data in Table S5A, rows R and S (i.e., mCAR binding with TCPOBOP or phenobarbital treatment). Results show that ~85% of DHS mapping to repressed genes and that have CAR bound are in a static enhancer or promoter state.

Table S5E. Robust CAR-binding and stringent CAR-free DHS (from Table S5A, column V). Mapping to TCPOBOP-induced and TCPOBOP-repressed genes, and their enrichment scores.

Table S6A. Two TCPOBOP-responsive gene sets (based on Bhushan et al, 2018 Hepatology) and their associated CAR binding sites. Columns A-H, Gene Set A: Genes induced in mouse liver by TCPOBOP after 2 days, but not significantly changed in expression in TCPOBOP-treated MET-knockout mouse liver in the presence of an EGFR inhibitor (MET-EGFR- double "knockout", DKO) compared to TCPOBOP-treated wildtype controls, as defined by Bhushan et al (2018). Columns J-Q, Gene Set D: Genes whose induction after 2 days of TCPOBOP exposure is inhibited in MET-knockout mouse liver in the presence of an EGFR inhibitor (MET-EGFR- double "knockout", DKO). Also shown for both gene sets are CAR binding sites associated with each gene. CAR binding sites were mapped to the nearest gene in the same TAD. If a gene contains a proximal CAR binding site, then the corresponding CAR peak ID is listed. Genes that lack a proximal CAR binding site are labeled as "none". CAR binding site data is shown for mouse CAR (mCAR) and human CAR (hCAR) based on puvblished CAR ChIp-seq data from Niu et al, NAR 2018.

Table S6B. Two other sets of TCPOBOP-responsive genes (based on Bhushan et al, 2018 Hepatology) and their associated CAR binding sites. Columns A-H, Gene Set A: Genes repressed in mouse liver by TCPOBOP after 2 days, but not significantly changed in expression in TCPOBOP-treated MET-knockout mouse liver in the presence of an EGFR inhibitor (MET-EGFR- double "knockout", DKO) compared to TCPOBOP-treated wild-type controls, as defined by Bhushan et al (2018). Columns J-Q, Gene Set D: Genes whose repression after 2 days of TCPOBOP exposure is increased in MET-knockout mouse liver in the presence of an EGFR inhibitor (MET-EGFR- double "knockout", DKO). Also shown for both gene sets are CAR binding sites associated with each gene. CAR binding sites were mapped to the nearest gene in the same TAD. If a gene contains a proximal CAR binding site, then the corresponding CAR peak ID is listed. Genes that lack a proximal CAR binding site are labeled as "none". CAR binding site data is shown for mouse CAR (mCAR) and human CAR (hCAR) based on published CAR ChIp-seq data from Niu et al, NAR 2018.

Table S6C. Gene sets, from Table S6A and S6B, used in DAVID analysis, shown inTables S6D and S6E

Table S6D. DAVID analysis of genes in set A + set B, shown in Table S6C (593 genes). The TCPOBOP responses of these genes are not dependent on MET, EGFR signaling.

Table S6E. DAVID analysis of genes in set C + set D, shown in Table S6C (369 genes). The TCPOBOP responses of these genes are dependent on MET, EGFR signaling.

Appendix 2 – List of supplemental tables related to Chapter 3.

Supplemental files are available upon request by email to djw@bu.edu

Table S1. Summary of DNase-seq analysis - Shown are the total and mapped read counts for the DNase-seq samples prepared and analyzed in this study, the number of DHS peaks discovered in each sample by MACS2, and the fraction of sequence reads found in the sample's respective peak list. The data for the pituitary-intact male liver (STAT5-high and STAT5-low) samples and the GH time-course mouse liver (hypophysectomized (hypox) and GH treated) samples are shown separately.

Table S2 - Set of 70,211 liver DHS regions used in this study.

Columns A-G: Set of 70,211 liver DHS regions and their DHS ID (column A), coordinates (columns B-D), DHS length (column E), and sex-specificity (column F) in mouse liver as determined by DNase-seq, as reported in (Ling et al., 2010). The set of 72,862 DHS identified from (Ling et al., 2010) was filtered to exclude those sites that overlap ENCODE blacklisted regions and those that were not classified according to a five class DHS model defined by promoter, weak promoter, enhancer, weak enhancer, and insulator DHS (Matthews and Waxman, 2018). The mouse liver TAD where the DHS is found is listed in column G.

Columns H-I: Characterization of DHS as static or dynamic. DNase-seq signal was compared between intact male liver STAT5 high and low samples to determine DHS response to an endogenous pulse of GH/STAT5. The DHS response was either static (labeled "STAT5_sDHS"), opened (labeled "STAT5_hi_dDHS") or closed ("STAT5_lo_dDHS") in response to an endogenous pulse of GH/STAT5 (column H). The sex-specificity of the DHS was simplified to be sex-independent, male-biased or female-biased comprised of the standard and robust subsets indicated from column F. The sex-specificity of the DHS (sex_indep, male_bias, female_bias) was then classified as being either static (based on the DHS being a STAT5_sDHS) or dynamic (based on the DHS being a STAT5_hi_dDHS).

Columns J-N: Characterization of DHS based the presence of STAT5 binding. ChIP-seq analysis of STAT5 binding in male and female mouse liver previously identified 15,094 merged peaks comprised of male and female enriched and male-female common STAT5 binding sites (Zhang et al., 2012). The DHS used in this study were labeled as either having STAT5 bound or not bound (column J). For DHS that contain a STAT5 binding site, the sex-specificity (male enriched, female enriched, or common) of the binding site is shown (column K). The normalized STAT5 ChIP-seq read counts for the male-high (MH) and female-high (FH) samples are shown (columns L-M) (obtained from Table S1F, (Zhang et al., 2012)). The average of the MH and FH STAT5 ChIP-seq read counts are shown in column N.

Column O: The frequency of the STAT5 motif in DHS. The STAT5B motif M00459 from the TRANSFAC motif database (Release 2011.1) (Matys et al., 2006) was used to determine the frequency of STAT5 motif occurrence in DHS sequences. Motifs present in DHS sequences were identified using FIMO (v4.12.0) (Grant et al., 2011) with the option (--thres 0.0005) to improve detection of short length motifs. The number of STAT5B motif occurrences in each of the 70,211 DHS sequences is shown in column O.

Columns P-S: Functional classification of DHS. DHS were classified according to a five class DHS model defined by promoter, weak promoter, enhancer, weak enhancer, and insulator DHS based on ChIP-seq signals for H3K4me1, H3K4me3, CTCF binding, and proximity to RefSeq gene transcription start sites (TSS) (Matthews and Waxman, 2018). The DHS used in this study were classified as either weak enhancer (WE), enhancer (Enh), insulator (Insul), weak promoter (WP), and promoter (Prom) based on overlap. The original class label, abbreviated class name, and full class name are shown in columns P-R respectively. A simplification of the class labels is shown in column S; enhancer and promoter labels include the subset of sites labeled as "weak" for the respective class.

Columns T-AE: Putative gene target of the DHS. Each of the 70,211 liver DHS was assigned a single putative gene target (RefSeq gene or multi-exonic lncRNA gene) that corresponds to the closest TSS within the same topologically associating domain (TAD). The nearest gene within the same TAD to each DHS is listed in column X, gene TSS coordinates in columns T-W, TAD assignment in column Y, and the linear genomic distance is listed in column Z. Columns AA-AE: summary results from RNA-seq between intact male (Intensity 2) and intact female (Intensity 1) liver samples. RNA-seq results include FPKM (fragments per kilobase per million reads) in male (column AB) and female (column AA) mouse liver, ratio, fold-change, and adjusted p-value (FDR).

Columns AF-AH: Sex bias and class of the putative gene target. RNA-seq data from columns AA-AE was used to define sex-specific and sex-independent genes as follows: sex-specific genes (labeled "Male" or "Female") were defined by |FC| > 1.5 (RefSeq) or |FC| > 2.0 (lncRNA), p-value < 0.05, and FPKM > 0.25 cutoff for the sex with greater signal intensity. Sex-independent genes (labeled "sex_indep") are defined by |FC| < 1.2, p-value > 0.1, and FPKM > 1 in both male and female liver. Genes labeled as "None" were neither sex-specific or sex-independent (column AF). Columns AG-AH indicate the class and subclass designations for the putative gene target using the same definitions and cutoffs described in (Connerney et al., 2017). Sex-biased genes responsive to hypox were defined by |fold-change| > 2 and an adjusted p-value < 0.05 for hypox versus intact mouse control in male liver, or separately, in female liver. Using these thresholds, RNA-seq data from intact and hypox male and female liver samples were used to classify sex-

specific genes into classes I and II and their corresponding subclasses (IA, IB, IC, IIA, and IIB) (Connerney et al., 2017). Class I sex-biased genes are those that are downregulated by hypox in the sex where they show higher expression in intact mice. Class II sex-biased genes are those that are upregulated by hypox in the sex where they show lower expression in intact mice. Subclasses A, B, and C indicate the response to hypox in the dominant sex (class II genes) or in the opposite sex (class I genes).

Columns AI-AJ: Mouse liver DHS response to hypophysectomy (hypox). The effect of hypox was determined by comparing hypox liver samples to the corresponding intact liver (control) samples in male and female liver. DHS signal was compared using diffReps with the nucleosome option (200 bp window). Significance was based on |fold-change| > 2 and FDR < 0.05 (Benjamini-Hochberg adjusted p-value) for diffReps-normalized signal intensity values. Each of the 70,211 liver DHS was then labeled based on whether it overlapped with a diffReps-identified dDHS that opens following hypox ("dDHS_open"), a diffReps dDHS that closes following hypox ("dDHS_close"), a static DHS ("static"), or "none", for DHS on the list of 70,211 that do not overlap the DHS peak union list for the corresponding dataset.

Columns AK-AM: Hypophysectomized (hypox) male mouse liver DHS response to a single injection of GH. Hypox male mice were given a single dose of GH and euthanized either 30, 90, or 240 minutes later. The DHS signal intensity of hypox+GH treated male liver samples were compared to male hypox controls with diffReps using the same

parameters described above; nucleosome option (200 bp window), |fold-change| > 2 and FDR < 0.05 (Benjamini-Hochberg adjusted p-value). The DHS response at each time point is indicated in columns AK-AM and is summarized using the same procedure described for columns AI-AJ.

Columns AN-AO: Chromatin state assignment for each of the 70,211 liver DHS. Chromatin state maps (14 state model) were previously developed for male mouse liver, and separately for female mouse liver, using a panel of six histone marks and DHS data, and used to identify sex differences in chromatin state and chromatin structure and their relationships to sex-biased gene expression (Sugathan and Waxman, 2013). Coordinates for the male liver chromatin state map were overlapped with the DHS used in this study, the chromatin state assignment for each DHS is shown in column AN. An abbreviated version of the chromatin state name is shown as E ("emission") followed by the state number, for example, chromatin state 12 is abbreviated as "E12" (column AO).

Column AP: A summary of the DHS response to hypophysectomy and a single injection of GH in hypox male mouse liver. A four-digit total flag sum (TFS) number was generated for each DHS (see Methods). Positions 1-4 of the TFS value correspond to the following data sets: MHx/Male control, MHx+GH/MHx at 30, 90, and 240 minutes, respectively. These TFS values were used to determine the overlap between STAT5-high Δ DHS that were closed by hypophysectomy and/or opened by a single exogenous GH pulse. A binary definition was used to summarize the DHS response in each data set. For
MHx/Male control: a value of 1 was assigned if the DHS was closed due to hypophysectomy or 0 otherwise, for the GH time-course data sets: a value of 1 was assigned if the DHS was opened by a single exogenous GH pulse or 0 otherwise. Shown are the TFS values for the subset of STAT5_hi_dDHS (n = 2,373).

Columns AQ-AY: RiPPM Normalized DNase-seq read counts in the DHS. Raw DNaseseq read counts were normalized using sequence reads in each DHS peak region per million mapped sequenced reads (reads-in-peaks-per-million, RiPPM) as a scaling factor. Raw read counts were divided by this per-million scaling factor to obtain RiPPM normalized read counts for male and female liver samples. RiPPM normalized DNase-seq read counts are shown for the combined samples for each data set. Data for pituitaryintact male liver samples are shown in columns AQ-AR with a log2(fold-change) calculated from the following ratio: (Male STAT5 Hi)/(Male STAT5 Lo) (column AS). Data for the hypox and hypox+GH male liver samples are shown in columns AT, AW-AY and female liver samples in columns AU (pituitary-intact) and AV (hypophysectomized).

Table S3 - The set of RefSeq (n=24,197) and multi-exonic lncRNA (n=3,152) genes and their corresponding sex-bias classification derived from gene expression data from mouse liver. Data shown are normalized differential expression ratios, calculated as intact-male/intact-female, corresponding fold change (FC) values, normalized read counts (FPKM; fragments per kilobase of region of interest per million mapped reads) for male and female liver, and FDR (adjusted p-values) for the comparison. Collapsed exon regions were considered for counting regions using FeatureCounts (Liao, 2014, Bioinformatics). Also shown are the mouse genomic coordinates (genome release mm9) (Mapping Location(s)), NCBI accession number, official gene symbol, and Gene Ontology (GO) terms for each gene. The Gene Notes column indicates which mouse TAD the gene is located in (TADs numbered sequentially from 1 to approx. 3660, based on their location in the mouse genome), which strand the gene is transcribed from, overlapping genes and isoforms, etc. The Gene Type (column H) indicates whether the gene is a RefSeq protein coding or multi-exonic lncRNA gene. Sex-specific genes (labeled "Male" or "Female") were defined by |FC| > 1.5 (RefSeq) or |FC| > 2.0(lncRNA), p-value < 0.05, and FPKM > 0.25 cutoff for the sex with greater signal intensity. Sex-independent genes (labeled "sex_indep") are defined by |FC| < 1.2, p-value > 0.1, and FPKM > 1 in both male and female liver. Genes labeled as "None" were neither sex-specific or sex-independent (column N). Class I sex-biased genes are those that are downregulated by hypox in the sex where they show higher expression in intact mice. Class II sex-biased genes are those that are upregulated by hypox in the sex where they show lower expression in intact mice. Subclasses A, B, and C indicate the response to hypox in the dominant sex (class II genes) or in the opposite sex (class I genes). Shown is the class (I or II) and subclass (IA, IB, IC, IIA, IIB) for each gene. Only sexspecific genes have class and subclass designations while all other genes are labeled as "none" (columns O-P). The following SEGEX experiments were used for defining sexspecific and sex-independent genes and those that respond to hypophysectomy in male and female liver for RefSeq and multi-exonic lncRNA genes.

Table S4 DHS Peak Summary – Sets of DNase-seq peaks discovered for the pituitary-intact male liver (STAT5-high and STAT5-low) samples and the GH timecourse mouse liver (hypophysectomized (hypox) and GH treated) samples. Each sheet contains the peak union list (see Methods) generated from the respective data set and indicates a single overlapping DHS for each peak union site.

Table S4A - STAT5-high/STAT5-low male liver DHS overlap with the set of 70,211 DHS. Merged list of 70,767 DHS regions were generated based on the MACS2 peaks called in each DNase-seq replicate sample (see Methods) and assigned a unique name (DHS Peak Number, column E). The genomic position of DHS region is listed in columns A-C, its DHS response (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of overlap between DHS regions.

Table S4B - MHx/Male Control liver DHS overlap with the set of 70,211 DHS.

Merged list of 57,053 DHS regions were generated based on the MACS2 peaks called in each DNase-seq replicate sample (see Methods) and assigned a unique name (DHS Peak Number, column E). The genomic position of DHS region is listed in columns A-C, its DHS response (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of overlap between DHS regions.

Table S4C - FHx/Female Control liver DHS overlap with the set of 70,211 DHS.

Merged list of 43,287 DHS regions were generated based on the MACS2 peaks called in each DNase-seq replicate sample (see Methods) and assigned a unique name (DHS Peak Number, column E). The genomic position of DHS region is listed in columns A-C, its DHS response (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of overlap between DHS regions.

Table S4D - MHx+GH 30 min/MHx liver DHS overlap with the set of 70,211 DHS.

Merged list of 33,241 DHS regions were generated based on the MACS2 peaks called in each DNase-seq replicate sample (see Methods) and assigned a unique name (DHS Peak Number, column E). The genomic position of DHS region is listed in columns A-C, its DHS response (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of overlap between DHS regions.

Table S4E - MHx+GH 90 min/MHx liver DHS overlap with the set of 70,211 DHS.

Merged list of 33,593 DHS regions were generated based on the MACS2 peaks called in each DNase-seq replicate sample (see Methods) and assigned a unique name (DHS Peak Number, column E). The genomic position of DHS region is listed in columns A-C, its DHS response (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of overlap between DHS regions.

Table S4F - MHx+GH 240 min/MHx liver DHS overlap with the set of 70,211 DHS.

Merged list of 33,563 DHS regions were generated based on the MACS2 peaks called in each DNase-seq replicate sample (see Methods) and assigned a unique name (DHS Peak Number, column E). The genomic position of DHS region is listed in columns A-C, its DHS response (column D), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of overlap between DHS regions.

Table S5A - Chromatin states in male mouse liver. Chromatin state (CS) maps (14 state model) were previously developed for male mouse liver, and separately for female mouse liver, using a panel of six histone marks and DHS data, and used to identify sex differences in chromatin state and chromatin structure and their relationships to sexbiased gene expression (Sugathan and Waxman, 2013). BEDTools was used to determine the overlap between the 14 chromatin states identified in male liver and the set of 70,211 DHS used in this study. The genomic position of chromatin state is listed in columns A-C, corresponding name (column D), a unique ID (CS Number, column E), and its

overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin state region.

Table S5B - Chromatin states in female mouse liver. Chromatin state (CS) maps (14 state model) were previously developed for male mouse liver, and separately for female mouse liver, using a panel of six histone marks and DHS data, and used to identify sex differences in chromatin state and chromatin structure and their relationships to sexbiased gene expression (Sugathan and Waxman, 2013). BEDTools was used to determine the overlap between the 14 chromatin states identified in female liver and the set of 70,211 DHS used in this study. The genomic position of chromatin state is listed in columns A-C, corresponding name (column D), a unique ID (CS Number, column E), and its overlapping DHS region (if exists) (column F-I). The number of overlapping base pairs is provided in column J; values of zero indicate lack of DHS overlap for that chromatin state region.

Table S6 - Summary of DNase-seq aggregate plots - Shown is the maximum value for each DNase-I cut site aggregate plot for each of the DNase-seq samples indicated; separated by figure number. Briefly, normalized DNase-I cut site aggregate plots were generated using input DNase-seq datasets (columns C-H) and a set of input genomic regions (DHS sequences) (columns A-B) used for read counting. A single replicate combined sample was generated for each of the following groups: male STAT5-high, male STAT5-low, female control, hypophysectomized male (MHx) and female (FHx), and MHx samples treated with GH (MHx+GH) at the following time points: 30, 90, 240 minutes. These replicate combined samples were then used for obtaining the number of DNase-I cuts at each nucleotide position of the 2 kb midpoint-centered regions of the DHS specified (see Methods). The maximum value of the smoothed cumulative DNase-I cutting profile in each of the figures are shown below.

Table S7A - Summary of genomic sites used for DHS enrichment calculations.

Enrichment calculations were performed using the following sets of DHS: (1) dynamic male DHS (834 sites), (2) static male DHS (1,895 sites), (3) dynamic sex-independent DHS (1,532 sites), and (4) static female DHS (1,359 sites) with static sex-independent DHS (64,584 sites) as the background. Shown are the biologically relevant regions used for the enrichment calculations; sex-biased transcription factor binding sites (columns A-C), sex-biased chromatin mark regions (columns E-F), and chromatin states in male and female liver (columns G-J). For each set of biologically relevant regions, the name, number of sites, and source publication is indicated. The STAT5-High (Male) (1,093 sites) and STAT5-High (Female) (1,174 sites) are the subset of STAT5 binding sites that also contain a STAT5 motif.

Table S7B - Summary of DHS enrichment calculations for the biologically relevant regions defined in Table S7A. Enrichment calculations were performed using the following sets of DHS: (1) dynamic male DHS (834 sites), (2) static male DHS (1,895 sites), (3) dynamic sex-independent DHS (1,532 sites), and (4) static female DHS (1,359 sites) with static sex-independent DHS (64,584 sites) as the background. DHS (64,584 sites) as the background. Shown is the calculated enrichment score (ES) and the Fisher Exact test p-value for the overlap between the DHS and each set of biologically relevant regions; sex-biased transcription factor binding sites (rows 2-11), sex-biased chromatin mark regions (rows 13-26), and chromatin states in male and female liver (rows 28-60).

Table S7C - Full listing of DHS enrichment calculations as shown in Table S7B.

Enrichment calculations were performed using the following sets of DHS: (1) dynamic male DHS (834 sites), (2) static male DHS (1,895 sites), (3) dynamic sex-independent DHS (1,532 sites), and (4) static female DHS (1,359 sites) with static sex-independent DHS (64,584 sites) as the background. Shown is the calculated enrichment score (ES), the number and percent of overlapping DHS, the number and percent of overlapping background DHS, and the Fisher Exact test p-value for the overlap between the DHS and each set of biologically relevant regions; sex-biased transcription factor binding sites (rows 2-11), sex-biased chromatin mark regions (rows 13-26), and chromatin states in male and female liver (rows 28-60).

 Table S7D - Coordinates for each of the sex-biased transcription factor binding sites

 defined in Table S7A. Shown is the chromosome, start, and end coordinates for each of

 the sites indicated (formatted as a 3-column BED file).

 Table S7E - Coordinates for each of the sex-biased chromatin mark regions in male

 liver defined in Table S7A. Shown is the chromosome, start, and end coordinates for

 each of the sites indicated (formatted as a 3-column BED file).

 Table S7F - Coordinates for each of the sex-biased chromatin mark regions in

 female liver defined in Table S7A. Shown is the chromosome, start, and end

 coordinates for each of the sites indicated (formatted as a 3-column BED file).

Table S8A - MAnorm comparative analysis of FoxA1 ChIP-seq peaks between male and female mouse liver. Raw data from (Li et al., 2012) was processed and reanalyzed with MAnorm to identify sex-specific FoxA1 binding sites in mouse liver. FoxA1 ChIPseq samples in male and female liver were defined as sample 1 and sample 2 respectively. Shown is the MAnorm output which lists the peak coordinates, raw reads, M and A values, and MA-norm p-value for the set of common peaks in each sample and for the peaks unique to each sample. M-values are defined as the log2 fold change and A-values are defined as the average signal strength of the of normalized read densities under comparison. Sex-specific FoxA1 binding sites defined in Table S7A were extracted as MAnorm differential sites based on the M-value. The subset of sites with M-value > +1 are defined as FoxA1 (Male) (3,280 sites) and the subset of sites with M-value < -1 are defined as FoxA1 (Female) (577 sites). Table S8B - MAnorm comparative analysis of FoxA2 ChIP-seq peaks between male and female mouse liver. Raw data from (Li et al., 2012) was processed and reanalyzed with MAnorm to identify sex-specific FoxA2 binding sites in mouse liver. FoxA2 ChIPseq samples in male and female liver were defined as sample 1 and sample 2 respectively. Shown is the MAnorm output which lists the peak coordinates, raw reads, M and A values, and MA-norm p-value for the set of common peaks in each sample and for the peaks unique to each sample. M-values are defined as the log2 fold change and A-values are defined as the average signal strength of the of normalized read densities under comparison. Sex-specific FoxA2 binding sites defined in Table S7A were extracted as MAnorm differential sites based on the M-value. The subset of sites with M-value > +1 are defined as FoxA2 (Male) (976 sites) and the subset of sites with M-value < -1 are defined as FoxA2 (Female) (780 sites).

LIST OF JOURNAL ABBREVIATIONS

Annu Rev Genomics Hum Genet	. Annual Review of Genomics and Human Genetics
Arch Biochem Biophys	Archives of Biochemistry and Biophysics
Chem Res Toxicol	Chemical Research in Toxicology
Clin Epigenetics	Clinical Epigenetics
Curr Drug Targets	Current Drug Targets
Drug Discov Today	Drug Discovery Today
Drug Metab Dispos	Drug Metabolism and Disposition
Drug Metab Rev	Drug Metabolism Reviews
Endocrinology	Journal of Endocrinology
Environ Epigenet	Environmental Epigenetics
Environ Health Perspect	Environmental Health Perspectives
Expert Opin Drug Metab Toxicol Expert Opinion on Drug Metabolism & Toxicology	
Front Genet	Frontiers in Genetics
Genome Biol	Genome Biology
Genome Res	Genome Research
Hepatology	Journal of Hepatology
Hum Genet	Human Genetics
J Biol Chem	Journal of Biological Chemistry
J Endocrinol	Journal of Endocrinology
Mamm Genome	Mammalian Genome
Mol Cell Biol.	Molecular and Cellular Biology

Mol Cell Endocrinol	Molecular and Cellular Endocrinology
Mol Endocrinol	Molecular Endocrinology
Mol Pharmacol	Molecular Pharmacology
Mutat Res.	Mutation Research
Nat Commun	Nature Communications
Nat Protoc	Nature Protocols
Nat Rev Genet	Nature Reviews Genetics
Nat. Protocols	Nature Protocols
Nucl Recept Signal	Nuclear Receptor Signaling
Nucleic Acids Res	Nucleic Acids Research
PPAR Res	
Proc Natl Acad Sci Proceedings of the	National Academy of Sciences of the USA
Sci Signal	Science Signaling
Semin Cell Dev BiolSe	eminars in Cell and Developmental Biology
TEM	Trends in endocrinology and metabolism
Toxicol Appl Pharmacol	Toxicology and Applied Pharmacology
Toxicol Sci	Toxicological Sciences
Tumour Biol	Tumor Biology

REFERENCES

- (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57-74.
- Aaronson, D.S., and Horvath, C.M. (2002). A road map for those who don't know JAK-STAT. Science 296, 1653-1655.
- Abe, T., Amaike, Y., Shizu, R., Takahashi, M., Kano, M., Hosaka, T., Sasaki, T., Kodama, S., Matsuzawa, A., and Yoshinari, K. (2018). Role of YAP Activation in Nuclear Receptor CAR-Mediated Proliferation of Mouse Hepatocytes. Toxicol Sci 165, 408-419.
- Al-Salman, F., and Plant, N. (2012). Non-coplanar polychlorinated biphenyls (PCBs) are direct agonists for the human pregnane-X receptor and constitutive androstane receptor, and activate target gene expression in a tissue-specific manner. Toxicol Appl Pharmacol 263, 7-13.
- Aleksunes, L.M., and Klaassen, C.D. (2012). Coordinated regulation of hepatic phase I and II drug-metabolizing genes and transporters using AhR-, CAR-, PXR-, PPARalpha-, and Nrf2-null mice. Drug Metab Dispos 40, 1366-1379.
- Ashrap, P., Zheng, G., Wan, Y., Li, T., Hu, W., Li, W., Zhang, H., Zhang, Z., and Hu, J. (2017). Discovery of a widespread metabolic pathway within and among phenolic xenobiotics. Proc Natl Acad Sci U S A 114, 6062-6067.
- Ayuk, J., and Sheppard, M.C. (2006). Growth hormone and its disorders. Postgraduate medical journal 82, 24-30.
- Baldwin, W.S., and Roling, J.A. (2009). A concentration addition model for the activation of the constitutive androstane receptor by xenobiotic mixtures. Toxicol Sci *107*, 93-105.
- Banerjee, M., Robbins, D., and Chen, T. (2015). Targeting xenobiotic receptors PXR and CAR in human diseases. Drug discovery today 20, 618-628.
- Barakat, T.S., Halbritter, F., Zhang, M., Rendeiro, A.F., Perenthaler, E., Bock, C., and Chambers, I. (2018). Functional Dissection of the Enhancer Repertoire in Human Embryonic Stem Cells. Cell stem cell 23, 276-288.e278.
- Benet, M., Lahoz, A., Guzman, C., Castell, J.V., and Jover, R. (2010). CCAAT/enhancerbinding protein alpha (C/EBPalpha) and hepatocyte nuclear factor 4alpha (HNF4alpha) synergistically cooperate with constitutive androstane receptor to transactivate the human cytochrome P450 2B6 (CYP2B6) gene: application to the

development of a metabolically competent human hepatic cell model. J Biol Chem 285, 28457-28471.

- Bhushan, B., Stoops, J.W., Mars, W.M., Orr, A., Bowen, W.C., Paranjpe, S., and Michalopoulos, G.K. (2018). TCPOBOP-induced hepatomegaly & hepatocyte proliferation is attenuated by combined disruption of MET & EGFR signaling. Hepatology.
- Blazquez, M., Carretero, A., Ellis, J.K., Athersuch, T.J., Cavill, R., Ebbels, T.M., Keun, H.C., Castell, J.V., Lahoz, A., and Bort, R. (2013). A combination of transcriptomics and metabolomics uncovers enhanced bile acid biosynthesis in HepG2 cells expressing CCAAT/enhancer-binding protein beta (C/EBPbeta), hepatocyte nuclear factor 4alpha (HNF4alpha), and constitutive androstane receptor (CAR). Journal of proteome research *12*, 2732-2741.
- Brocker, C.N., Yue, J., Kim, D., Qu, A., Bonzo, J.A., and Gonzalez, F.J. (2017). Hepatocyte-specific PPARA expression exclusively promotes agonist-induced cell proliferation without influence from nonparenchymal cells. American journal of physiology. Gastrointestinal and liver physiology *312*, G283-g299.
- Buchman, C.D., Chai, S.C., and Chen, T. (2018). A current structural perspective on PXR and CAR in drug metabolism. Expert Opin Drug Metab Toxicol *14*, 635-647.
- Chang, T.K.H., and Waxman, D.J. (2006). Synthetic drugs and natural products as modulators of constitutive androstane receptor (CAR) and pregnane X receptor (PXR). Drug Metab Rev *38*, 51-73.
- Chatterjee, S., and Ahituv, N. (2017). Gene Regulatory Elements, Major Drivers of Human Disease. Annu Rev Genomics Hum Genet 18, 45-63.
- Chen, W.D., Fu, X., Dong, B., Wang, Y.D., Shiah, S., Moore, D.D., and Huang, W. (2012). Neonatal activation of the nuclear receptor CAR results in epigenetic memory and permanent change of drug metabolism in mouse liver. Hepatology 56, 1499-1508.
- Choi, H.K., and Waxman, D.J. (2000). Plasma growth hormone pulse activation of hepatic JAK-STAT5 signaling: developmental regulation and role in male-specific liver gene expression. Endocrinology *141*, 3245-3255.
- Christensen, B.C., and Marsit, C.J. (2011). Epigenomics in environmental health. Front Genet 2, 84.
- Clodfelter, K.H., Holloway, M.G., Hodor, P., Park, S.-H.H., Ray, W.J., and Waxman, D.J. (2006). Sex-dependent liver gene expression is extensive and largely dependent upon signal transducer and activator of transcription 5b (STAT5b): STAT5b-

dependent activation of male genes and repression of female genes revealed by microarray analysis. Mol Endocrinol 20, 1333-1351.

- Clodfelter, K.H., Miles, G.D., Wauthier, V., Holloway, M.G., Zhang, X., Hodor, P., Ray, W.J., and Waxman, D.J. (2007). Role of STAT5a in regulation of sex-specific gene expression in female but not male mouse liver revealed by microarray analysis. Physiol Genomics 31, 63-74.
- Conforto, T.L., Zhang, Y., Sherman, J., and Waxman, D.J. (2012). Impact of CUX2 on the Female Mouse Liver Transcriptome: Activation of Female-biased Genes and Repression of Male-biased Genes. Mol Cell Biol.
- Connerney, J., Lau-Corona, D., Rampersaud, A., and Waxman, D.J. (2017). Activation of Male Liver Chromatin Accessibility and STAT5-Dependent Gene Transcription by Plasma Growth Hormone Pulses. Endocrinology 158, 1386-1405.
- Cortessis, V.K., Thomas, D.C., Levine, A.J., Breton, C.V., Mack, T.M., Siegmund, K.D., Haile, R.W., and Laird, P.W. (2012). Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. Hum Genet 131, 1565-1589.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., *et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America 107, 21931-21936.
- Cui, J.Y., and Klaassen, C.D. (2016). RNA-Seq reveals common and unique PXR- and CAR-target gene signatures in the mouse liver transcriptome. Biochimica et biophysica acta 1859, 1198-1217.
- DeKeyser, J.G., Laurenzana, E.M., Peterson, E.C., Chen, T., and Omiecinski, C.J. (2011). Selective phthalate activation of naturally occurring human constitutive androstane receptor splice variants and the pregnane X receptor. Toxicol Sci 120, 381-391.
- Delgado, I., Fresnedo, O., Iglesias, A., Rueda, Y., Syn, W.K., Zubiaga, A.M., and Ochoa, B. (2011). A role for transcription factor E2F2 in hepatocyte proliferation and timely liver regeneration. American journal of physiology. Gastrointestinal and liver physiology 301, G20-31.
- Dempsey, J.L., and Cui, J.Y. (2018). Regulation of Hepatic Long Non-coding RNAs by PXR and CAR Agonists in Mouse Liver. Drug Metab Dispos.

- Dixit, S.G., Tirona, R.G., and Kim, R.B. (2005). Beyond CAR and PXR. Current drug metabolism *6*, 385-397.
- Dong, B., Lee, J.S., Park, Y.Y., Yang, F., Xu, G., Huang, W., Finegold, M.J., and Moore, D.D. (2015). Activating CAR and beta-catenin induces uncontrolled liver growth and tumorigenesis. Nat Commun 6, 5944.
- Eckstein, M., Eleazer, R., Rea, M., and Fondufe-Mittendorf, Y. (2017). Epigenomic reprogramming in inorganic arsenic-mediated gene expression patterns during carcinogenesis. Reviews on environmental health *32*, 93-103.
- Edwards, T.M., and Myers, J.P. (2007). Environmental exposures and gene regulation in disease etiology. Environmental health perspectives *115*, 1264-1270.
- Ernst, J., and Kellis, M. (2017). Chromatin-state discovery and genome annotation with ChromHMM. Nat Protoc *12*, 2478-2492.
- Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat Protoc 7, 1728-1740.
- Frank, C., Gonzalez, M.M., Oinonen, C., Dunlop, T.W., and Carlberg, C. (2003). Characterization of DNA complexes formed by the nuclear receptor constitutive androstane receptor. The Journal of biological chemistry 278, 43299-43310.
- Fu, S., Wang, Q., Moore, J.E., Purcaro, M.J., Pratt, H.E., Fan, K., Gu, C., Jiang, C., Zhu, R., Kundaje, A., *et al.* (2018). Differential analysis of chromatin accessibility and histone modifications for predicting mouse developmental enhancers. Nucleic Acids Res 46, 11184-11201.
- Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics 27, 1017-1018.
- Grossman, S.R., Zhang, X., Wang, L., Engreitz, J., Melnikov, A., Rogov, P., Tewhey, R., Isakova, A., Deplancke, B., Bernstein, B.E., *et al.* (2017). Systematic dissection of genomic features determining transcription factor binding and enhancer function. Proc Natl Acad Sci U S A *114*, E1291-E1300.
- Guerrero-Bosagna, C., and Skinner, M.K. (2012). Environmentally induced epigenetic transgenerational inheritance of phenotype and disease. Mol Cell Endocrinol *354*, 3-8.
- Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. Genome Biol 8, R24-R24.

- He, H.H., Meyer, C.A., Chen, M.W., Jordan, V.C., Brown, M., and Liu, X.S. (2012). Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. Genome research 22, 1015-1025.
- Heindel, J.J., McAllister, K.A., Worth, L., Jr., and Tyson, F.L. (2006). Environmental epigenomics, imprinting and disease susceptibility. Epigenetics : official journal of the DNA Methylation Society *1*, 1-6.
- Helsley, R.N., and Zhou, C. (2017). Epigenetic impact of endocrine disrupting chemicals on lipid homeostasis and atherosclerosis: a pregnane X receptor-centric view. Environ Epigenet 3.
- Hernandez, J.P., Mota, L.C., Huang, W., Moore, D.D., and Baldwin, W.S. (2009). Sexually dimorphic regulation and induction of P450s by the constitutive androstane receptor (CAR). Toxicology 256, 53-64.
- Herrington, J., Smit, L.S., Schwartz, J., and Carter-Su, C. (2000). The role of STAT proteins in growth hormone signaling. Oncogene *19*, 2585-2597.
- Holloway, M.G., Laz, E.V., and Waxman, D.J. (2006). Codependence of growth hormone-responsive, sexually dimorphic hepatic gene expression on signal transducer and activator of transcription 5b and hepatic nuclear factor 4alpha. Mol Endocrinol 20, 647-660.
- Honkakoski, P., Zelko, I., Sueyoshi, T., and Negishi, M. (1998). The nuclear orphan receptor CAR-retinoid X receptor heterodimer activates the phenobarbital-responsive enhancer module of the CYP2B gene. Molecular and cellular biology *18*, 5652-5658.
- Hu, C., Liu, D., Zhang, Y., Lou, G., Huang, G., Chen, B., Shen, X., Gao, M., Gong, W., Zhou, P., *et al.* (2014). LXRalpha-mediated downregulation of FOXM1 suppresses the proliferation of hepatocellular carcinoma cells. Oncogene *33*, 2888-2897.
- Hurst, C.H., and Waxman, D.J. (2004). Environmental phthalate monoesters activate pregnane X receptor-mediated transcription. Toxicol Appl Pharmacol *199*, 266-274.
- Ito, Y., Nakamura, T., Y., Y., Ramdhan, D.H., Yamagishi, N., Naito, H., Kamijima, M., Gonzalez, F.J., and Nakajima, T. (2012). Plasticizers May Activate Human Hepatic Peroxisome Proliferator-Activated Receptor alpha Less Than That of a Mouse but May Activate Constitutive Androstane Receptor in Liver. PPAR Res 2012, 11.

- Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J.A., Mahony, S., Pugh, B.F., Lee, D., Kaestner, K.H., and Zaret, K.S. (2016). The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. Molecular cell 62, 79-91.
- Jackson, J.P., Ferguson, S.S., Negishi, M., and Goldstein, J.A. (2006). Phenytoin induction of the cyp2c37 gene is mediated by the constitutive androstane receptor. Drug Metab Dispos *34*, 2003-2010.
- Jansson, J.O., Ekberg, S., Isaksson, O., Mode, A., and Gustafsson, J.A. (1985). Imprinting of growth hormone secretion, body growth, and hepatic steroid metabolism by neonatal testosterone. Endocrinology 117, 1881-1889.
- Ji, H., Jiang, H., Ma, W., and Wong, W.H. (2011). Using CisGenome to analyze ChIPchip and ChIP-seq data. Current protocols in bioinformatics *Chapter 2*, Unit2.13.
- Jimenez-Chillaron, J.C., Nijland, M.J., Ascensao, A.A., Sardao, V.A., Magalhaes, J., Hitchler, M.J., Domann, F.E., and Oliveira, P.J. (2015). Back to the future: transgenerational transmission of xenobiotic-induced epigenetic remodeling. Epigenetics 10, 259-273.
- Jin, J., Wang, G.L., Iakova, P., Shi, X., Haefliger, S., Finegold, M., and Timchenko, N.A. (2010). Epigenetic changes play critical role in age-associated dysfunctions of the liver. Aging Cell 9, 895-910.
- Kadoch, C., Copeland, R.A., and Keilhack, H. (2016). PRC2 and SWI/SNF Chromatin Remodeling Complexes in Health and Disease. Biochemistry *55*, 1600-1614.
- Kanno, Y., Suzuki, M., Nakahama, T., and Inouye, Y. (2005). Characterization of nuclear localization signals and cytoplasmic retention region in the nuclear receptor CAR. Biochimica et biophysica acta 1745, 215-222.
- Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Cheneby, J., Kulkarni, S.R., Tan, G., *et al.* (2018). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic acids research 46, D260-D266.
- Kim, J., Min, G., and Kemper, B. (2001). Chromatin assembly enhances binding to the CYP2B1 phenobarbital-responsive unit (PBRU) of nuclear factor-1, which binds simultaneously with constitutive androstane receptor (CAR)/retinoid X receptor (RXR) and enhances CAR/RXR-mediated activation of the PBRU. The Journal of biological chemistry 276, 7559-7567.
- Kobayashi, K., Hashimoto, M., Honkakoski, P., and Negishi, M. (2015). Regulation of gene expression by CAR: an update. Archives of toxicology *89*, 1045-1055.

- Lake, B.G. (2018). Human relevance of rodent liver tumour formation by constitutive androstane receptor (CAR) activators. Toxicology research 7, 697-717.
- Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., *et al.* (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res 22, 1813-1831.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memoryefficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.
- Latchney, S.E., Fields, A.M., and Susiarjo, M. (2018). Linking inter-individual variability to endocrine disruptors: insights for epigenetic inheritance. Mamm Genome 29, 141-152.
- Lau-Corona, D., Suvorov, A., and Waxman, D.J. (2017). Feminization of male mouse liver by persistent growth hormone stimulation: Activation of sex-biased transcriptional networks and dynamic changes in chromatin states. Molecular and cellular biology.
- Laurenzana, E.M., Coslo, D.M., Vigilar, M.V., Roman, A.M., and Omiecinski, C.J. (2016). Activation of the Constitutive Androstane Receptor by Monophthalates. Chem Res Toxicol 29, 1651-1661.
- Lempiainen, H., Muller, A., Brasa, S., Teo, S.S., Roloff, T.C., Morawiec, L., Zamurovic, N., Vicart, A., Funhoff, E., Couttet, P., *et al.* (2011). Phenobarbital mediates an epigenetic switch at the constitutive androstane receptor (CAR) target gene Cyp2b10 in the liver of B6C3F1 mice. PLoS One 6, e18216.
- Li, Z., Tuteja, G., Schug, J., and Kaestner, K.H. (2012). Foxa1 and Foxa2 are essential for sexual dimorphism in liver cancer. Cell *148*, 72-83.
- Libbrecht, M.W., Ay, F., Hoffman, M.M., Gilbert, D.M., Bilmes, J.A., and Noble, W.S. (2015). Joint annotation of chromatin state and chromatin conformation reveals relationships among domain types and identifies domains of cell-type-specific expression. Genome Res 25, 544-557.
- Lickteig, A.J., Csanaky, I.L., Pratt-Hyatt, M., and Klaassen, C.D. (2016). Activation of Constitutive Androstane Receptor (CAR) in Mice Results in Maintained Biliary Excretion of Bile Acids Despite a Marked Decrease of Bile Acids in Liver.

Toxicological sciences : an official journal of the Society of Toxicology *151*, 403-418.

- Ling, G., Sugathan, A., Mazor, T., Fraenkel, E., and Waxman, D.J. (2010). Unbiased, genome-wide in vivo mapping of transcriptional regulatory elements reveals sex differences in chromatin structure associated with sex-specific liver gene expression. Mol Cell Biol *30*, 5531-5544.
- Lodato, N.J., Melia, T., Rampersaud, A., and Waxman, D.J. (2017). Sex-Differential Responses of Tumor Promotion-Associated Genes and Dysregulation of Novel Long Noncoding RNAs in Constitutive Androstane Receptor-Activated Mouse Liver. Toxicological sciences : an official journal of the Society of Toxicology 159, 25-41.
- Lodato, N.J., Rampersaud, A., and Waxman, D.J. (2018). Impact of CAR Agonist Ligand TCPOBOP on Mouse Liver Chromatin Accessibility. Toxicological sciences : an official journal of the Society of Toxicology *164*, 115-128.
- Ma, W., Noble, W.S., and Bailey, T.L. (2014). Motif-based analysis of large nucleotide data sets using MEME-ChIP. Nat. Protocols 9, 1428-1450.
- Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics (Oxford, England) 27, 1696-1697.
- Macisaac, K.D., Gordon, D.B., Nekludova, L., Odom, D.T., Schreiber, J., Gifford, D.K., Young, R.A., and Fraenkel, E. (2006). A hypothesis-based approach for identifying the binding specificity of regulatory proteins from chromatin immunoprecipitation data. Bioinformatics (Oxford, England) 22, 423-429.
- MacLeod, J.N., Pampori, N.A., and Shapiro, B.H. (1991). Sex differences in the ultradian pattern of plasma growth hormone concentrations in mice. J Endocrinol *131*, 395-399.
- Maricque, B.B., Chaudhari, H.G., and Cohen, B.A. (2018). A massively parallel reporter assay dissects the influence of chromatin structure on cis-regulatory activity. Nature biotechnology.
- Masuyama, H., Hiramatsu, Y., Kunitomi, M., Kudo, T., and MacDonald, P.N. (2000). Endocrine disrupting chemicals, phthalic acid and nonylphenol, activate Pregnane X receptor-mediated transcription. Mol Endocrinol *14*, 421-428.
- Matthews, B.J., and Waxman, D.J. (2018). Computational prediction of CTCF/cohesinbased intra-TAD loops that insulate chromatin contacts and gene expression in mouse liver. eLife 7.

- Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., *et al.* (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic acids research 34, D108-110.
- Melia, T., and Waxman, D.J. (2019). Sex-Biased lncRNAs Inversely Correlate With Sex-Opposite Gene Coexpression Networks in Diversity Outbred Mouse Liver. Endocrinology 160, 989-1007.
- Mellor, C.L., Steinmetz, F.P., and Cronin, M.T.D. (2016). The identification of nuclear receptors associated with hepatic steatosis to develop and extend adverse outcome pathways. Critical reviews in toxicology *46*, 138-152.
- Misra, P., Viswakarma, N., and Reddy, J.K. (2013). Peroxisome proliferator-activated receptor-alpha signaling in hepatocarcinogenesis. Sub-cellular biochemistry *69*, 77-99.
- Mutoh, S., Sobhany, M., Moore, R., Perera, L., Pedersen, L., Sueyoshi, T., and Negishi, M. (2013). Phenobarbital indirectly activates the constitutive active androstane receptor (CAR) by inhibition of epidermal growth factor receptor signaling. Sci Signal 6, ra31.
- Nagy, C., and Turecki, G. (2012). Sensitive periods in epigenetics: bringing us closer to complex behavioral phenotypes. Epigenomics *4*, 445-457.
- Naugler, W.E., Sakurai, T., Kim, S., Maeda, S., Kim, K., Elsharkawy, A.M., and Karin, M. (2007). Gender disparity in liver cancer due to sex differences in MyD88dependent IL-6 production. Science (New York, N.Y.) 317, 121-124.
- Negishi, M. (2017). Phenobarbital Meets Phosphorylation of Nuclear Receptors. Drug Metab Dispos 45, 532-539.
- Ngan, C.H., Beglov, D., Rudnitskaya, A.N., Kozakov, D., Waxman, D.J., and Vajda, S. (2009). The structural basis of pregnane X receptor binding promiscuity. Biochemistry 48, 11572-11581.
- Nilsson, E.E., Sadler-Riggleman, I., and Skinner, M.K. (2018). Environmentally induced epigenetic transgenerational inheritance of disease. Environ Epigenet 4, dvy016.
- Niu, B., Coslo, D.M., Bataille, A.R., Albert, I., Pugh, B.F., and Omiecinski, C.J. (2018). In vivo genome-wide binding interactions of mouse and human constitutive androstane receptors reveal novel gene targets. Nucleic acids research 46, 8385-8403.

- O'Connor, C., and Waxman, D.J. (2006). Growth Hormone Regulation of Sex-Dependent Liver Gene Expression. Molecular Endocrinology 20, 2613-2629.
- Omiecinski, C.J., Coslo, D.M., Chen, T., Laurenzana, E.M., and Peffer, R.C. (2011). Multi-species analyses of direct activators of the constitutive androstane receptor. Toxicol Sci *123*, 550-562.
- Osabe, M., and Negishi, M. (2011). Active ERK1/2 protein interacts with the phosphorylated nuclear constitutive active/androstane receptor (CAR; NR1I3), repressing dephosphorylation and sequestering CAR in the cytoplasm. The Journal of biological chemistry 286, 35763-35769.
- Oshida, K., Vasani, N., Jones, C., Moore, T., Hester, S., Nesnow, S., Auerbach, S., Geter, D.R., Aleksunes, L.M., Thomas, R.S., *et al.* (2015). Identification of chemical modulators of the constitutive activated receptor (CAR) in a gene expression compendium. Nucl Recept Signal *13*, e002.
- Ou, J., Wolfe, S.A., Brodsky, M.H., and Zhu, L.J. (2018). motifStack for the analysis of transcription factor binding site evolution. Nature methods *15*, 8-9.
- Paranjpe, S., Bowen, W.C., Mars, W.M., Orr, A., Haynes, M.M., DeFrances, M.C., Liu, S., Tseng, G.C., Tsagianni, A., and Michalopoulos, G.K. (2016). Combined systemic elimination of MET and epidermal growth factor receptor signaling completely abolishes liver regeneration and leads to liver decompensation. Hepatology 64, 1711-1724.
- Parfett, C.L., and Desaulniers, D. (2017). A Tox21 Approach to Altered Epigenetic Landscapes: Assessing Epigenetic Toxicity Pathways Leading to Altered Gene Expression and Oncogenic Transformation In Vitro. International journal of molecular sciences 18.
- Phillips, J.M., Burgoon, L.D., and Goodman, J.I. (2009). The constitutive active/androstane receptor facilitates unique phenobarbital-induced expression changes of genes involved in key pathways in precancerous liver and liver tumors. Toxicological sciences : an official journal of the Society of Toxicology 110, 319-333.
- Poland, A., Mak, I., Glover, E., Boatman, R.J., Ebetino, F.H., and Kende, A.S. (1980). 1,4-Bis[2-(3,5-dichloropyridyloxy)]benzene, a potent phenobarbital-like inducer of microsomal monooxygenase activity. Mol Pharmacol 18, 571-580.
- Poland A, M.I., Glover E, Boatman RJ, Ebetino FH, Kende AS (1980). 1,4-Bis[2-(3,5-Dichloropyridyloxy)]Benzene, a Potent Phenobarbital-Like Inducer of Microsomal Monooxygenase Activity. Mol Pharmacol 18, 571-580.

- Ren, H., Aleksunes, L.M., Wood, C., Vallanat, B., George, M.H., Klaassen, C.D., and Corton, J.C. (2010). Characterization of peroxisome proliferator-activated receptor alpha--independent effects of PPARalpha activators in the rodent liver: di-(2-ethylhexyl) phthalate also activates the constitutive-activated receptor. Toxicol Sci 113, 45-59.
- Reveron-Gomez, N., Gonzalez-Aguilera, C., Stewart-Morgan, K.R., Petryk, N., Flury, V., Graziano, S., Johansen, J.V., Jakobsen, J.S., Alabert, C., and Groth, A. (2018).
 Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. Mol Cell 72, 239-249 e235.
- Rooney, J.P., Oshida, K., Kumar, R., Baldwin, W.S., and Corton, J.C. (2018). Chemical Activation of the Constitutive Androstane Receptor (CAR) Leads to Activation of Oxidant-Induced Nrf2. Toxicol Sci.
- Ross, P.K., Woods, C.G., Bradford, B.U., Kosyk, O., Gatti, D.M., Cunningham, M.L., and Rusyn, I. (2009). Time-course comparison of xenobiotic activators of CAR and PPARalpha in mouse liver. Toxicol Appl Pharmacol 235, 199-207.
- Ruggieri, A., Barbati, C., and Malorni, W. (2010). Cellular and molecular mechanisms involved in hepatocellular carcinoma gender disparity. International journal of cancer 127, 499-504.
- Ruiz-Hernandez, A., Kuo, C.C., Rentero-Garrido, P., Tang, W.Y., Redon, J., Ordovas, J.M., Navas-Acien, A., and Tellez-Plaza, M. (2015). Environmental chemicals and DNA methylation in adults: a systematic review of the epidemiologic evidence. Clin Epigenetics 7, 55.
- Scacchi, M., Pincelli, A.I., and Cavagnini, F. (1999). Growth hormone in obesity. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity 23, 260-271.
- Schwartz, J.B. (2007). The current state of knowledge on age, sex, and their interactions on clinical pharmacology. Clinical pharmacology and therapeutics *82*, 87-96.
- Sexton, B.S., Avey, D., Druliner, B.R., Fincher, J.a., Vera, D.L., Grau, D.J., Borowsky, M.L., Gupta, S., Girimurugan, S.B., Chicken, E., *et al.* (2014). The spring-loaded genome: Nucleosome redistributions are widespread, transient, and DNA-directed. Genome research, 251-259.
- Shao, Z., Zhang, Y., Yuan, G.C., Orkin, S.H., and Waxman, D.J. (2012). MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. Genome Biol 13, R16.

- Shen, L., Shao, N.-Y., Liu, X., Maze, I., Feng, J., and Nestler, E.J. (2013). diffReps: detecting differential chromatin modification sites from ChIP-seq data with biological replicates. PLoS One 8, e65598.
- Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet *15*, 272-286.
- Shu, W., Chen, H., Bo, X., and Wang, S. (2011). Genome-wide analysis of the relationships between DNaseI HS, histone modifications and gene expression reveals distinct modes of chromatin domains. Nucleic Acids Res 39, 7428-7443.
- Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Scholer, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., *et al.* (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 480, 490-495.
- Strazzullo, M., and Matarazzo, M.R. (2017). Epigenetic Effects of Environmental Chemicals on Reproductive Biology. Curr Drug Targets *18*, 1116-1124.
- Sugathan, A., and Waxman, D.J. (2013). Genome-wide analysis of chromatin states reveals distinct mechanisms of sex-dependent gene regulation in male and female mouse liver. Molecular and cellular biology *33*, 3594-3610.
- Suino, K., Peng, L., Reynolds, R., Li, Y., Cha, J.-Y., Repa, J.J., Kliewer, S.A., and Xu, H.E. (2004). The nuclear xenobiotic receptor CAR: structural determinants of constitutive activation and heterodimerization. Molecular cell 16, 893-905.
- Suvorov, A., and Waxman, D.J. (2015). Early programing of uterine tissue by bisphenol A: Critical evaluation of evidence from animal exposure studies. Reproductive toxicology (Elmsford, N.Y.) 57, 59-72.
- Swinstead, E.E., Paakinaho, V., Presman, D.M., and Hager, G.L. (2016). Pioneer factors and ATP-dependent chromatin remodeling factors interact dynamically: A new perspective: Multiple transcription factors can effect chromatin pioneer functions through dynamic interactions with ATP-dependent chromatin remodeling factors. BioEssays : news and reviews in molecular, cellular and developmental biology 38, 1150-1157.
- Tabb, M.M., Kholodovych, V., Grun, F., Zhou, C., Welsh, W.J., and Blumberg, B. (2004). Highly chlorinated PCBs inhibit the human xenobiotic response mediated by the steroid and xenobiotic receptor (SXR). Environ Health Perspect *112*, 163-169.
- Thomson, J.P., Hunter, J.M., Lempiainen, H., Muller, A., Terranova, R., Moggs, J.G., and Meehan, R.R. (2013a). Dynamic changes in 5-hydroxymethylation signatures

underpin early and late events in drug exposed liver. Nucleic acids research 41, 5639-5654.

- Thomson, J.P., Lempiainen, H., Hackett, J.A., Nestor, C.E., Muller, A., Bolognani, F., Oakeley, E.J., Schubeler, D., Terranova, R., Reinhardt, D., *et al.* (2012). Nongenotoxic carcinogen exposure induces defined changes in the 5hydroxymethylome. Genome Biol *13*, R93.
- Thomson, J.P., Moggs, J.G., Wolf, C.R., and Meehan, R.R. (2013b). Epigenetic profiles as defined signatures of xenobiotic exposure. Mutat Res.
- Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., *et al.* (2012). The accessible chromatin landscape of the human genome. Nature 489, 75-82.
- Tian, J., Huang, H., Hoffman, B., Liebermann, D.A., Ledda-Columbano, G.M., Columbano, A., and Locker, J. (2011). Gadd45beta is an inducible coactivator of transcription that facilitates rapid liver growth in mice. The Journal of clinical investigation 121, 4491-4502.
- Tian, J., Marino, R., Johnson, C., and Locker, J. (2018). Binding of Drug-Activated CAR/Nr1i3 Alters Metabolic Regulation in the Liver. iScience *9*, 209-228.
- Timsit, Y.E., and Negishi, M. (2007). CAR and PXR: the xenobiotic-sensing receptors. Steroids 72, 231-246.
- Timsit, Y.E., and Negishi, M. (2014). Coordinated regulation of nuclear receptor CAR by CCRP/DNAJC7, HSP70 and the ubiquitin-proteasome system. PLoS One 9, e96092.
- Vijayakumar, A., Yakar, S., and Leroith, D. (2011). The intricate role of growth hormone in metabolism. Frontiers in endocrinology 2, 32.
- Wada, T., Gao, J., and Xie, W. (2009). PXR and CAR in energy metabolism. Trends in endocrinology and metabolism: TEM 20, 273-279.
- Wagner, C.E., Jurutka, P.W., Marshall, P.A., and Heck, M.C. (2017). Retinoid X Receptor Selective Agonists and their Synthetic Methods. Current topics in medicinal chemistry 17, 742-767.
- Walker, C.L. (2016). Minireview: Epigenomic Plasticity and Vulnerability to EDC Exposures. Mol Endocrinol *30*, 848-855.
- Wallace, B.D., and Redinbo, M.R. (2013). Xenobiotic-sensing nuclear receptors involved in drug metabolism: a structural perspective. Drug Metab Rev 45, 79-100.

- Wang, X., He, L., Goggin, S.M., Saadat, A., Wang, L., Sinnott-Armstrong, N., Claussnitzer, M., and Kellis, M. (2018). High-resolution genome-wide functional dissection of transcriptional regulatory regions and nucleotides in human. Nat Commun 9, 5380.
- Wang, Y.M., Ong, S.S., Chai, S.C., and Chen, T. (2012). Role of CAR and PXR in xenobiotic sensing and metabolism. Expert Opin Drug Metab Toxicol 8, 803-817.
- Wauthier, V., Sugathan, A., Meyer, R.D., Dombkowski, A.A., and Waxman, D.J. (2010). Intrinsic sex differences in the early growth hormone responsiveness of sexspecific genes in mouse liver. Mol Endocrinol 24, 667-678.
- Wauthier, V., and Waxman, D.J. (2008). Sex-specific early growth hormone response genes in rat liver. Mol Endocrinol 22, 1962-1974.
- Waxman, D.J. (1999). P450 gene induction by structurally diverse xenochemicals: central role of nuclear receptors CAR, PXR, and PPAR. Arch Biochem Biophys 369, 11-23.
- Waxman, D.J., and Holloway, M.G. (2009). Sex differences in the expression of hepatic drug metabolizing enzymes. Mol Pharmacol *76*, 215-228.
- Wiwi, C.A., and Waxman, D.J. (2004). Role of hepatocyte nuclear factors in growth hormone-regulated, sexually dimorphic expression of liver cytochromes P450. Growth factors (Chur, Switzerland) 22, 79-88.
- Wu, B., Li, S., and Dong, D. (2013). 3D structures and ligand specificities of nuclear xenobiotic receptors CAR, PXR and VDR. Drug Discov Today 18, 574-581.
- Wu, M.-H., Ma, W.-L., Hsu, C.-L., Chen, Y.-L., Ou, J.-H.J., Ryan, C.K., Hung, Y.-C., Yeh, S., and Chang, C. (2010). Androgen receptor promotes hepatitis B virusinduced hepatocarcinogenesis through modulation of hepatitis B virus RNA transcription. Science translational medicine 2, 32ra35.
- Xie, S., Duan, J., Li, B., Zhou, P., and Hon, G.C. (2017). Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. Mol Cell 66, 285-299 e285.
- Xin, F., Susiarjo, M., and Bartolomei, M.S. (2015). Multigenerational and transgenerational effects of endocrine disrupting chemicals: A role for altered epigenetic regulation? Semin Cell Dev Biol 43, 66-75.
- Xu, S., Grullon, S., Ge, K., and Peng, W. (2014). Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. Methods in molecular biology (Clifton, N.J.) 1150, 97-111.

- Yamamoto, Y., Moore, R., Goldsworthy, T.L., Negishi, M., and Maronpot, R.R. (2004). The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice. Cancer research 64, 7197-7200.
- Yang, H., and Wang, H. (2014). Signaling control of the constitutive androstane receptor (CAR). Protein & cell 5, 113-123.
- Yang, H., Zheng, W., Shuai, X., Chang, R.M., Yu, L., Fang, F., and Yang, L.Y. (2015). MicroRNA-424 inhibits Akt3/E2F3 axis and tumor growth in hepatocellular carcinoma. Oncotarget 6, 27736-27750.
- Yu, M., Tang, Z., Meng, F., Tai, M., Zhang, J., Wang, R., Liu, C., and Wu, Q. (2016). Elevated expression of FoxM1 promotes the tumor cell proliferation in hepatocellular carcinoma. Tumour Biol 37, 1289-1297.
- Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome research 21, 1273-1283.
- Zhang, Y., Klein, K., Sugathan, A., Nassery, N., Dombkowski, A., Zanger, U.M., and Waxman, D.J. (2011). Transcriptional profiling of human liver identifies sexbiased genes associated with polygenic dyslipidemia and coronary artery disease. PLoS One 6, e23506.
- Zhang, Y., Laz, E.V., and Waxman, D.J. (2012). Dynamic, sex-differential STAT5 and BCL6 binding to sex-biased, growth hormone-regulated genes in adult mouse liver. Mol Cell Biol 32, 880-896.

CURRICULUM VITAE















