Abstract

Travis C. Book: Interstrain Differences in Transcription Factor Activity in Mouse Liver

(Under the direction of Ivan Rusyn, M.D., Ph.D.)

Transcription factors (TFs) are responsible for regulating gene expression. By simultaneously measuring the activity of multiple transcription factors one should be able to create a snapshot which functionally characterizes the TF regulatory network. As the primary site for metabolism of nutrients, understanding the regulatory mechanisms of the liver could provide great insight for pharmacological and toxicological studies. This study applied a method which allows simultaneous measurement of multiple transcription factors to an in-vivo design consisting of 20 inbred mouse strains. Our hypothesis was that, among inbred strains, regarding TF activity, interstrain variation would be high – that is, noticeably different among the strains. In addition, we also investigated the extent of intrastrain correlation for each strain and the utility of the method to return data which could be used in the discovery of yet described, cooperative TF-TF and TF-gene interactions. In order to accomplish this, we measured the TF activity of each strain, investigated mRNA correlations with the transcription factor data, and created transcription factor activity profiles of each strain to illuminate new directions of research while improving knowledge on basal TF activities.

The major finding of our work shows that there are significant interstrain differences in TF activity. In addition, we show that there is high intrastrain correlation for each strain considered. Using the TF data from each strain we were able to construct TF activity profiles for each mouse strain as well as determine highly correlated TF-TF interactions which represent both well known and currently not characterized relationships. Finally, we confirmed several TF-gene interactions by correlating TF activity with microarray mRNA levels.
Acknowledgements

This work would not have been possible without the collaboration of the following individuals and institutions. I would like to thank:

Dr. Dan Gatti\(^1\) for invaluable direction and aid regarding data analysis as well as tissue collection,
Dr. Masato Tsuchiya\(^2\) for sacrificing the mice used in the study and tissue collection,
UNC Department of Laboratory Animal Facility for housing and feeding the mice used in the study,
Dr. Matthew Martin\(^3\) and Dr. David Dix\(^3\) for project advice and direction regarding practicum work,
Dr. Pierre Bushel\(^4\) for data mapping the mouse TFs to the genes represented on the whole mouse Agilent array

---

\(^1\) Jackson Laboratory, 600 Main St. Bar Harbor Maine 04609
\(^2\) Visiting Scientist from University of Yamanashi, Japan to the Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599
\(^3\) National Center for Computational Toxicology, US EPA, 109 T.W. Alexander Dr. Durham NC 27711
\(^4\) NIEHS, Research Triangle Park, NC 27709, P.O. Box 12233, Mail Drop E1-02
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vi</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter 1: Introduction</td>
<td>10</td>
</tr>
<tr>
<td>Chapter 2: Literature Review</td>
<td>12</td>
</tr>
<tr>
<td>Chapter 3: Experimental Methods</td>
<td>33</td>
</tr>
<tr>
<td>Chapter 4: Results and Discussion</td>
<td>36</td>
</tr>
<tr>
<td>Chapter 5: Conclusions</td>
<td>46</td>
</tr>
<tr>
<td>Chapter 6: Future Directions</td>
<td>47</td>
</tr>
<tr>
<td>Chapter 7: Practicum</td>
<td>49</td>
</tr>
<tr>
<td>References</td>
<td>84</td>
</tr>
</tbody>
</table>
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td>Mice Removed from analysis with removal criteria</td>
<td>63</td>
</tr>
<tr>
<td>2:</td>
<td>Significant RTU-Gene pairs using Pearson Correlation</td>
<td>64</td>
</tr>
<tr>
<td>3:</td>
<td>Significant RTU-Gene pairs using Spearman Rank Correlation</td>
<td>64</td>
</tr>
<tr>
<td>4:</td>
<td>RTU/TF Association Indicated by TRU-RTU Correlation Represented in Current Literature</td>
<td>66</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1:</td>
<td>Strain Correlation Heatmap</td>
<td>52</td>
</tr>
<tr>
<td>2:</td>
<td>Barplot showing Intrastrain Correlation</td>
<td>53</td>
</tr>
<tr>
<td>3:</td>
<td>RTU Profiles of Inbred Mouse Strains</td>
<td>54</td>
</tr>
<tr>
<td>4:</td>
<td>Organized Heatmap with Cluster Dendrogram and Histogram showing RTU-RTU correlation</td>
<td>62</td>
</tr>
<tr>
<td>5:</td>
<td>Cluster Dendrogram of RTU Library</td>
<td>63</td>
</tr>
</tbody>
</table>
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AhR</td>
<td>Aryl-hydrocarbon Receptor</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen Receptor</td>
</tr>
<tr>
<td>BSEP</td>
<td>Bile Salt Export Pump</td>
</tr>
<tr>
<td>cAMP</td>
<td>cyclic Adenosine Monophosphate</td>
</tr>
<tr>
<td>CAR</td>
<td>Constitutive Androstane Receptor</td>
</tr>
<tr>
<td>CDCA</td>
<td>Chenodeoxycholic acid</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>C/EBPβ</td>
<td>CCAAT-enhancer binding protein β</td>
</tr>
<tr>
<td>Cox-2</td>
<td>Cyclooxygenase 2</td>
</tr>
<tr>
<td>CREB</td>
<td>Cyclic AMP Response Element Binding Protein</td>
</tr>
<tr>
<td>CYP</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>EF</td>
<td>Edema Factor</td>
</tr>
<tr>
<td>ELIDA</td>
<td>Enzyme Linked protein/DNA Interaction Assay</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme Linked Immunosorbent Assay</td>
</tr>
<tr>
<td>EMMA</td>
<td>Efficient Mixed Model Association</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electrophoretic Mobility Shift Assay</td>
</tr>
<tr>
<td>ER</td>
<td>Estrogen Receptor</td>
</tr>
<tr>
<td>FasL</td>
<td>Fas Ligand</td>
</tr>
<tr>
<td>FDR</td>
<td>False Discovery Rate</td>
</tr>
<tr>
<td>FOXa2</td>
<td>Forkhead Box Protein a2</td>
</tr>
<tr>
<td>FOXa3</td>
<td>Forkhead Box Protein a3</td>
</tr>
<tr>
<td>FOX-M-1-B</td>
<td>Forkhead Box protein M1</td>
</tr>
<tr>
<td>FXR</td>
<td>Farnesoid X Receptor</td>
</tr>
<tr>
<td>GSH</td>
<td>Glutathione</td>
</tr>
<tr>
<td>HGF</td>
<td>Hepatocyte Growth Factor</td>
</tr>
<tr>
<td>HNF1</td>
<td>Hepatocyte Nuclear Factor 1</td>
</tr>
<tr>
<td>HNF4</td>
<td>Hepatocyte Nuclear Factor 4</td>
</tr>
<tr>
<td>IDO1</td>
<td>Indoleamine 2,3-Dioxygenase-like Protein 1</td>
</tr>
<tr>
<td>IDO2</td>
<td>Indoleamine 2,3-Dioxygenase-like Protein 2</td>
</tr>
<tr>
<td>IkB</td>
<td>Inhibitor of kB</td>
</tr>
<tr>
<td>IL-1β</td>
<td>Interleukin 1β</td>
</tr>
<tr>
<td>IL-6</td>
<td>Interleukin-6</td>
</tr>
<tr>
<td>Keap1</td>
<td>Kelch-like ECH associated Protein 1</td>
</tr>
<tr>
<td>LXR</td>
<td>Liver X receptor</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen Activated Protein Kinase</td>
</tr>
<tr>
<td>Mdm2</td>
<td>Murine Double Minute 2</td>
</tr>
<tr>
<td>Mdr2</td>
<td>Multidrug Resistance Protein 2</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger Ribonucleic Acid</td>
</tr>
<tr>
<td>MRTU</td>
<td>Multiplexing Reporter Transcription Unit</td>
</tr>
<tr>
<td>MTBE</td>
<td>Methyl Tert-Butyl Ether</td>
</tr>
<tr>
<td>Ngn3</td>
<td>Neurogenin 3</td>
</tr>
<tr>
<td>NR</td>
<td>Nuclear Receptor</td>
</tr>
</tbody>
</table>
NRC – National Research Council
NRF2 – Nuclear Factor (erythroid-derived 2)-like 2
PCR – Polymerase Chain Reaction
PI3K – Phosphatidylinositol 3-kinase
Plau – Urokinase-type Plasminogen Activator
PPARα – Peroxisome Proliferator Activated Receptor α
PPARγ – Peroxisome Proliferator Activated Receptor γ
QTL – Qualitative Trait Loci
RAR – Retinoic acid receptor-α
RNA – Ribonucleic Acid
RNS – Reactive Nitrogen Species
ROS – Reactive Oxygen Species
RTU – Reporter Transcription Unit
RT-PCR – Real Time Quantitative Reverse Transcription PCR
RXR – Retinoid X Receptor
Sdc-1 – Syndecan-1
SLC39A6 – Solute Carrier Family 39, (Zinc Transporter) member 6
αSM ACT – α-Smooth Muscle Actin
SNP – Single Nucleotide Polymorphism
Stat3 – Signal Transducer and Activator of Transcription 3
TCDD – 2,3,7,8-Tetrachlorodibenzo-p-dioxin
TF – Transcription Factor
TGFα – Tumor Growth Factor α
TNFα – Tumor Necrosis Factor α
UBASH3B – Ubiquitin-Associated and SH3 Domain-Containing Protein B
uPAR – Urokinase Plasminogen Activator Receptor
VEGF – Vascular Endothelial Growth Factor

RTU Abbreviations

AhRE – Aryl Hydrocarbon Receptor Element
AP-1 – Activating Protein 1
AP-2 – Activating Protein 2
ARE – Antioxidant Response Element
BRE – B recognition Element
C/EBP – CCAAT-enhancer binding protein β
CMV – Cytomegalovirus
CRE – Cyclic AMP Response Element
DR5 – Direct Repeat 5 / Retinoic Acid Receptor Element
E-Box – Enhancer Box
EGR – Early Growth Response Factor
ERE – Estrogen Response Element
Ets – E-twenty six
E2F – [Transcription Factor family name; not an abbreviation]
FoxA – Forkhead Box Protein Family A
FoxO – Forkhead Box Protein Family O
FXRE – Farnesoid X-Receptor-Responsive Element
GATA – ‘GATA’ Binding Transcription Factors
GLI – Glioblastoma
GRE – Glucocorticoid Receptor
HIF1α – Hypoxia-Inducible Factor 1 α
HNF6 – Hepatocyte Nuclear Factor 6
HSE – Heat Shock Element
IR1 – Inverted Repeat Element 1
ISRE – IFN Stimulation Response Element
LXRE (DR4/LXR) – Liver X Receptor Response Element
MRE – Metal Response Element
MT-Delta – Metallothionein Delta
Myb – [Name; not an abbreviation]
Myc – [Name; not an abbreviation]
M-06 – [Name; not an abbreviation]
M-18 – [Name; not an abbreviation]
M-19 – [Name; not an abbreviation]
M61 – [Name; not an abbreviation]
NF-Kb – Nuclear Factor kappa-light-chain-enhancer of activated B cells
NF1 – Nuclear Factor 1
NRF1 – Nuclear Respiratory Factor 1
Oct – Octamer-Binding Transcription Factor
Pax – Paired Box
PBREM – Phenobarbital Responsive Enhancer Module
PPRE – Peroxisome Proliferator Hormone Response Element
PXRE – Pregnan X Response Element
PS3 – Tumor Protein S3
RARE – Retinoic Acid Response Element
RORE – Retinoid-related Orphan Receptor Response Elements
SOX – Sry-related High-Mobility Group Box
SP1 – Specificity Protein 1
SREPB – Sterol Regulatory Element Binding Protein
STAT – Signal Transducers and Activators of Transcription
TA – [Nucleotide Repeat]
TAL – Transcription Activator-like Effector
TCF/β-catenin – Transcription Factor β-catenin
TGFβ – Transforming Growth Factor β
VDRE – Vitamin D Response Element
Xbp1 – X-box binding Protein 1
CHAPTER 1

Introduction

There are an estimated 3000 TFs coded in the human genome (8% of the organisms total encoded proteins) each with the ability to activate or suppress transcription of a particular gene(s). Because of their ability to regulate transcription based on cellular stimuli (e.g. xenobiotic exposure), understanding TF activity is a subject of considerable interest. [1] As such, many TFs have been well characterized by their functions and interactions as well as their roles in organized regulatory networks. These networks are important to understanding how an organism regulates function and copes with exposure. Because of the ability of TFs to respond to xenobiotics, the liver is a major organ of interest: it is the primary site for metabolism, responsible for both xenobiotic and nutrient metabolism due to its abundance of phase I/II metabolic enzymes.[2]

Prior work has shown that regulatory interactions are conserved to varying degrees in closely related organisms.[1, 3] However, TF activity has not yet been measured across multiple individuals of inbred mice strains. Given their genetic similarity, individuals from a single inbred strain were expected to show highly similar TF activity levels. In addition, variation among strains was expected to be significant. With the methodology devised by Romanov et al it is now feasible to accurately measure activity levels of multiple TFs.[4] Application of the plasmid reporter system will allow determination of the inherent interstrain variability of basal TF activity across multiple, simultaneously measured TFs in the mouse liver using a panel of inbred strains. Because the measurements are made in-vivo they can account for more complex TF-TF interactions which may not be observable in cultured cell lines as whole organism interactions are much more complex than isolated cells and/or tissue. Elucidating these activities will provide a much clearer picture of the complex regulatory and responsive interactions of TFs.

Our hypothesis was that, among inbred strains, regarding TF activity, interstrain variation would be high - that is, noticeably different among the strains. In addition, we also investigated the extent of
intrastrain correlation for each strain and whether it was feasible to construct differentiable TF activity profiles for each strain. Further, we were interested in the utility of the method to return data which could be used in the discovery of yet described, cooperative TF-TF and TF-gene interactions.

We applied the high throughput plasmid reporter system developed by Romanov et al [4] to 20 strains of inbred mice to procure in-vivo TF activity levels. Here, we show that said is functionally robust when applied to in vivo mouse experiments. Our results showed that intrastrain TF activity variation is fairly low while there is significant interstrain variation of TF activity across the mouse panel. From the data, we were able to build TF activity profiles for each strain as well as correlate TF activity to gene expression. Our results are well representative of known TF-TF relationships while also suggesting novel TF-TF interactions. Further, such information will contribute to intelligent strain selection and provide insight regarding the translation of results across strain and species studies.
CHAPTER 2

Literature Review

Transcription factors (TF) are sequence specific DNA binding proteins responsible for regulating gene transcription[5], thereby denoting responsibility for maintaining homeostasis and stimulus response. In order to regulate the vast quantity of functions and genes, TFs have multiple modes of operation and can function independently or within a complex of other TFs through either promoting or preventing transcription of the corresponding gene.[6]

Major hepatic functions regulated by selected TF activity include cell proliferation, oxidative stress, inflammation, and metabolism. Without the precise regulation and mediated response to relevant stimuli provided by TFs such functions could not operate as specifically or accurately as they do. These pathways are discussed below regarding their molecular function in liver tissue and the importance of TFs in their performance.

Cell Proliferation

Cell proliferation is the increase of the number of cells within the context of tissue growth. As the primary site for xenobiotic metabolism, the liver is constantly exposed to potentially harmful substances which can cause cellular damage and/or death. Therefore, it is necessary for the cells of said tissue to be equipped with regenerative capacity. Hepatic tissue has a particularly high ability to grow and regenerate, both naturally and in response to physical or chemical damage. Mediating proliferation requires control over the cell cycle which is not only mediated by organ-specific TFs but also is dependent on the stage of development.[7]

During embryonic and postnatal growth, hepatocyte differentiation is regulated by HNF1α, NHF4α, Foxa2, and Foxa3. These master regulatory factors maintain a presence in differentiated cells, acting as gatekeepers of proliferation and the cell cycle. In mature liver, cell cycle progression is mediated
by a network of TFs regulated by FOX-M-1-B.[7] In mice, a FOX-M-1-B knockout is embryonic lethal due to diminished DNA replication and inability to enter mitosis (a 75% reduction in hepatoblasts was observed prior to death).[8] Following injury, Tumor Growth Factor α (TGFα), Tumor Necrosis Factor α (TNFα), interleukin-6 (IL-6), and hepatocyte growth factor (HGF) induce proliferation via reentry of terminally differentiated hepatocytes into the cell cycle.[9-11] Without IL-6, immediate induction of the necessary TFs (NF-kB, AP-1(c-Fos/c-Jun), Stat3, c-Myc, and the C/EBPβ family) fails, resulting in a decrease in proliferation.[12, 13] A decrease in the functional mass of the liver due to a toxic insult would decrease the metabolic rate of the liver thereby compromising its ability to function.

**Oxidative Stress**

Oxidative stress is an imbalance between pro-oxidant production and the resulting damage and the ability of the organism to detoxify the pro-oxidants and repair any damage. Pro-oxidants are commonly referred to as reactive oxygen species (ROS) and can originate from both external to and internal to an organism. Primary Endogenous sources of ROS include endoplasmic reticulum (flavoproteins, CYP enzymes), lysosomes (myeloperoxidase), peroxisomes (oxidases, flavoproteins), and mitochondria (electron chain protein complexes). Mitochondria are the single largest source of endogenous ROS. As the primary site of metabolism the liver requires considerable energy and thus high number of mitochondria. Other sources of ROS include xenobiotic metabolism, injury, and external causes such as heat.[14-16]

Hepatic tissue can tolerate mild oxidative stress due to sufficient antioxidant presence. Damage occurs when the oxidant-antioxidant imbalance becomes too severe for the antioxidant defense mechanisms to cope.[14] ROS can cause cellular damage in two ways: physical damage to the cell and altered cell signaling. Physically, ROS can damage lipids, protein, and/or DNA via oxidation. Pathways impacted by ROS damage include those responsible for gene expression, cell adhesion, metabolism, and
cell cycle regulation. In addition to cell damage, ROS can affect DNA directly and cause the formation of DNA adducts. These adducts can result in replication errors and detrimental mutations which will interfere with proper cellular functioning.

In the liver, ROS can also activate inflammatory pathways mediated by TFs and kinases such as STAT-3, NF-kB, AP-1, MAPK, and PI3K, as well as inhibit the tumor-suppressor gene P53.\[16, 17\] P53 acts as a proofreader for errors in the genetic code; pausing the cell cycle until necessary repairs are completed. If unmitigated by DNA repair mechanisms, ROS will cause permanent DNA damage potentially resulting in cell death or tumorigenesis. Loss of P53 function will lead to abnormalities in the cell cycle and likely tumor development (see section on P53) while chronic inflammation will result in extensive fibrosis due to tissue destruction and subsequent repair. While ROS expose is unavoidable, mitigating any damage is important to cellular functioning because ROS damage can not only be carcinogenic but disrupt the functioning of undamaged tissues by interfering with cell-signaling.

**Inflammation**

Inflammation is an innate immune response which demonstrates rapid onset and is characterized by the presence of heat, pain, swelling, redness, and the loss of function. It is an integral aspect of wound healing and is beneficial in the short term by promoting tissue regeneration and the migration of leukocytes (primarily neutrophils) and plasma to the inflamed area. When inflammation persists over a extended time period it becomes known as chronic inflammation, which is characterized by simultaneous repair and destruction of the inflamed tissue and presents with delayed onset, prolonged duration, and the presence of mononuclear cells and fibroblasts as the primary cell types.

When inflammation occurs in the liver, it is called hepatitis. Chronic hepatitis is strongly linked to the development of fibrosis, cirrhosis, and hepatocellular carcinoma.\[18, 19\] In the liver, inflammation is mediated by a cytokine response which can be triggered by physical injury, viral and bacterial infections,
toxins, and autoimmune diseases. NF-kB and AP-1 are ubiquitous first responder TFs responsible for initiating and maintaining inflammation through their pro-inflammatory cytokine production pathways.[20] As the cause of the inflammation is assuaged by the appropriate process, other TFs will antagonize NF-kB and AP-1 (as well as other pro-inflammatory TFs). PPARα, PPARγ, and LXR negatively regulate (inhibit) the effects of NF-kB, AP-1, and STAT resulting in an anti-inflammatory response.[21, 22] Mice lacking the ability to inhibit inflammation – such as LXR knockouts – show increased inflammatory damage such as release of hepatic enzymes and fibrotic tissue deposition.[23, 24] Inflammation as an innate reaction is vital to the defense and recovery of liver tissue but if allowed to progress unchecked the inflammation will ultimately destroy the tissue.

**Metabolism (Nuclear Receptors)**

Nuclear receptors (NRs) are a class of proteins which act as sensors of ligands for hormones, lipids, vitamins, xenobiotics, bile acids, and metabolic intermediates. Due to their ability to bind DNA directly, NRs are classified as TFs; however, NRs are unable to bind DNA until they undergo a conformational change induced through ligand binding. NRs are classified as either type 1 or type 2 depending on intracellular location (nucleus or cytoplasm) and NRs with unknown ligands are known as “orphan” NRs. Type 1 NRs undergo a conformational change after binding their ligands in the cytoplasm which allows them to translocate to the nucleus and bind DNA. Type 2 NRs are restrained to the nucleus and bind DNA as heterodimers, usually with Retinoid X receptor (RXR).

NRs regulate gene expression related to development (embryonic and post-natal), homeostasis, and metabolism[25], thereby signaling genetic changes based on the environmental needs of the cell.[26] There are multiple families of NR categorized by sequence homology. Among others, the PPAR family, RAR, LXR, HNF4, RXR, ER, AR, are all considered NRs.
Specific to the liver, NRs regulate key hepatic functions such as bile secretion and homeostasis, lipid, glucose, and drug metabolism as well as tumor formation, fibrosis, cell differentiation, regeneration, and inflammation.[23, 27, 28] Given such a broad list of regulatory responsibilities, NRs are particularly important to proper functional maintenance. Without their proper functioning, the liver would be less able to regulate itself in relation to the extracellular environment and therefore the organ would be significantly more susceptible to damage and disease.[26] For example, NR function determines the risk for gallstone formation as the NR FXR induces bile acid (BSEP) and phospholipid (Mdr2) export pumps. If these pumps malfunction and bile components aggregate, gallstones may form. Because NRs have multiple functions, the impact of one failure can have a domino effect. Faulty export of bile acids may lead to gallstones but because bile acids also regulate hepatic lipid and glucose homeostasis through NRs a decrease in exported bile acids will reverberate through the organism.[25]

In response to toxins and potential toxins, the PPAR family, FXR, PXR, and LXR have the ability to act as sensors and initiate transcription of the necessary cytoprotective genes. FXR, for example, in the liver, is activated by Chenodeoxycholic acid (CDCA) which causes the induction of antioxidant and xenobiotic metabolizing enzymes through activation of C/EBPβ. However, an FXR knockout attenuated the ability of CDCA to activate C/EBPβ and therefore prevented the induction of cytoprotective enzymes.[29] Further, FXR, LXR, and the PPAR family, also have significant anti-inflammatory effects which may be relevant to liver pathogenesis and treatment.[25] Without properly functioning NRs cells would not be able to readily adapt to a changing environment and would likely not survive.

Response of transcription factors to xenobiotic stress in liver tissue

In addition to the roles played in normal homeostasis, TFs are responsible for mediating stress responses throughout the body. These responses are particularly important as they represent the damage
mitigating efforts of the organism. Stress can be a result of physical, chemical, and radiological (e.g. sunlight) damage therefore the responses mediated by TFs are incredibly diverse. Depending on the particular stressor, a response can be very general, very specific, or require multiple steps of different complexities and specificities.

As this work focuses on liver tissue, the TFs considered here are particularly relevant to xenobiotic stress in the liver and were chosen to demonstrate functional diversity of such responses. This does not preclude, however, the possibility that the following TFs have functions beyond what is considered here. The eight TFs to be further considered are the PPAR family, CAR, AhR, NFKB, HNF6, AP-1, P53, and NRF2.

**PPARs – Peroxisome Proliferator Activated Receptors**

The Peroxisome Proliferator-Activated Receptors (PPARs) are a family (α,β/δ,y) of NRs (see above section) most highly expressed in brown adipose and liver tissue which act as lipid sensors to regulate gene expression. Prior to binding DNA, PPARs must first heterodimerize with RXR: PPARs not bound to RXR are not able to bind DNA. The sequences where PPARs bind are known as peroxisome proliferator hormone response elements (PPREs) and are composed of two AGGTCA sequences separated by one nucleotide. Due to the importance of lipids to multiple functions, PPARs are associated with pleiotrophic responses, specifically playing major roles in metabolism, cellular differentiation, development, and tumorigenesis, and relate to chronic disorders such as diabetes, inflammation, and fatty acid liver disease.[30]

PPARs role in the liver primarily regards lipid homeostasis. The liver performs several functions in lipid metabolism, specifically lipogenesis and lipoprotein synthesis, both of which can be regulated by PPARs. In response to normal variation in fatty acid concentration, PPARs function as lipid sensors with the ability to align local gene expression to metabolic rate, thereby allowing the organism to regulate its metabolism according to resource availability. PPARα and PPARγ control lipoprotein fatty acid release by
regulating the expression of lipoprotein lipase (LPL). Should intracellular fatty acid levels increase dramatically (such as seen in type II diabetes), PPARα and PPARγ will upregulate carnitine palmitoyltransferases to increase the amount of activated fatty acids being imported to the mitochondria.[31]

PPAR activity can be diminished or increased in response to toxic stimuli and both result in damage to the organism. Xenobiotics which induce PPAR activity in rats and mice can lead to adverse health effects such as enlargement of the liver via proliferation of hepatocyte peroxisomes and induction of the fatty acid oxidation enzyme CYP4A.[32] In addition to liver enlargement, prolonged exposure to peroxisome proliferator compounds is associated with liver tumors in mice and rats.[32] Conversely, a decrease in PPAR activity results in damage to the immune response.[33, 34] PPARα knockout mice develop hypoglycemia and hypertriglyceridemia.[35] PPARα ligands also repress bile acid synthesis and promote phospholipid secretion into bile, counteracting liver disease.[36] When PPARs function improperly, homeostatic lipid metabolism becomes dysregulated and can result in liver and systemic damage while also weakening the organism to further insult.

**CAR – Constitutive Androstane Receptor**

CAR is a constitutively active NR which acts as a sensor of endobiotic and xenobiotic substances and targets genes involved with bilirubin conjugation as well as xenobiotic and energy metabolism (particularly cytochrome P450 expression).[23] As it is constitutively active, regulation of CAR is achieved through agonists and inverse agonists. Prior to binding a ligand, CAR is unable to translocate to the nucleus, thereby rendering it unable to bind DNA and influence transcription. In response to the presence of relevant ligands, CAR translocates to the nucleus where it dimerizes with retinoid X receptor (RXR). The CAR-RXR dimer binds specific response elements resulting in the upregulation of genes responsible for metabolism and the excretion of metabolic products.[32, 37]
Due to CAR’s upregulation of P450 enzymes, the liver is a primary site for CAR mediated
detoxification of xenobiotics. One such xenobiotic that acts as a prototypical CAR activator is
phenobarbital. When activated, CAR induces CYP2B which causes hepatic hyperplasia and hypertrophy
through proliferation of smooth and rough endoplasmic reticulum to increase the synthesis of necessary
metabolizing proteins. The hypertrophy generally affects all centrilobular hepatocytes and 45% of
periportal hepatocytes in rats. Once the stimuli has been metabolized the tissue returns to normal as the
increased protein production is no longer required.[32] However, prolonged exposure to phenobarbital
can results in induction of hepatocellular neoplasms in mice and rats.[32]

CAR null mice have significantly impaired drug metabolism as CAR regulates CYP3A, 1A and 2E, in
addition to 2B.[23, 37] Without proper ability to metabolize and detoxify compounds, an organism will be
unable to mitigate xenobiotic threats as well as properly utilize necessary compounds. Chen et al[38]
showed that without functional CAR, alcohol induced liver damage is achieved with less exposure and
greater tissue damage. Relative to wild type, CAR deficient mice also experienced more severe steatosis
and greater susceptibility to ethanol induced hepatocyte apoptosis.[38]

\textit{AhR – Aryl hydrocarbon Receptor}

Aryl hydrocarbon receptor (AhR) is a cytosolic member of the basic helix loop helix family of TFs
originally isolated and described as a sensor of environmental toxicants. Prior to activation through ligand
binding, AhR is bound to co-chaperone molecules which keep it in an inactive state. Ligand binding causes
the co-chaperone molecules to disassociate, thereby activating AhR. When activated, AhR functions as a
regulator of metabolism by inducing phase 1 (CYP1A and 1B) and phase II (glutathione S-transferase and
UDP-glucuronosyltransferase) metabolic enzymes.[39]

When activated, AhR can have numerous protective effects on the liver such as liver enlargement
through hepatocyte hypertrophy, increased replication, and decreased apoptosis.[38] In response to
toxins AhR mediated induction of CYP1A1/1A2 acts to protect the liver against chronic inflammation.[37, 38] However, some of the xenobiotics which AhR functions to metabolize would not elicit toxic responses without the presence of AhR – AhR induction leads to bioactivation of these compounds. Such compounds include dioxins and benzo[a]pyrene: AhR knockout mice show resistance to dioxin and benzo[a]pyrene-induced toxicity.[40] AhR has a significant role in metabolism through its ability to induce CYP enzymes and has generally cytoprotective activity but can also play a role in toxicity of select compounds by facilitating their metabolism to reactive moieties.

**NFkB – Nuclear Factor kappa-b**

Nuclear Factor Kappa-b is a ubiquitous, pro-inflammatory rapid acting transcription factor – it is constitutively present in an inactive state and does not require protein synthesis to be activated. This allows NF-kb to quickly respond to cellular stimuli such as ROS, interleukin 1β (IL-1β), and bacterial lipopolysaccharides. Prior to activation, NF-kb is located in the cytoplasm and kept in an inactive state by inhibitors of kB (IκBs). Activation of NF-kb occurs through the degradation of IκB proteins in response to extracellular signals. Once disassociated from the IκB protein, NF-kb enters the nucleus and activates transcription of the appropriate gene(s).[41-44]

The ubiquity of NF-kb – it is present in most cell types – is reflected by diversity of stimuli that can activate the TF and by the multitude of cellular pathways it is involved in; particularly, oxidative stress, inflammation, immunity (innate and adaptive), and as a regulator of clock controlled genes and memory. Within each of these pathways, NF-kb mediates a spectrum of responses with the most prominent being inflammation. Indeed, when NF-kb deficient mice (p50(-/-)) were treated with phenobarbital, the p50/-mice liver had higher levels of cell proliferation and apoptosis: showing a decreased ability to cope with the exposure.[45]
Normally, when NF-kb is activated in the liver— in response to xenobiotic and toxic stimuli (such as acetaminophen) — it activates the transcription of genes such as tumor necrosis factor (TNF), interleukin 1 and 6 (IL-1, IL-6), chemokines, vascular endothelial growth factor (VEGF), cell, surface adhesion molecules, etc.[15, 19, 46-48] Once the stimulus is neutralized, NF-kb activity is suppressed in order to decrease inflammation as well as limit carcinogenic proliferation. Should NF-kb become dysregulated (generally through mutation or chronic activation) it can be a powerful initiator of tumorigenesis.[48] NF-kb’s role in tumorigenesis is further strengthened as NF-kb is constitutively active in most tumor cells, almost all gene products linked to inflammation are regulated by NF-kb, and NF-kb is activated by known carcinogenic risk factors including tobacco, stress, obesity, alcohol, irradiation, and environmental stimuli.[47] While NF-kB has a major role in inflammation, is an important TF regarding numerous other cellular functions and it is tightly linked to carcinogenesis.

**HNF6 — Hepatic Nuclear Factor 6**

HNF6 is a member of a set of liver-enriched transcription factors (HNF1-6) found in the nuclei of hepatocytes. Its function is to regulate and mediate biological processes best known to occur in the liver and pancreas. As a tissue specific gene, HNF6 can respond to microenvironmental factors such as oxygen supply, glucose, concentration, local pH, etc. and readily alter its expression as needed by different areas of a single tissue.[7] In addition to micro-environmental adaptations, HNF6 is extremely important to proper liver function and development as its activity has been correlated to cell proliferation and migration, cell cycle regulation and cell-matrix adhesion, organogenesis, bile homeostasis, inflammation, and glucose metabolism.[49-51]

Although HNF6 is able to regulate normal liver function via direct DNA binding, HNF6 also expands its regulatory impact on liver function by influencing the activity of FOXA2, which requires HNF6 as a co-factor for transcription. HNF6 binds FOXA2 resulting in the recruitment of p300/C/EBP coactivator
proteins; thereby activating FOXA2.[7] Once activated, FOXA2 has numerous functions, including regulation of glucose metabolism. The importance of FOXA2, and therefore HNF6 as its regulator, is shown with FOXA2 knockout mice, which die during gestation due to severe defects in gastrulation prior to liver formation. Conversely, HNF6 knockouts survive gestation but fail to develop a gallbladder and show severe defects in the formation of intrahepatic and extrahepatic bile ducts.[49]

In adults, there is little known about the function of HNF6 in response to injury. What has been observed is that during liver injury and repair, up-regulated HNF6 expression was hepato-protective. One such impact seems to be the promotion of inflammatory action by repressing glucocorticoid action through direct HNF6/glucocorticoid receptor interaction. However, any damage must be mitigated quickly as chronic and sub-acute inflammation can significantly downregulate expression of HNF6 thereby eliminating its hepatoprotective properties.[49-51] Overall, knockout studies have established HNF6 as a vital TF in liver physiology and homeostatic maintenance but the full extent of its functions is yet to be well characterized.

AP-1 – Activating/ion Protein 1

AP-1 is a pro-inflammatory TF family which consists of structurally and functionally related members of the Jun (c-Jun, JunB, JunD) and Fos (c-Fos, FosB, Fra-1, Fra-2) protein families.[46] Each of the subunit families share a basic leucine zipper domain which allows the proteins to form dimers capable of regulating different genes depending on which proteins form the complex. Such diversity allows AP-1 to target a broad range of DNA-binding sites and genomic functions.[52]

Of all the pathways AP-1 is known to be involved in as an actor, its best characterized function is regulating the liver’s response to cellular damage. Whether by chemical (xenobiotic) or physical trauma, liver damage will induce AP-1, as an immediate cellular responder, to upregulate Cox-2, initiating an inflammatory response. Once AP-1 is actively up-regulating inflammation the TF also has considerable
influence on cell survival and apoptosis: depending on the extent of damage or the duration of stimuli, AP-1 can induce either proliferation or apoptosis. Proliferation occurs through AP-1 mediated increases in growth factors such as cyclin D1 and suppression of genes such as p21/waf and p53. Apoptosis is achieved through phosphorylation of the AP-1 subunit Jun resulting in the enhanced transcription of pro apoptotic genes such as Fas ligand (FasL) and Tumor Necrosis Factor α (TNFα).[15, 52, 53]

Given the influence AP-1 possesses for determining cell fate, it is unsurprising that AP-1 also has an important role in the development of normal liver. Not only are certain AP-1 protein subunit knockouts embryonic lethal (Fra-1, c-Jun, JunB)[54], but interference with other subunits can have significant effects on hepatocyte proliferation. For example, the subunit Jun is induced during a postnatal developmental phase noted for rapid liver growth and expansion[52] and is critical for liver regeneration. This is observed in Junα mice – a strain with a perinatally introduced liver specific deletion of Jun – which show reduced hepatocyte proliferation and impaired liver regeneration after partial heptectomy.[52, 55]

P53

P53 is a tumor suppressor protein which is continually produce and degraded by healthy cells. In the absence of an insult, P53 is held inactive by its inhibitor, Mdm2. While bound to Mdm2, P53 becomes targeted for proteasomal degradation. In response to stimuli this inhibition is relieved and P53 is able to bind DNA and activate the appropriate genes. The outcomes of P53 activation are as diverse as the stimuli which activate it and include senescence, cell cycle arrest, apoptosis, self (P53) regulation, and others.[56, 57]

Because the liver has a high capacity for proliferation and regeneration, proper P53 function is vital to the maintenance of healthy tissue. When the liver suffers a detrimental exposure to a genotoxic chemical any damaged cells must be repaired or replaced. Within these cells P53 acts to regulate the cell cycle, preventing it from progressing if DNA damage is present, and if so, activating DNA repair
mechanisms. If the damage is not repairable, p53 will initiate apoptosis. Failure of this system often leads to uncontrolled cell proliferation and tumorigenesis. Indeed, the importance of P53 in carcinogenesis is evident in that most cancerous cells have a defective p53 response[56] and that p53 controlled functions such as DNA damage repair, and apoptosis are selected against in tumorigenesis.[58] Unsurprisingly, knockout p53 mice showed enhanced susceptibilities to spontaneous and carcinogen induced tumors.[59] Without functional P53, the cell has a decreased capacity to repair DNA damage and therefore is significantly more susceptible to carcinogenesis.

NRF2 - Nuclear Factor (erythroid-derived 2)-like 2

NRF2 acts as a sensory control switch for a cytoprotective network which attenuates toxicity during electrophile and oxidative stress and inflammation. Under basal conditions, NRF2 is bound to its cytosolic repressor Keap1 which targets the complex for proteosomal degradation. Both internal and xenobiotic stressors (ROS, RNS, lipid aldehydes, heavy metals, electrophiles, etc) can trigger a conformational change in Keap1, thereby releasing NRF2. No longer inhibited by Keap1, NRF2 migrates to the nucleus where it dimerizes with small Maf proteins before binding to promoters containing antioxidant response elements (ARE).[15, 60] In the liver, when NRF2 is activated it promotes the transcription of cytoprotective genes. Specifically, when activated by acetaminophen, NRF2 protects against liver injury by upregulating UDP-glucuronosyl-transferase 1a6 (Ugt1a6), multidrug resistant proteins (Mdr2-4), and glutamate-cysteine ligases (Gclc and Gclm).[61] Acetaminophen conjugates with glucuronic acid, as catalyzed by Ugt1a6, and glutathione, for which Gclc and Gclm are enzymes promoting its synthesis. Once conjugated, acetaminophen is excreted into bile via Mdr2.[61]

In their review, Osburn and Kensler (2008) aggregated studies utilizing NRF2 deficient mice. The mice showed increased susceptibility to hepatotoxicity, pneumotoxicity, and neurotoxicity as well as carcinogenicity and inflammation. Each study considered investigated a particular xenobiotic and the
aggregation included acetaminophen, bieomycin (an anti-cancer drug), Malonate, Diesel exhaust, and tobacco smoke, respectively to each condition. Without functional NRF2, the mice were unable to induce the necessary cytoprotective enzymes which include electrophile conjugating enzymes, antioxidant enzymes and glutathione (GSH) homeostasis, nor are they able to produce reducing equivalents and proteasome components.[62]

Inbred Mice

As shown in the preceding section, TFs have an incredible functional diversity even in the liver alone. Because of their diversity and overlapping integration (a TF may have numerous functions which require interactions with other, varied TFs) studies utilizing chemical exposures and/or knockout/null animals may not result in complete depiction of a TF’s activities. In order to better understand the impact of specific TFs on an organism we will consider them on a genetic level. TFs control gene expression as DNA binding proteins but TF binding is highly sequence specific and polymorphisms may have a major impact on binding affinity and therefore function.[63, 64]

To eliminate this variability, we can use inbred mice – which have known and fixed genotypes – to consistently investigate TF-DNA binding. A strain is considered inbred when it has gone through 20 or more generations of uninterrupted brother-x-sister matings. After 20 generations, on average, 98.6% of the loci in each mouse are homozygous and further brother-x-sister matings increase that percentage. Many strains are essentially homozygous at all loci as they have been bred for more than 150 uninterrupted generations. Genetic profiles of strains can then be created via sequencing allowing mice to be identified through genotyping.[65] The homozygosity of the strains makes them ideal for experimentation as the decreased genetic diversity of a single strain begets greatly reduced phenotypic diversity. Phenotypic homogeneity within and inbred strain decreases the number of mice needed to detect significant phenotypic differences. For the experiment detailed in materials and methods the
genetic similarity of the inbred mice is ideal, as we are interested in the conservation of expression within multiple individuals of a single inbred strain.

Genetical Genomics

Because the mice within each inbred strain are, effectively, genetically identical, the TF expression, function, and regulatory networks within each strain should be highly conserved. Because of this, it is feasible that a genetical genomics styled approach be applied to determine TF activity correlation with gene expression. Genetical genomics is the combination of genetic and gene expression approaches. It is applied to the study of the genetic basis of gene expression, due to the fact that mRNA levels of many genes are heritable. Qualitative trait loci (QTL) analysis determines the chromosomal regions responsible for steady state, basal expression of the genes of interest. QTL mapping is greatly expanded by genetical genomics, as it submits the entire genome for analysis and shows hundreds of heritable traits as well as the location of any significant QTLs.[66-68] If mRNA levels are heritable then TF expression levels may be as well, therefore the expression levels of interlinked genes and the TF(s) that regulate them should be very similar.

Genetic regulatory networks

Another reason TF activity should be consistent among individuals is their organization into regulatory networks. TF regulatory networks are multilayered systems stretching from specific DNA binding to the full blueprint for gene expression in an organism. Regulatory networks function to regulate the expression in an organism, activating and de-activating pathways as needed in response to stimuli or to maintain homeostasis. Babu et al[1] organize TF networks into 4 categories: Basic units, Motifs, Modules, and Transcriptional regulatory networks. A basic unit comprises the TF, its target gene with DNA binding site, and the interaction between them. Motifs are the simplest unit of network structure,
comprising specific patterns of regulation. These patterns can be very diverse: multiple genes can be regulated by a single TF or multiple TFs can cooperatively regulate a single gene. Modules are interconnected motifs responsible for regulation of cellular processes. Modules are discrete and separable based on the distinct processes they regulate, though not fully isolated. A Transcriptional regulatory network is the aggregation of all regulatory interactions and it acts as a blueprint for gene expression in an organism.[1] Because the animals used have identical genetic blueprints, their regulatory networks should be identical as well.

Differences in these individuals are unlikely as TF/target gene interactions evolve via general (eg. duplication) and lineage specific mechanisms (e.g., domain shuffling). Duplications, with or without divergence, are the mechanism responsible for more than half of the evolved interactions. TFs are less conserved than target genes, therefore TFs evolve faster than the genes they regulate.[6, 63, 64, 69] Given the genetic consistency of the inbred mice, such divergences should not interfere.

Methods to Measure Transcription Factor Activity

Regardless how consistent the TF activity may be or how identical the inbred mice are, they are of no practical value here without an accurate, repeatable method to measure the TF activity. To this end, there have been numerous methods introduced and we will now review those techniques leading up to the method we apply in this work.

**EMSA**

The gel electrophoresis mobility shift assay (EMSA) is a rapid and sensitive method which is used to detect and return qualitative data regarding protein-nucleic acid complexes. EMSA is based on the observation that the electrophoretic mobility of a protein-nucleic acid complex is generally less than that of unbound nucleic acid.[70] The assay is performed by combining solutions of proteins and (usually) radio-isotope labeled nucleic acids and subjecting the mixtures to electrophoresis through polyacrylamide
or agarose gel. After completion of electrophoresis, the distribution of species containing nucleic acid is determined by autoradiography of the labeled nucleic acid. Although it provides the greatest sensitivity, radio-labeled nucleic acid is not required for the assay. Other methods of measuring non-radioactively labeled nucleic acid include fluorescence, chemiluminescence, or immunohistochemical techniques.[70]

EMSA can be performed on a wide range of nucleic acid sizes from small oligopeptides to transcription complexes. Not only does the assay accommodate a wide range of binding conditions, it also works well with both crude cellular extracts and highly purified proteins. However, there are limitations to EMSA: nucleic acid-protein complexes are not at equilibrium during electrophoresis and may disassociate during that step. Further, the electrophoretic mobility of a compound depends on multiple factors beyond the size of a bound protein, therefore limiting the ability of the assay to measure molecular weights or protein identities.[70]

**ELISA based methods**

Improving on the sensitivity and breadth of EMSA, Benotmane et al developed two ELISA (Enzyme Linked Immunosorbent Assay) based methods to study Protein-DNA interactions.[71] The methods, termed direct and reverse ELIDA (Enzyme Linked protein/DNA Interaction Assay) did not require radioactively labeled isotopes and were shown to be more sensitive than the previous standard: EMSA. The first method consisted of immobilizing the DNA-binding proteins to microtiter plates, incubating the proteins with biotinylated DNA, and finally quantification the bound DNA with streptavidin and biotin-substituted horseradish peroxidase. The second method immobilized biotinylated DNA on streptavidin-substituted microtiter plates, incubated them with DNA-binding protein, and quantified the bound protein with specific antibodies.[71]

While these methods allowed for measurement of multiple TFs (DNA-binding proteins) in a single assay, further adaptation of ELISA techniques into chemiluminescent assays conferred greater robustness, reproducibility and a simpler adaptation to automation.[72] Combined, these methods advanced the
measurement on TFs with their ability to measure multiple TFs simultaneously and therefore provide data regarding TF relationships from individual samples.

However, problems remained. While these methods solved the problem of measuring multiple TFs simultaneously, other issues impacting the accuracy and robustness of the assay still existed. First, requiring multiple cultures introduced variables which did not allow for great accuracy in determining the relative activities of multiple TFs. More importantly, ELISA based methods depend completely on DNA/protein interactions, but despite being required for Transcription, DNA-binding is not a direct measure of TF activity.[73] TFs can be regulated by ligands which influence their interactions with co-activators or co-repressors independent of DNA binding. Some TFs constitutively bound to DNA but remain inactive until bound by an activator ligand. Therefore, DNA binding assays are not directly indicative of the TF activity, merely its interaction with DNA.

**Reporter Gene Assays**

Reporter gene assays overcome the limitations of DNA-binding assays by shifting the assay target from the DNA-protein complex to a direct indicator of TF activity. Through the introduction of a TF responsive DNA sequence, reporter gene assays measures a TF’s activity via a TF-responsive promoter which controls the expression of a gene encoding an easily assayable reporter protein. Because the TF will activate both the intrinsic gene(s) and the reporter sequence, the rate of transcription of the reporter gene is a 1:1 representation of the rate of transcription of the intrinsic gene(s) the TF of interest regulates. While ease of use and assay robustness helped these assays became the standard for functional TF measurement, there are notable concerns with the method.[73]

One such concern is that all reporter assays must achieve a high degree of certainty because a negative result due to low-level expression cannot be differentiated from a negative result due to low assay sensitivity. Further, the sensitivity and effectiveness of a reporter assay depends on the expression level within the cell, the stability of the reporter gene within the cell and the ability to accurately measure
the reporter gene.[73] Finally, a limited reporter protein library retards the use of these assays beyond a few TFs at a time. While this can be overcome by fabricating new reporter proteins specific for each additional TF, said proteins are vulnerable to differential alteration by unpredictable post-transcriptional mechanisms which raise questions about the accuracy and reliability of reporter gene/protein assays.[74]

Array Hybridization Methods

Array hybridization is able to overcome translational modifications because array hybridization of multiple unique reporter transcripts allows the sequences to be evaluated by their rate of transcription only, thereby preventing translational modifications from interfering with the output. To perform this assay, a sample of interest (TF) is mixed with a library of pre-labeled DNA-cis elements. Each TF binds its specific cis-element (sequence) to form a DNA-protein complex. DNA-TF complexes and unbound TFs are then separated using a spin-column. The DNA aspect of the complexes are then separated from their bound protein and eluted. Subsequently, this DNA will be hybridized to a DNA array to determine the identity and quantity of the sequences which the TF in question formed a complex with. This technology can be used to profile multiple TFs simultaneously.[75]

However, there are still shortcomings inherent to this method. During the mixing and complex forming step, different sequences will be transcribed at different rates simply because they are different. The secondary structures of the transcripts will also suffer differing levels of stability, leading to considerable uncertainty impacting the accuracy and reproducibility of the assay.[76, 77] Therefore, the heterogeneity of the reporter sequences is a problem.

Plasmid Reporter System

Romanov et al[4] describe a high capacity system capable of overcoming these shortfalls. Romanov’s system is dependent upon a library of reporter genes, each having a unique TF-responsive promoter linked to a very similar downstream reporter sequence. These reporter constructs are called “reporter transcription units” (RTUs). The only difference among each reporter transcript is a unique
endonuclease recognition site acting as an identifier. Using highly similar reporter constructs to generate nearly identical reporter sequences the method overcomes the problems inherent in the previously mentioned methods.[4]

Further, the use of highly similar reporter constructs to generate nearly identical reporter sequences evaluated by capillary electrophoresis, virtually eliminated non-specific detection background while achieving a robust, accurate, and repeatable multiplexing assay.[4] The system designed by allows for simultaneous measurement of multiple TFs/TF families in a single assay. Targets of the assay are only limited by the sequences that comprise the RTU library. As new TF promoters are sequenced they will expand the capacity of the RTU library. Near identical reporter sequences simplifies transcript isolation and processing into steps that affect all sequences simultaneously. Assays measuring TF activity are utilized for multiple purposes including drug discovery, predictive toxicology modeling, and the expansion of gene regulatory networks.

Overall Summary

Transcription factors function in a variety of ways to regulate the genetic and cellular actions of an organism. While this paper has largely considered their role in hepatic functions, the universality of cell proliferation, oxidative stress, inflammation, and metabolism speak to the importance of TFs on an organismal scale. Regulatory networks evolved to ensure the proper response to stimuli and insult as well as maintain necessary homeostatic function. Without a well defined, specific method of gene regulation an organism would neither be able to properly utilize the genetic information it consists of nor respond in a pro-survival manner given a harmful exposure.

Because TF binding is very specific we use inbred mice to control for any genetic polymorphisms within a single strain. These consistencies across the individuals of a strain combined with the knowledge that mRNA levels are heritable allow us to consider that TF activity should have the capacity to be meaningfully correlation with mRNA levels. Further understanding of what genes which TFs regulate will
help expand and add detail to known TF regulatory networks. However, without an accurate method to
measure TF activity, none of these goals can be reached. By reviewing the history of assays used to
measure TF activity we have shown how the method utilized here overcomes past limitations and is an
advancement of the field.
CHAPTER 3
Experimental Methods

Animals and Tissues. C57BL/6J, DBA/2J, 129S1/SvImJ, C3H/HeJ, BALB/CByJ, AKR/J, NZW/LacJ, A/J, FVB/NJ, NOD/ShiLtJ, LP/J, C57L/J, SM/J, PL/J, SJL/J, SWR/J, RIIIS/J, JF1/Ms, BUB/BnJ, and CBA/J are inbred strains of mice available from the Jackson Laboratory (Bar Harbor, ME). Male mice, 6 to 8 weeks of age, were maintained at 20-24 degrees C on a 14/10 hour light/dark cycle in a pathogen free colony at UNC. Mice were fed Agway Prolab 3000 rat and mouse chow (5% fat) and given tap water in glass bottles. Mice were given one week to acclimate prior to plasmid injection and two weeks post injection prior to sacrifice. Sacrifice was performed by 0.1ml intraperitoneal injection of Nembutal (Lundbeck Inc. Deerfield IL, 60015 USA). Whole liver was immediately removed and the left lobe and remainder was placed in separate 2mL Eppendorf tubes and frozen in liquid nitrogen. All animal studies for this project were approved by Animal Care and Use committee at the University of North Carolina at Chapel Hill.

Reporter Plasmid Injection. Plasmids were originally produced by Attagene (Morrisville, NC). Individually, each plasmid is referred to as a reporter transcription unit (RTU). As a group, they are referred to as a library of multiple RTUs (MRTU library). A 3mL solution containing the MRTU library was introduced into each mouse via high pressure hydrodynamic tail vein (HTV) injection [78] lasting no longer than 5 seconds. High pressure hydrodynamic tail vein injection will heretofore be referenced as 'hydrodynamic tail vein injection' unless otherwise noted.

Luciferase Assay. The pCMV-GLuc control plasmid was introduced into each mouse by hydrodynamic tail vein injection along with the MRTU library. Introduction by this method effectively forces the plasmids into the mouse tissues, including liver. 24 hours after injection, 10μl of blood was drawn and blood luciferase concentrations were measured in duplicate. Blood luciferase values were determined using the
Gaussia Luciferase Assay Kit (E3300L, New England BioLabs, Ipswitch, MA) according to manufacturer’s instructions. Luminescence was measured on a Beckman Coulter DTX 880 Multimode Detector.

**RNA Isolation.** Total RNA was isolated from liver samples (left lobe, ~45mg) using the RNeasy maxi kit (Qiagen, Valencia, CA) as detailed by the manufacturer. RNA quality and quantity was determined spectrophotometrically from the absorbance at 260nm and 280nm on a Nanodrop ND-1000 spectrophotometer. Aliquots of RNA samples were frozen at −80 degrees C until mRNA isolation.

**mRNA Isolation.** mRNA was isolated from Total RNA aliquots using the mRNA midi kit (Qiagen, Valencia, CA) as detailed by the manufacturer. RNA quality and quantity was determined spectrophotometrically from the absorbance at 260nm and 280nm. Aliquots of mRNA samples were frozen at −80 degrees C until reverse transcription.

**Reverse Transcription.** mRNA was reverse transcribed using Ambion RETROscript Kit (Applied Biosystems, Foster City, CA) as detailed by the manufacturer. Aliquots of cDNA were frozen at −80 degrees C until further processing. Prior to PCR amplification, in order to remove primers, excess dNTPs, salts and any other low molecular weight material, samples were run through Performa DRT Gel Filtration Cartridges (Gaithersburg, MD) as detailed by the manufacturer.

**PCR Amplification.** PCR amplification was performed using Ambion SuperTaq Plus DNA polymerase (Applied Biosystems, Foster City, CA) as detailed by the manufacturer. Primers sequences are described in Romanov et. al [4]. Finished samples were immediately run through QIAquick PCR Purification Kit (Qiagen, Valencia, CA) as detailed by the manufacturer.
Fluorescent Labeling. Sample labeling was performed with labeled primers designed by Romanov et al.[4] but fabricated and purchased from TIB MOLBIOL (Berlin, Germany). Samples were no longer exposed to light for the remainder of the methods and stored at -20 degrees C until endonuclease digestion.

Endonuclease Digestion. cDNA digestions were performed using the restriction enzymes HPA1 followed by Proteinase K, both purchased from New England Bio Labs (Ipswitch, MA), as detailed by the manufactures. Samples were then sent to the UNC sequencing core to undergo Capillary electrophoresis. Capillary electrophoresis allows the determination of base pair length, and therefore RTU identity, via elution against a known ladder, and concentration quantity of the RTU transcript via the size of the fluorescent peak.

Transcription Factor Analysis. Data received from the UNC sequencing core consisted of the order of elution against a standard (fragment size/height) and associated fluorescence (peak area). Using the size/height of the fragments, we were able to determine which RTU each fragment corresponded to. Peaks were then scaled to the maximum and normalized to the RTU M-32. The Transcription Factor Analysis was performed using R Statistical Analysis software (http://www.r-project.org/). False Discovery Rate cutoffs from 0.20 to 0.05 were used for Pearson and Spearman rank correlations as well as EMMA analysis. mRNA data used was from Harrill at al[79] whose array (Agilent-011978 Mouse Microarray G45121A) contained 20,868 mRNA transcripts. Their array data was edited to only include information from the inbred mouse strains also used in this experiment. Correlations were performed using the R statistical package using the normalized data. The online software 'Pscan'[80] (http://159.149.109.9/pscan/) was used to identify overrepresented TF binding motifs in genes of interest. The Transfac Mouse TF database was used to compare significant mRNA-RTU pairs for known associations.[81]
CHAPTER 4
Results and Discussion

In-vivo introduction of plasmid reporter system

Previous use of the MRTU method was in-vitro only.[4] While previous work has shown successful transfection though hydrodynamic tail vein injections we needed to determine if and how well the plasmids were entering cells. In order to determine transfection efficiency, we used expression of secretable form of luciferase, introduced as part of the plasmid mixture. Blood Luciferase served as an indicator of plasmid transfection rate. Luciferase levels show a positive correlation with total RTU counts from isolated tissue, suggesting luminescence corresponds to transfection. These measurements are not standardized across strains because the CMV promoter controlled luciferase plasmid will be transcribed at strain specific constitutive activity levels.

After all samples were processed, it was noticed that 46 samples had either poor (21) or no (25) data. It was presumed that these results were due to either human error during processing or a failure of the plasmids to properly transfect and/or the injection was unsuccessful. To determine the cause, the 46 samples were reprocessed from their total RNA isolate forward and the new results were reintroduced to consideration for the final analysis while also used a method to assess the reproducibility of the assay. Most of the sample data improved from reprocessing, although there were certain animals which returned high correlation to the original data. The latter results were likely not due to human error. For example, 3 RIII/J samples were reprocessed for obvious batch effects [technical sources of variation that may have been added to the samples during handling] and returned the same data. Assuming no error, this begins to display the high degree of reproducibility inherent with this system.

The reprocessed samples were added to the full data set, replacing the corresponding original data. Using this more complete dataset, individual animals and circumstances were considered to remove any animals that failed to properly complete the protocol. Individual mice were removed due to failed or
poor injection, death during injection, low luciferase level (indicative of poor transfection), failure to return data, and obvious batch effect. Finally, in order to apply a highly objective, mathematical cutoff, animals showing an intrastrain correlation value outside 2 standard deviations from the mean of its strain were also removed. Due to this cut off, 11 additional mice were removed from the data set. 4 strains (A/J, BUB/BnJ, JF1/Ms, LP/J) were completely removed for poor correlation due to a combination of the above factors – a total of 45 mice. The number of mice removed from the analysis for each exclusion criteria is shown in Table 1. Of the 214 mice across 20 strains initially used in this experiment, 112 across 16 strains were included in the final analysis (Figure 1).

The most important result is that the protocol was successfully applied to an in-vivo study. Mice survived the treatment and returned useable data. In past investigation, hydrodynamic tail vein injection has been shown to have diminishing transfection results with larger DNA sequences and plasmids [78] but the MRTU library was successfully introduced into the liver. The chosen method of measuring transfection efficiency – luciferase levels – is not specific for hepatocytes. Instead, it represents transfection across all cells in the liver. Such cellular non-discrimination is supported by previous investigation of hydrodynamic tail vein injection.[78] While this non-discrimination is a road block to in-vivo TF measurements in individual cell types, it also results in the ability to encompass tissue-specific interactions. Because tissue function requires the synergy of multiple cell types, transfecting multiple cell types will provide greater insights into TF interactions across the liver. In summary, this experiment not only validates in-vivo application of the protocol but expands the breadth of what genetic mechanisms can be catalogued by in-vivo studies.

Intrastrain Correlation

Our results returned high intrastrain correlations which denote a high degree of consistency of TF activity in each inbred strain (Figure 2). Such is to be expected as the individuals within each inbred strain
are effectively clones. The passing of genetic information from parent to offspring is the hallmark of sexual reproduction. In this manner, anything encoded into one's DNA can be inherited through generations. Classically, heritability is reserved for direct transfersences such as alleles and mutations. However, genetical genomics — a field which combines gene expression studies (microarray studies, etc) with genetic linkage analysis — shows that mRNA levels are heritable and thus amenable to genetic analysis.[66] Regarding inbred animal strains, in which animals are genetically identical, assuming equivalent treatment, intrastrain TF activity should, theoretically, also be identical.

Combining the ability to simultaneously measure the activity of multiple TFs and the ideas of genetical genomics yields the possibility that TF activity levels are heritable. Regarding the mouse strains utilized, intrastrain correlation ranges from 0.630 (PLJ) to 0.873 (C56LJ) with an average value of 0.748. Such a correlation level is consistent with expectations and shows that individuals within an inbred strain highly similar TF activity. Given such inter-animal similarity in TF activity; if mRNA levels of certain genes and the regulatory mechanisms that control them are heritable perhaps it is possible that certain TF activity levels are also heritable.

Presently, there is scant evidence for this conjecture. Regarding regulation by QTL, it is known that cis-acting effects seem to occur non-specifically while trans-acting effects are more likely due to tissue-specific regulation. Also, trans-effects are typically broader and more complex than cis-effects, in order to reflect the culmination of genetic, epigenetic, and environmental factors an organism is subject to. As a result, it has been suggested that trans-acting polymorphisms in QTLs may be coding variants in TFs.[66] Further, in yeast, hundreds of co-regulated target genes affected by trans-effect regulators were transcription factors.[82]

While the most likely explanation for differential intrastrain TF activity is the difference in injection/transfection efficiency among mice, it is possible there is a regulatory hierarchy with which at least some TFs are involved. As more information regarding TF activity within strains is elucidated, certain
patterns of similarity may become evident. If so, investigating the heritability of TF activity could become 
a compelling endeavor. Currently, understanding that there are TF activity differences resulting in 
differential genetic regulation should help lay a foundation for idiosyncratic responses.

**Interstrain Correlation**

Our results show that there are significant interstrain differences in TF activity across multiple 
strains of inbred mice (Figure 1). It has long been known that genetic differences exist among animals, but 
the extent of interstrain differences in TF activity within inbred mice remained unmeasured. By applying 
the plasmid reporter system described by Romanov et al [4] to multiple strains of inbred mice, we were 
able to quantitatively measure differences in TF activity by strain (Figure 1). Each pixel in figure 1 
represents an individual animal, and the groups of animals comprising each strain are delineated. 
Therefore, figure 1 shows us the correlation of the data from each animal/strain with the data from the 
remaining animals/strains.

Because the inbred strains used in this experiment were chosen to maximize genetic diversity, it is 
not unexpected that the data from certain strains does not correlate well overall. Regardless, interstrain 
correlation is relatively high and consistent with the known genetic relatedness of the chosen strains. 
Consistent negative TF activity correlations are limited to AKR/J, BalB/CByJ, C3H/HeJ, SJL/J, and SM/J. All 
of these strains except SJL/J are classified as ‘Castle’s mice’ and are not closely related to each other. 
C57L/J and C57BL/6J show similar TF activity (interstrain correlation of 0.500) and are separated by only 4 
generations. The closest related strains based on breeding are SWR/J and SJL/J, which share a great 
grandparent. As expected, the interstrain correlation between SWR/J and SJL/J is fairly high (0.633).

The more closely related inbred strains showing higher interstrain correlation is suggestive for TF 
activity being, to some extent, of a genetic basis. While it is currently known that TFs are less conserved 
than the genes they regulate – suggesting an evolutionary preference for regulatory plasticity – as we
learn more regarding TF activity in animal models, certain TF pathways are known to be well conserved across species.[1] As more TF activity data is procured across different strains of a model animal (e.g., mice) or even different species (e.g., mice and dogs), profiles of TF activity can become useful in strain/animal selection and possibly the understand of seemingly idiopathic effects of stimuli.

Transcription Factor Activity Profiles

Our results showed both varying interstrain correlation (Figure 1) and high intrastrain correlation (Figure 2). Based on these differences (high variation between strains and low variation within a strain) being indicative of differential activity of the TF families considered in this study we investigated the TF activity profiles of the strains used in the analysis. From the normalized data we isolated each individual strain’s MRTU data, and were able to construct RTU activity profiles for each strain of inbred mice used in the analysis (Figure 3). From these profiles it is clear that different strains show variable basal TF activity levels.

By creating a RTU activity profile for different strains/animals, we will be able to compare the TF activity of each strain/animal and potentially determine how likely a mechanism of action is to be shared between the strains/animals, thereby establishing a degree of biological plausibility. The impacts of a chemical exposure to a strain of mice, for example, will carry greater consideration to human exposure if the genes/proteins involved in mouse toxicity also exist in humans. Well conserved pathways and/or proteins, such as WNT signaling[83] and P53/Mdm2 function[84] successfully translate from animal models to human models due to their high conservation across species. Because of this, research done on animals regarding these pathways can be confidently translated to humans. Non-conserved proteins and mechanisms of action do not and cannot. For example, MTBE has various modes of toxicity in rodents which are not applicable to humans because humans lack the proteins involved in the rodent pathways.[85]
Given that transcription factors are known to have binding site motifs conserved across species, and how the specificity of TF binding is essential to proper regulatory network function and gene expression, models with similar RTU activity profiles to human should be more indicative of how a human TF response to a chemical of interest will proceed.[63, 64] However, expression differences among strains demonstrate that the functionality of these TFs are not necessarily reliant on strict inter-related expression levels. This implies that even highly conserved mechanisms exist on a gradient across different genetic backgrounds.

Regardless, understanding the differences in how organisms respond to stimuli (physical, chemical, etc.) will provide a regulatory comparison between animal model data. As knowledge of TF activity grows, more data can be considered for the formation of TF regulatory networks. As these networks expand across same-animal strains and different species they can be used to make more specific comparisons as to how organisms will respond to a stimuli, or if that comparison is valid on a shared genetic level.

**Illuminating Transcription Factor Interactions via RTU-RTU correlation**

Transcription Factors can have multiple and varied pathways of involvement which overlap with other Transcription Factors. Using the RTU data from each animal within the 16 inbred stains which completed the analysis we were able to show the organized correlation of the MRTU library with itself (Figure 4). Given the breadth of knowledge on TFs currently available in the literature, it was expected that RTUs with high correlations would be well represented in the literature. However, not all TF relationships are currently known, so this data can potentially reveal TFs which have yet to be investigated as partners in yet-described pathways. Roughly 100 pairs and clusters of highly correlated RTUs were subjected to literature reviews. Table 4 summarizes the results of this search. Well known TF relationships such as NK-kb and AP-1, p53 and Myc, and ISRE and E-Box are well represented.
There is a great diversity within table 4 regarding function and pathway involvement, including adipogenesis, cellular differentiation, stress response, DNA damage, oncogenesis, metabolism, and limb regeneration. Because the RTU-RTU correlations returned such depth and breadth regarding known TF interactions, it is reasonable to consider that highly correlated pairs/groups which returned no published material, likely have integrated functions which have yet to be described.

Based on the RTU expression data used to construct figure 4, a cluster dendrogram diagramming the relatedness of the MRTU library was created (Figure 5). Figure 5 visualizes the correlation of the RTU library across its members. Known relationships such as NF-kb and AP-1 are represented in the figure, which provides a snapshot of overall MRTU correlation. Figures 4 and 5 represent a step forward in TF regulatory networks by providing in-vivo data on multiple, interacting TFs from a single assay. The data is a direct measurement of the genetic regulatory mechanisms within mouse liver and effectively a snapshot of the basal TF activity within the inbred strains. In conclusion, we show that our data is representative of confirmed TF-TF interactions and therefore likely to be an accurate representation of the basal TF activity of inbred mice.

RTU and mRNA correlation

mRNA values used in this study originated from the control mice of a past study as noted in chapter 3. The mice were from identical strains and therefore should be acceptable substitutes for the mice bred for this experiment. Normalized and non-normalized data were both analyzed by False Discovery Rate (FDR) corrected Pearson and Spearman rank correlations. Pearson correlation was utilized so that the full range of the data could be considered. Spearman rank correlation was utilized to compress the data range and therefore prevent very high or low values from exerting a disproportionate weight on the correlations. Significant gene-RTU matches are show in Tables 2 and 3. Matched genes were then input into the web-based software 'Pscan' for a known relationship between the gene and RTU, effectively
creating a circle of association: RTU data was used to identify significant genes which were then investigated independently to determine if they had known TF regulation consistent with the RTU for which the gene was originally matched. Significant genes were also compared against the Transfac Mouse TF database for matching TF association.

Using the Transfac database, significantly correlated genes were investigated for known transcriptional control mechanisms. One gene returned a match; AP-1 was matched to GRE via the gene Plau (urokinase-type plasminogen activator). Okan et. al. (2001) showed that transcription of the urokinase plasminogen activator receptor (uPAR) gene is activated via an AP1-dependant mechanism using an AP1/GRE TF binding site.[86]

Literature searches were performed on all gene-RTU matches. The result of the searches led to matches being separated into direct (Xbp1+Slc39a6 and CREB+Sdc1) and indirect (Ubash3b+CREB5 and Ido2+AhR) match categories. Indirect matches were those where both the gene and RTU were involved in an overarching condition such as disease but did not have a direct relationship.

**Ubash3b+CREB5**

A genome wide association study in Behcet’s disease using 152 patients and 172 ethnically matched healthy controls showed that both CREB5 (Cyclic AMP-responsive element-binding protein 5) and UBASH3B were associated with Behcet’s disease, though CREB5 association was non-significant (p>0.05). Single Nucleotide Polymorphisms (SNPs) in UBASH3B were associated with Behcet’s disease. UBASH3B negatively regulates T-cell receptor signaling.[86]

**Ido2+AhR**

Aryl Hydrocarbon Receptor (AhR) activation by 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) results in immune suppression in C57BL/6J mice. AhR activation by TCDD induces indoleamine 2,3-dioxygenase-
like protein 1 and 2 (IDO1, IDO2). IDO plays a critical role in leading to the generation of regulatory T-cells which mediate immune suppression through AhR activation.[87]

**Xpb1+Slc39a6**

Real-time Quantitative Reverse Transcription (RT)-PCR was used to compare mRNA expression of 560 genes in Estrogen Receptor $\alpha$ positive and Estrogen Receptor $\alpha$ negative human breast tumors. XBP1 and SLC39A6 co clustered with ER$\alpha$ in a microarray analysis of breast tumor biopsies.[88]

**CREB+Sdc1**

Transcriptional Targets activated by Edema Factor (EF) and CREB activates the heparan sulfate transmembrane proteoglycan Syndecan-1 (Sdc-1) gene which is critically required for pathogen induced host cell migration. Sdc1 is a target of CREB. CREB is activated through EF induced cAMP signaling due to anthrax infection. CREB activates Sdc1 which plays a crucial role in Edema Toxin-induced macrophage migration.[89]

Another question impacted by the low number of significant gene-RTU pairs was whether differential gene expression can be explained by differential TF activity. RTU-gene Pearson correlations returned 8 RTUs total, with 4 returned for both normalized and non-normalized data (Table 2). Spearman rank correlation returned 13 RTUs total with 8 returned for both normalized and non-normalized data (Table 3). Each RTU which returned significant gene correlation from both normalized and non-normalized data (from both Pearson and Spearman rank methods) except one shared significantly correlated gene(s). This suggests that the mechanisms regulating these interactions are strongly conserved among strains.
However, relationships between TF activity and gene expression are difficult to elucidate here given the low number of significantly correlated genes. The initial FDR cutoff of 0.05 was relaxed up to 0.20 in order to expand the gene list but even that resulted in fairly small additions. Further relaxing of the FDR cutoff expanded the significant gene-RTU pairs into the hundreds but the cutoff was too great to warrant significance. Given the dimensions of the comparison (16 strains, 52 RTUs, and 20,868 mRNA values) it is not incredibly surprising that fewer than expected correlations were found. Numerous attempts to normalize the data (log scales, rank vs. non, etc.) designed to uncover more, stronger correlations failed. Further, the mRNA array data used here was from a different study. They were of the same strain, but not the same individuals, and as this work has shown, the gene expression among individuals in a single inbred strain can differ by nearly 40% (Figure2). Such a result should not diminish the application of the protocol, but instead raise questions into the use of array data which did not originate from the same animals used within the study.

Summary

The method devised by Romanov et al[4] was successfully translated from in-vitro to in-vivo while retaining the functionality which marked it as an advancement in the measurement of TF activity. Application of this method demonstrated that interstrain TF activity differed among 16 strains of inbred mice. Further, the data showed that intrastrain TF activity was highly similar in these strains. Using the full MRTU library data set we were able to construct TF activity profiles for each mouse strain as well as determining highly correlated TF-TF interactions which represent both well known and yet characterized relationships. Finally, by correlating the TF activity with microarray mRNA levels we found known TF-gene interactions implying a utility for this method (identifying gene-TF interactions) beyond solely measuring TFs.
CHAPTER 5
Conclusions

In conclusion, this study describes the successful transition of an in-vitro method to in-vivo systems allowing for accurate, repeatable, and robust measurement of TF activity within mouse liver tissue. As expected, we have found differences in TF activity across different strains of inbred mice. These differential TF activities allow for the creation of TF activity profiles for unique strains which can be utilized for identification and intelligent strain selection. Using this, transcriptional mechanistic considerations can be introduced when investigating the regulation of an individual gene or a complex pathway. Further, genetic similarity among individuals within an inbred strain (intrastrain correlation) is high but not equivalent which has implications for pharmacological and toxicological studies utilizing these models. This result shows not only the potential that, like mRNA, certain TF levels are heritable, but also contributes to the understanding of idiosyncratic responses.

The application of this method to correlating gene expression (mRNA levels) as an indicator of TF activity via RTU regulation is not yet demonstrated, though the potential is evident. Integration of this method into exposure studies will provide information on impacts to the gene regulatory network as well as altered gene expression. Finally RTU-RTU correlations prominently show known, strong TF-TF interactions. Therefore, TF regulatory networks can be expanded in scope and involvement through the RTU-RTU correlations showing yet investigated interactions.
CHAPTER 6
Future Directions

Expansion of the RTU Library

The current library of RTUs available to this protocol is 52. As more RTUs are developed more TFs as well as TF interactions can be investigated. Expanding the RTU library would also allow for study-specific RTU selection. The current library is a fairly diverse group of major TFs. Future studies could select from an expansive RTU library to investigate specific TF interactions under specific circumstances. As the library of RTUs grow and catches up to our understanding of individual TFs, specific sets of RTUs can be constructed to investigate certain tissues or functions. There may be physical limitations regarding how many RTUs can be introduced into a single animal which would prevent full scale measurement of most known TFs from a single animal.

TF Profiles for in-vivo study animals

This is the first instance of this protocol being used in-vivo. Expansion beyond the 16 inbred mouse strains into other mouse strains as well as other small rodent models (rat, guinea pig etc) would provide information on the genetic regulatory mechanism of a considerable number of animal models. TF Activity profiles of each strain/species should be created and catalogued. This would allow for more intelligent strain selection in experiments as the strains can be chosen based on genetic expression profiles.

Effects of Differential TF activity on gene expression

Further mRNA-RTU studies are necessary to determine the full ability of this protocol to illuminate gene-TF interaction. Ideally, a study could be performed in which the mice injected with the RTU library
are also the mice tested for mRNA levels on a microarray. Using a larger sample of mice per strain should increase the number of significant correlations by increasing the number of data points available.

With more genes returning for each RTU, we will be able to build TF regulatory networks focusing on individual genes, TFs, and/or known pathways, thus providing a genetic map of the organism/strain utilized. These genetic regulatory maps or profiles could then be used to further investigate gene/TF function as well as contribute to pharmacological and toxicological research through intelligent strain selection based on the varying genetic pathways in different animal models.

Novel Transcription Factor Associations

Transcription factors are ubiquitous and complex in their regulation; considerable breadth and overlap exist in the function of each. RTU-RTU correlations represented well established TF relationships; therefore it is reasonable to consider strong RTU correlations for TFs yet linked to each other as a guide for future investigations. Data from this work should illuminate relationships yet discovered and catalyze the specificity of TF research through directed TF selection and the ability to measure multiple TFs simultaneously. In addition to the full maps shown by Figures 4 and 5, individual RTU maps can be constructed showing the level of correlation among the library directed at a single RTU. This would allow for individual RTUs to be linked to pathways where their potential involvement is currently unknown.
Identifying primary and secondary pharmacological and toxicological targets across the ToxCast chemical library

For my practicum I worked with the National Center for Computational Toxicology (NCCT) of the United States Environmental Protection Agency (EPA) at Research Triangle Park in North Carolina (NC) under the direction of David Dix, Ph.D and Matt Martin, Ph.D. My project was to identify primary and secondary pharmacological and toxicological targets across the ToxCast chemical library for use in quality control (QC) assessment, assay selection, and as an internal reference.

In response to National Research Council (NRC) report Toxicity Testing in the 21st Century: A Vision and a Strategy (2007) federal agencies began work on large scale chemical toxicity screening. Two projects this created were Toxicity Forecast (ToxCast) and Toxicity Reference Database (ToxRefDB).

ToxCast is an EPA program designed to predict the toxicity of a compound based on bioactivity profiling and therefore act as a method to prioritize the thousands of chemicals that require toxicity testing.[90] ToxCast utilizes data from state-of-the-art high throughput assays capable of rapidly screening many genes for toxicity due to a compound of interest.[90] ToxRefDB is a database devoted to in-vivo animal toxicity studies which contains over 30 years and $2 billion worth of study design, dosing, and observed treatment related effects.[91]

High throughput studies for the ToxCast database are performed by 3rd party contractors. When the NCCT receives study data it has a limited time period to determine whether the data is of acceptable quality. Because of the size of the datasets, it is very difficult to review and isolate individual gene-chemical interactions of questionable plausibility. Therefore, the focus was shifted to the identification of well known, expected interactions which could be easily isolated in the datasets and serve as positive controls for the assay.
In order to create a reference list of such interactions, literature and industry knowledge investigations were performed on all ToxCast phase II chemicals – 776 distinct chemicals. Priority was given to pharmaceutical compounds with primary gene targets as the top interest of the investigations. To maximize the information on each chemical, CASRNs, chemical names, and any known synonyms (e.g. 103-90-2, Tylenol, Acetaminophen, Paracetamol, etc) were considered in each search. Sources of information primarily consisted of literature sources and aggregations of such, such as ToxRefDB, Pubmed, The national library of medicine, PubChem, and the Comparative Toxicogenomics Database. When available, studies performed by the US EPA were considered, as were industrial studies and manufacturer’s details regarding function, use, and toxicity.

Data logged for each chemical went beyond primary and secondary pharmacological and toxicological targets. The following data was considered and recorded (when available) for each chemical in the library: Mode of action, Target assay, Affected gene(s) (beyond secondary), Description of affected genes, Intended target classification, whether the compound is metabolically active (prodrug), Species the studies were performed on, Functionality of the chemical, Relevant detail regarding its use and toxicity, and an Assessment of the available data/evidence strength regarding the compound. Gene abbreviations recorded in the list are consistent with those found in ToxCast and ToxMiner (an interface for visualizing and analyzing ToxCast data).

Completion of the project resulted in a searchable data-set of the ToxCast chemical library complete with known gene targets and assays as well as functional uses of each chemical. Of the 776 chemicals investigated, 91 returned high confidence primary gene targets. This list is meant to be used internally as a general resource to quickly identify a chemical’s potential to interact with a gene target. By design, the list is amenable to the creation of chemical subsets by gene targets or assays or vice-versa. Because of this ability, the list can be functionally applied to quality control assessment of 3rd party chemical exposure data, thereby fulfilling its promise. Further, the list can be used for assay.
selection regarding a gene of interest, and chemical selection for a particular assay. Overall, the work should stand as a reference tool for quality control which can be expanded as more chemicals are investigated and can contribute to future work by guiding assay and chemical selection in exposure studies concerned with genetic targets.

This work is highly relevant to public health because it helps in the assessment and prioritization of toxicity testing while operating within the cost-effective approach of the ToxCast project. Increased quality control regarding high throughput toxicity data strengthens the identification of chemicals which pose the greatest human toxicity threat. The reference list can also be used as a guide for assay and/or gene selection for further study of a compound. Finally, by contributing to stricter quality control, the data may allow funds to be used more efficiently and increase investigations into chemicals of greater human toxicological relevance.
Figure 1: Strain Correlation Heatmap. Each row/column represents a single animal. Animals are divided and labeled by strain. The diagonal bisecting the figure is the self correlation of each animal (1.00). Boxes along the diagonal show intrastrain correlation.
Figure 2: Barplot showing intrastrain correlation. Plot was created using R Statistical Analysis software (http://www.r-project.org/).
Figure 3a: 129S1.SvImJ RTU Profile

Figure 3b: AKR.J RTU Profile
Figure 3c: BALB.CbYJ RTU Profile

Figure 3d: C3H.HeJ RTU Profile
**Figure 3e: C57BL.6J RTU Profile**

**Figure 3e: C57L.J RTU Profile**
Figure 3g: DBA.2J RTU Profile

Figure 3h: FVB.NJ RTU Profile
Figure 3i: LP.J RTU Profile

Figure 3j: NOD.ShiLtJ RTU Profile
Figure 3k: NZW.LacJ RTU Profile

Figure 3l: PLJ RTU Profile
Figure 3m: RIIS.J RTU Profile

Figure 3n: SJL.J RTU Profile
Figure 3a-p: RTU Profiles of 16 inbred mouse strains. Polar Plots for the 16 inbred strains which completed the experiment. Transcription Factor activity values were normalized to the constitutively active TF M32. Data was log₂ transformed. Plot was created using R statistical analysis software (http://www.r-project.org/).
Figure 4: Organized heatmap with Cluster Dendrogram and Histogram showing RTU-RTU correlation. RTU data from all 16 RI strains was incorporated into the plot. RTU data was normalized to the highly conserved and constitutively active transcription factor M32. Plot was created using R statistical analysis software (http://www.r-project.org/).
Figure 5: Cluster Dendrogram of RTU Library. Dendrogram show the relatedness hierarchy of the RTUs measured in the experiment using data from all 16 RI strains. Height is representative of the level of relatedness among the RTUs. Plot was created using R statistical software (http://www.r-project.org/).

Table 1: Mice removed from analysis with removal criteria

<table>
<thead>
<tr>
<th>Removal Criteria</th>
<th>Number Removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch Effect</td>
<td>3</td>
</tr>
<tr>
<td>Death</td>
<td>5</td>
</tr>
<tr>
<td>Poor/Failed Injection</td>
<td>23</td>
</tr>
<tr>
<td>Failure to Return Data</td>
<td>21</td>
</tr>
<tr>
<td>Low Luciferase Level</td>
<td>39</td>
</tr>
<tr>
<td>&gt;2SD from Intrastrain Correlation</td>
<td>11</td>
</tr>
</tbody>
</table>

102 animals across 20 strains were removed. 4 strains (A/J, BUB.BnJ, JF1.Ms, LP.J) were completely removed, accounting for 45 mice. Individuals removed due to poor/failed injection are different from individuals removed due to low Luciferase level. Low Luciferase levels can be caused by poor injection but are not classified as such because the plasmid library injection showed no abnormalities in these mice.
### Table 2: Significant RTU-Gene pairs using Pearson Correlation.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>FDR Cutoff(%)</th>
<th>RTU</th>
<th>FDR Cutoff(%)</th>
<th>Gene Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stk32a</td>
<td>0.2</td>
<td>PXR</td>
<td>0.1</td>
<td>Stk32a</td>
</tr>
<tr>
<td>Zfp449</td>
<td>0.2</td>
<td>PXR</td>
<td>0.05</td>
<td>Colr4</td>
</tr>
<tr>
<td>II28ra</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dsg1b</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defb41</td>
<td>0.1</td>
<td>CRE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldh111</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fkbp6</td>
<td>0.2</td>
<td>Ahr</td>
<td>0.15</td>
<td>Htr7</td>
</tr>
<tr>
<td>Entpd5</td>
<td>0.15</td>
<td>Ahr</td>
<td>0.15</td>
<td>Entpd5</td>
</tr>
<tr>
<td>Giyd2</td>
<td>0.15</td>
<td></td>
<td>0.2</td>
<td>Giyd2</td>
</tr>
<tr>
<td>2810004N23Rik</td>
<td>0.15</td>
<td></td>
<td>0.2</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amz2</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Htr7</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1700049J03Rik</td>
<td>0.15</td>
<td>DR4/LXR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>0.1</td>
<td>p53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ptpkrk</td>
<td>0.1</td>
<td>Pax6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ctsb</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ammecr1l</td>
<td>0.1</td>
<td>M-18</td>
<td>0.05</td>
<td>Ammecr1l</td>
</tr>
<tr>
<td>4933427G23Rik</td>
<td>0.1</td>
<td>M-18</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Cyp46a1</td>
<td>0.15</td>
<td>FoxO</td>
<td>0.05</td>
<td>Cyp46a1</td>
</tr>
</tbody>
</table>

Data was normalized by mean-centering. Returned genes significant for an RTU from both normalized and non-normalized data are written in red. Gene-RTU pairs shared between Pearson and Spearman correlation results are written in blue.

### Table 3: Significant RTU-Gene pairs using Spearman Rank Correlation.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>FDR Cutoff(%)</th>
<th>RTU</th>
<th>FDR Cutoff(%)</th>
<th>Gene Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ngb</td>
<td>0.1</td>
<td>EBOX</td>
<td>0.2</td>
<td>Ngb</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EBOX</td>
<td>0.2</td>
<td>Cpne8</td>
</tr>
<tr>
<td>Ngb</td>
<td>0.05</td>
<td>NFI</td>
<td>0.15</td>
<td>Ngb</td>
</tr>
<tr>
<td>Nt5c31</td>
<td>0.15</td>
<td>PXR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jmjd4</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gpr68</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atox1</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>0.15</td>
<td>GRE</td>
<td>GRE</td>
<td>0.2</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Cpne8</td>
<td>0.15</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>isoc1</td>
<td>0.15</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Cair4</td>
<td>0.15</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Socs7</td>
<td>0.1</td>
<td>GRE</td>
<td>GRE</td>
<td>0.2</td>
</tr>
<tr>
<td>H2afb1</td>
<td>0.1</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Plau</td>
<td>0.2</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Olfr1240</td>
<td>0.2</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>NA</td>
<td>0.2</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Apitd1</td>
<td>0.2</td>
<td>STAT3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1700111E14Rik</td>
<td>0.15</td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Ii13ral</td>
<td>0.05</td>
<td>TAL</td>
<td>TAL</td>
<td>0.05</td>
</tr>
<tr>
<td>Olfr799</td>
<td>HYPEXbp1</td>
<td>Xbp1</td>
<td>Xbp1</td>
<td>0.05</td>
</tr>
<tr>
<td>Adrb1</td>
<td>0.05</td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Ubash3b</td>
<td>0.2</td>
<td>CRE</td>
<td>CRE</td>
<td>0.2</td>
</tr>
<tr>
<td>Crip3</td>
<td>0.2</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Vopp1</td>
<td>0.2</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Hsbp1</td>
<td>0.2</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Ly6g6c</td>
<td>0.2</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Wdr8</td>
<td>0.05</td>
<td>TA</td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Cdk17</td>
<td>0.05</td>
<td>DR4/LXR</td>
<td>DR4/LXR</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Paip2b</td>
<td>0.15</td>
<td>VDRE</td>
<td>VDRE</td>
<td>0.05</td>
</tr>
<tr>
<td>Rab11fip4</td>
<td>0.05</td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Scd2</td>
<td>0.05</td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Inadl</td>
<td>0.05</td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>E130303B06Rik</td>
<td>0.05</td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>NA</td>
<td>0.1</td>
<td>M-61</td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>NA</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data was normalized by mean-centering. Returned genes significant for an RTU from both normalized and non-normalized data are written in red. Gene-RTU pairs shared between Pearson and Spearman rank correlation results are written in blue. Genes correlated to different RTUs from either normalized or non-normalized data are written in purple. Genes correlated to different RTUs from both normalized and non-normalized data are written in green.
Table 4 – RTU/TF Associations Indicated by RTU-RTU Correlation Represented in Current Literature

<table>
<thead>
<tr>
<th>Associated RTU/TFs</th>
<th>Associated Gene(s)</th>
<th>Pathways</th>
<th>Tissue</th>
<th>Functional Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.EBP, HNF6,1,4</td>
<td>Multidrug resistance-associated protein 2 (MRP2)</td>
<td>Hormone Response</td>
<td>Apical surface of hepatocytes</td>
<td>HNF1a,4a activate MRP2. HNF6,3, C.EBP inhibit HNF1a,4a. Acts at proximal GRE site in gene promoter for a hormone response.[92]</td>
</tr>
<tr>
<td>C.EBP, HNF6a,4a, 1b</td>
<td>Transcriptional repressor Tbx3 (T-box family)</td>
<td>Hepatocyte Differentiation</td>
<td>Liver</td>
<td>Tbx3 controls liver bud expansion by suppressing cholangiocyte (HNF6a, 1b) and favoring hepatocyte differentiation (HNF4a, C.EBP) in the liver bud.[93]</td>
</tr>
<tr>
<td>HNF4a, 6, C.EBPa,b</td>
<td></td>
<td>Liver Maturation</td>
<td>Liver</td>
<td>TFs used as markers of liver maturation. Other markers: tryptophan 2,3-dioxygenase (TO), and serine dehydratase (SDH).[94]</td>
</tr>
<tr>
<td>TGFb, HNF4a, HNF1b</td>
<td>HSP105B, CYP51, and C/EBPα</td>
<td>Hepatocyte-to-BEC (biliary epithelial cells) transdifferentiation</td>
<td>Bile Ducts/Liver</td>
<td>Loss of HNF4a and gain of HNF1b by hepatocytes with increased expression of TGF1b in perportal region appear to be mechanism of hepatocyte-to-BEC transdifferentiation.[95]</td>
</tr>
<tr>
<td>HNF6,1a,4a, FOXA2,3</td>
<td></td>
<td>HNF6 inhibited cell cycle progression in the G2/M and G1 phase in Caco-2 and HepG2 cell lines</td>
<td>Human colon Caco-2 and HepG2 carcinoma cell lines</td>
<td>FoxA2 inhibits HNF6, removal of FoxA2 allows HNF6 to re regulate Cell Cycle.[96]</td>
</tr>
<tr>
<td>C.EBP, GRE</td>
<td>GR, C/EBP, ApoD and Monoamine oxidase-A</td>
<td>Hormone Response</td>
<td>Skeletal muscle</td>
<td>Estrogen Related Receptors control GR (Glucocorticoid Receptor) which induces C/EBP, ApoD and Monoamine oxidase-A.[97]</td>
</tr>
<tr>
<td>C.EBP, GRE, Ahr, PXR, ARE,</td>
<td>GSTA2</td>
<td>GSTA2 repression</td>
<td></td>
<td>C/EBPα,b are involved in GC-dependent repression of GSTA2 expression and ARE sequences that bind C/EBPs appear to be critical for these responses. PXR dependant transactivation.[98]</td>
</tr>
<tr>
<td>C.EBP, GRE, HNF3, CREB</td>
<td>Carbamoylphosphate synthetase I (CPS) gene</td>
<td>Regulation of CPS expression</td>
<td>Liver</td>
<td>GRE, C.EBP, and HNF3 interact with enhancer for regulation of CPS expression. Cyclic AMP depends on integrity of GR pathway.[99]</td>
</tr>
<tr>
<td>C.EBP, PPRE, STAT</td>
<td>STAT</td>
<td>Adipogenesis</td>
<td>Human bone marrow derived-stromal cells (hBMSCs)</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------</td>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>C.EBP, PPRE, Nrf2, RXR</td>
<td>Glutathione S-transferase (GSTA2)</td>
<td>Glutathione S-transferase induction</td>
<td>PPARγ - RXR heterodimer promotes GSTA2 induction by activating PPREM in the GSTA2 gene, as well as inducing Nrf2 and C.EBPβ activation. [101]</td>
<td></td>
</tr>
<tr>
<td>C.EBP, SREBP, STAT5, KLF, Wnts, E2F, GATA</td>
<td>GCN5-mediated lysine acetylation, Sumo, SMAD proteins</td>
<td></td>
<td>C.EBP, SREBP1c, and STAT5 promote adipogenesis, while Wnts, E2F, and GATA inhibit the process. [102]</td>
<td></td>
</tr>
<tr>
<td>C.EBP, SREBP, Foxo1, TGFβ, BMP</td>
<td>Runx2, p57</td>
<td>Skeletogenesis</td>
<td>Wnt and BMP regulate Runx2. C.EBPβ up-regulates Runx2/p57. C.EBPβ helps regulate Runx2, which is the master regulator of osteoblast differentiation. [104]</td>
<td></td>
</tr>
<tr>
<td>C.EBP, AP1, NFKb</td>
<td>IL6, p38, XIAP</td>
<td>IL6 transcription</td>
<td>X-linked inhibitor of apoptosis (XIAP) regulates IL-6 transcription via NF-kappaB in cooperation with AP1 and C/EAST-beta. p38 reduces XIAP activity. [105]</td>
<td></td>
</tr>
<tr>
<td>C.EBP, AP1, NFKb, IL6, p21</td>
<td></td>
<td>Inflammatory/immune clustering response</td>
<td>Mouse urothelial carcinoma cell line MB49</td>
<td></td>
</tr>
<tr>
<td>C.EBP, AP1, OCT, NFKb,</td>
<td>IL6, IL17</td>
<td>Inflammatory/immunity</td>
<td>Deletion of NFKb or C.EBP eliminated 24p3 promoter activity in response to IL-17. API-OCT sites are overrepresented in IL17 target genes. [107]</td>
<td></td>
</tr>
<tr>
<td>C.EBP, CMV, DBP, CMV-DBP, CREB/ATF</td>
<td>Angiotensinogen I and II</td>
<td>Circadian expression of the human angiotensinogen gene</td>
<td>Angiotensin II circadian rhythm is regulated by CMV-DBP, CREB/ATF, and C.EBP. CREB binds CMV-DBP, and the resulting transcript binds C.EBP. [108]</td>
<td></td>
</tr>
<tr>
<td>C.EBP, CMV, YB-1</td>
<td>Heat response</td>
<td>SW480 human colon carcinoma cells (SW480/CMV CD)</td>
<td>YB-1 and C.EBPβ might mediate heat responsiveness of the CMV-IE promoter which could be used to construct heat inducible vectors. Cytomegalovirus-immediate early (CMV-IE) promoter is widely used as a strong and constitutively active promoter.[109]</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---------------</td>
<td>-----------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>C/EBP alpha expression vector (CMV alpha)</td>
<td>Cell proliferation</td>
<td>Cancerous cell lines</td>
<td>C.EBPα may play a role in maintaining the quiescent state of hepatocytes and other cells; effects not mediated through p53 or Rb or T-antigen. C.EBP alpha expression is high in non-dividing hepatocytes, but decreases during liver regeneration.[110]</td>
<td></td>
</tr>
<tr>
<td>SREBP, GRE/PRE, SP1, and ERE</td>
<td>Epitheliasin</td>
<td>Epitheliasin regulation</td>
<td>Epitheliasin transcripts: potential regulatory elements include SP1, SREBP, GRE/PRE and ERE.[111]</td>
<td></td>
</tr>
<tr>
<td>GRE, TCF.b.cat, AP.1,2, SP.1, CREB, E2F, EGR, TCF.1</td>
<td>Neuron-specific enolase (NSE)</td>
<td>Rat neuronal cells</td>
<td>Listed TFS are found in promoter region for Neuron-specific enolase (NSE).[112]</td>
<td></td>
</tr>
<tr>
<td>GRE, CRE, NF-kB, NF-IL6, and TRE</td>
<td>IL6, ARE, Nrf2</td>
<td>Hematopoietic and inflammatory processes</td>
<td>Mice</td>
<td></td>
</tr>
<tr>
<td>GRE, CRE, AP.1, CREB, SP1, c-Jun, c-Fos, GR</td>
<td>Thryotrophin-releasing hormone (TRH)</td>
<td>TRH transcription</td>
<td>Rat hypothalamic primary cultures</td>
<td></td>
</tr>
<tr>
<td>GRE, AP1, FGF2 response element (FRE)</td>
<td>Bone sialoprotein (BSP), fibroblast growth factor 2 (FGF2)</td>
<td>Osteoblast differentiation</td>
<td>Rat osteoblastic cell line ROS17/2.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FGF2 stimulates BSP gene transcription by targeting the FRE and AP1/GRE elements in the rat BSP gene promoter.[115]</td>
<td></td>
</tr>
<tr>
<td>Gene Expression</td>
<td>Function</td>
<td>Cells/Tissues</td>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>----------</td>
<td>---------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>GRE, AP1, Androgen Receptor (AR), CREB, c-Fos, c-Jun</td>
<td>BSP gene transcription</td>
<td>Murine F9 cells</td>
<td>AR stimulates BSP gene transcription by targeting the CRE and AP1/GRE elements in the promoter of the rat BSP gene. CREB, c-Fos, c-Jun antibodies disrupted complex formation.[116]</td>
<td></td>
</tr>
<tr>
<td>GRE, AP1, GR, c-Fos, c-Jun, c-Jun-c-Fos</td>
<td>GRE activation</td>
<td>Mice (liver)</td>
<td>GR activates GRE and AP1 if sites are close to each other (overlapping). If sites are distant GR will either activate or repress depending on subunit of AP1. GR synergized with either c-Jun or c-Jun-c-Fos at paired elements with GRE-AP1 site separations of &gt; = 26 base pairs. In contrast, paired elements with separations of 14-18 base pairs mimicked the composite GRE, i.e. GR synergized with c-Jun and repressed c-Jun-c-Fos.[117]</td>
<td></td>
</tr>
<tr>
<td>GRE, RORE</td>
<td>Bmal1, Tieg1</td>
<td>Circadian rhythm</td>
<td>GC box/TIEG1-mediated repression of Bmal1 promoter was additive to RORE-dependent repression by REV-ERBAα.[118]</td>
<td></td>
</tr>
<tr>
<td>E. Box</td>
<td>Period (Per1-3) and Cryptochrome (CRY1-2)</td>
<td>Circadian rhythm</td>
<td>E-box region of Cry1 promoter is necessary and sufficient for circadian Cry1 expression with an appropriate delay (4 h) relative to Per2. RORE not necessary. E-box (CACGTTG) is crucial for daytime transcriptional activity.[119]</td>
<td></td>
</tr>
<tr>
<td>SREBP, LXR, PPARα, DR1</td>
<td>Lipids and insulin metabolic regulation</td>
<td>Human liver cells</td>
<td>PPARα agonists enhance the activity of the SREBP1c promoter (a DR1 element) and act in cooperation with LXR or insulin to induce lipogenesis. SREBPs regulate the expression of a number of enzymes which catalyze the synthesis of fatty acids, cholesterol, triglycerides, and phospholipids.[120]</td>
<td></td>
</tr>
<tr>
<td>LXR, Foxo1, SREBP</td>
<td>SREBP-1c expression</td>
<td>HepG2 cells</td>
<td>LXRα binds to the LXRE region in SREBP-1c promoter elevating transcription, Foxo1 strongly inhibits LXRα. Inhibition can be partially alleviated by insulin.[121]</td>
<td></td>
</tr>
<tr>
<td>LXR, Myc, MMP7, BMP4</td>
<td>Beta-catenin (Wnt)</td>
<td>Colon Carcinogenesis</td>
<td>Colon cancer HCT116 cells</td>
<td>LXR ligand decreased mRNA expression of beta-catenin (Wnt) targets, MYC, MMP7 and BMP4, and recruited LXRs to MYC and MMP7 promoters.[122]</td>
</tr>
<tr>
<td>LXRα, C-Myc, PPARα, CD36</td>
<td>Low Density Lipoprotein receptor</td>
<td>Atherosclerosis</td>
<td>Inside arterial wall</td>
<td>GTPs downregulate PPARα, CD36, LXRα, C-myc, and upregulate LDR and PPARα (these TFs/genes control lipid metabolism, cytokine production and cellular activity within the arterial wall).[123]</td>
</tr>
<tr>
<td>p53, Oct-1, NfkB, AP1, cMyb</td>
<td>NO synthases</td>
<td>Inflammatory reactions, Nitrosative stress</td>
<td>In un-stimulated cells nitrosative stress increases NF-kB or AP-1 dependent transcription, while in activated cells nitrosative stress abolishes NF-kB or AP-1 dependent transcription.[124]</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------</td>
<td>---------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>p53, c-myc, Sox9 and Hnf1b</td>
<td>p53 Inactivation / oncogenesis</td>
<td>Pancreatic acinar cells</td>
<td>p53 (-/-) cells show increased expression of Ptf1a, Pdx1, Cpa1, c-myc, Sox9 and Hnf1b.[125]</td>
<td></td>
</tr>
<tr>
<td>TP 53, c-Myc, cyclin D2, p21WAF1/CIP1 and p27KIP1</td>
<td>G1 cell cycle regulation</td>
<td>B-cell chronic lymphocytic leukaemia</td>
<td>Progression of disease was accompanied by increases in p53, cyclin D2 and c-Myc mRNA expression.[126]</td>
<td></td>
</tr>
<tr>
<td>p53, PPARγ</td>
<td>GADD45 and p21(WAF1/Cip1)</td>
<td>Cell Cycle regulation</td>
<td>p53-dependent mechanism in the PPARγ ligand-mediated inhibition of cholangiocarcinoma growth (potential therapy implications).[127]</td>
<td></td>
</tr>
<tr>
<td>progesterone receptor (PGR), FOXO1A, cAMP</td>
<td>Insulin-like growth factor-binding protein-1 (IGFBP1)</td>
<td>Human cholangiocarcinoma cells</td>
<td>FOXO1A may be a target of cAMP and FOXO1A and PGR and FOXO1A interact; each influences the transactivating potential of the other.[128]</td>
<td></td>
</tr>
<tr>
<td>PPARγ, GRE, RXRa, NfkB, c-Myc</td>
<td>Immune Regulation</td>
<td>Nontransformed pro/pre-B (BU-11) and transformed immature B cells</td>
<td>Activation of PPARγ-RXR initiates a potent apoptotic signaling cascade in B cells (BU-11, and WEHI-231), potentially through NF-kB activation.[129]</td>
<td></td>
</tr>
<tr>
<td>SREBP-1c, Tcf.b.cat, PPARγ, Wnt, Glut 4</td>
<td>Adipogenic differentiation</td>
<td>Adipose tissue-derived mesenchymal stem cells</td>
<td>Chemicals inhibited expression of adipogenic markers PPARγ, SREBP-1c and Glut4, from mid-phase differentiation via activation of Wnt signalling via oestrogen receptor-dependent pathways.[130]</td>
<td></td>
</tr>
<tr>
<td>SREBP-1c, Wnt/beta-cat</td>
<td>Akt/PKB AMPK and MAPK pathways</td>
<td>Insulin resistance</td>
<td>SREBP-1c and Wnt10b display inverse expression patterns (they inhibit each other) during muscle ontogenesis and regeneration.[131]</td>
<td></td>
</tr>
<tr>
<td>SREBP-1, Tcf.b.cat, C/EBPa,b,d, PPARγ</td>
<td>Wnt signalling, WNT1, FZD1, DVL1, GSK3beta, beta-catenin and TCF1</td>
<td>Adipogenesis and insulin resistance</td>
<td>Expression of Wnt signalling genes and adipogenic signalling TFs (C/EBPa,b,d, PPARγ, and SREBP-1) were reduced in adipose tissue. Insulin resistant cells were enlarged; cell size was inversely correlated to expression of the Wnt signalling genes, adiponectin, and aP2.[132]</td>
<td></td>
</tr>
</tbody>
</table>

71
<table>
<thead>
<tr>
<th>CRE, BRE, cAMP, CREB</th>
<th>Bone morphogenetic protein (BMP)</th>
<th>Induced osteoblastic gene expression</th>
<th>C2C12 cells</th>
<th>cAMP-PKA/CREB/CRE signaling potentially enhances BMP-induced transcription through a BRE site in the promoter of the BMP-responsive gene through a PKA-mediated pathway.[133]</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFKB, IRF, MAFB, EGR, IRF, BCL6, AP1, Nur77</td>
<td>MEF2, BRI, HLX, HDAC5, H2AV, TCF7L2, NFIL3</td>
<td>Differentiation and inflammation</td>
<td>Monocytes to macrophages</td>
<td>LPS stimulation upregulates IRF, NFKb, NUR77, FOS, JUN, STAT3 and SMAD – classic role in macrophage function of inflammatory response.[134]</td>
</tr>
<tr>
<td>NFKB, E.box, c-Myb,</td>
<td>S phase entry, NFKb, c-Myb expression</td>
<td>Stellate cell activation, Hepatic fibrosis</td>
<td>Hepatic stellate cells (lipocytes)</td>
<td>Through the induction of c-myb and NFKb, Oxidative stress plays an essential role on stellate cell activation leading to hepatic fibrosis (c-myb and NFKb disrupt E.box promoter). c-myb and NFKb expression was inhibited by antioxidants.[135]</td>
</tr>
<tr>
<td>NFKB, AP1, ERE</td>
<td>Cyclo-oxygenase (COX) isoforms (NSAIDs block them for Arthritis relief)</td>
<td>Innate and adaptive immune responses</td>
<td>Estrogen receptors are cofactors for NFKB/AP1.[136]</td>
<td></td>
</tr>
<tr>
<td>NFKB, AP1, C/EBPb, phospho-CREB, NF-IL6, TCF-4/LEF-1 and c-Myb</td>
<td>IL1, TH1+TH2 cytokines, glucocorticoids</td>
<td>Cytokine transcription (inflammation)</td>
<td>Joint tissue</td>
<td>Transcription factors that activate COX-2 expression include C/EBP-beta, phospho- CREB, NF-IL6, AP1, NFKb, TCF-4/LEF-1 and c-Myb.[137]</td>
</tr>
<tr>
<td>NFKB, AP.1</td>
<td>TNF receptor associated factors</td>
<td>Apoptosis in cancer cells</td>
<td>Glucocorticoids inhibit IL1 and cytokine-induced transcription factors (NFKb, AP1) while cytokines enhance glucocorticoid receptor (GR) transcriptional activity.[138]</td>
<td></td>
</tr>
<tr>
<td>NFKB, CMV</td>
<td></td>
<td></td>
<td>NFKB-mediated CMV promoter activation in stably transfected cells can lead to increased target gene levels after transfection instead of siRNA-mediated knockdown.[139]</td>
<td></td>
</tr>
<tr>
<td>NFKB, c-myb, E2F, YY1, AP.1, c-Jun, FosB, MafG, p65, E2F-1</td>
<td></td>
<td></td>
<td>Knockdown of p65, MafG, c-Myb, and E2F-1 lowered basal MAT2A levels. Mutation of AP-1, NFKb, E2F, and USF/c-Myb individually lowered the IGF-1 mediated increase in MAT2A promoter activity (combination mutations fully prevent the increase). YY1 represses MAT2A expression.[140]</td>
<td></td>
</tr>
<tr>
<td>ISRE, Ebox, p53, Oct-1, NFκB</td>
<td>Human mimcan gene</td>
<td>Cellular growth</td>
<td>MG-63 cells-human osteosarcoma</td>
<td>Mimcan is transcriptionally activated by p53 through a conserved intronic recognition site that also contains Ebox, Oct-1, NFκb, and ISRE recognition sites. [141]</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ISRE, Ebox, GAS, IRF-E</td>
<td>MHC class II transactivator (CIITA)</td>
<td>Immune regulation</td>
<td>Human B lymphocytes</td>
<td>Induction of MHC II and CIITA is by IFN-gamma, proteins bind at the GAS, IRF-E and E box sites of CIITA promoter IV. [142]</td>
</tr>
<tr>
<td>ISRE, Ebox, USF-1, IRF, ARE-1, GAS, IRF-1</td>
<td>MHC class II transactivator (CIITA)</td>
<td>Immune regulation / metastatic cells</td>
<td>MDA-MB-435 human breast cancer cell line</td>
<td>Partial or hemi-methylation of promoter IV (has IFNgamma-responsive elements such as GAS, E-box or IRF-1) is sufficient to silence the CIITA expression in highly metastatic cells. This epigenetic mechanism is responsible for the lack of MHC-II expression. [143]</td>
</tr>
<tr>
<td>ISRE, Ebox, STAT1, GAS, IRF-E</td>
<td>MHC class II transactivator (CIITA)</td>
<td>Immune regulation / cancer</td>
<td>Multiple myeloma cells</td>
<td>In vivo protein/DNA binding studies demonstrate protein binding at the GAS, E box and IRF-E sites. In vitro studies confirm the binding of IRF-1 and IRF-2 to CIITA pIV. Although multiple myeloma cells express PRD1-BF1/Blimp-1, a factor that represses both the CIITA type III and IV promoters, they retain the capability to upregulate CIITA pIV and MHC II expression in response to IFN-gamma treatment. [144]</td>
</tr>
<tr>
<td>ISRE, NF-1-like factor, IRF-2, NF-Y</td>
<td>IL-4</td>
<td>Immune response, cytokine release, B cells</td>
<td>Human genome</td>
<td>Mutations of the ISRE site increased overall IL-4 promoter activity twofold. IRF-2 represses INF genes. NF-Y is a transcription activator, NF-1-like factor enhances transcription activated by NF-Y. [145]</td>
</tr>
<tr>
<td>ISRE, AP1, 3, 4 PPARα, c.myb, ETF, Myb(2), p53, Myc-Max, NFκb, STAT1, 3, 4, FAST-1, Pax-5, and others</td>
<td>Dystrophin gene</td>
<td>Dystrophin muscular dystrophy</td>
<td>Diaphragm muscle in mdx mice</td>
<td>TFs upregulated: inflammatory (Stat1,3,4 NF-κB, and AP-1) oxidative stress (PPARα, NF-κB, AP-1, Myc). PAX5 and AP1 may be involved in Satellite cell proliferation or muscle regeneration. P53 plays a critical role in homeostasis of tissues including skeletal muscle. p53, Stat1, Stat4, EVI-1, and lactoferrin BP demonstrated the highest increase. [146]</td>
</tr>
<tr>
<td>ISRE, AP-1, NFκb, CREB, E2F, CETP/CRE, MSP1, Pax6</td>
<td>Cell Proliferation and Apoptosis</td>
<td>HeLa cells</td>
<td>Said TF's activities were regulated by both TNFα- and PMA-induced pathways: GATA, NF-E1, and ISRE. Also PMA only, and E47 TNFα only. PMA and TNFα mediate many biological functions including cell proliferation and apoptosis. [147]</td>
<td></td>
</tr>
<tr>
<td>ISRE, CMV, NFkb,</td>
<td>Immune regulation</td>
<td>Jurkat cells, (T-lymphocyte cell line)</td>
<td>MHC class I regulatory or enhancer elements such as the interferon-stimulated response element (ISRE), NF-kappa B and H2TF1 were not needed for CMV mediated activation of HLA A2 gene promoter.[148]</td>
<td></td>
</tr>
</tbody>
</table>
| ISRE, CMV, IRF-7, HSV-1, TLR7,9 | Type 1 INF genes | INF production | Plasmacytoid dendritic cells (pDC) from cord blood
Profound defect in cord blood capacity to produce IFN-alpha/beta in response to TLR9 as well as to TLR7 ligation or human CMV or HSV-1 exposure.[149] |
<p>| ISRE, E2F, AP-1, CREB, NFkb | Inflammation and ulceration and atrophy | Gastric mucosa of Mongolian gerbils | AP-1 and CREB levels were early responders related to inflammation and ulceration. NF-kB and ISRE were late responders related to atrophy.[150] |
| ISRE, E2F, GAS, Ets, Myb, MyoD, NFkb, NF-IL6, APRF | Mouse RNA-dependent eIF-2 alpha protein kinase (PKR) | Interferon inducibility, growth reg and differentiation. Acute-phase, immune, and inflammatory response genes | 5' flanking region of PKR gene contains binding sites for ISRE+GAS (can induce interferon), Ets, Myb, MyoD, E2F (growth control regulation and differentiation), NFkb, and NF-IL6+APRF (acute phase, immune, and inflammatory response).[151] |
| E2F1 | Interferon regulatory factor 3 (IRF-3) | DNA damage signaling, tumor suppression, and virus-induced apoptosis | E2F1 negatively regulates the basal transcriptional activity of IRF-3 through binding to the E2F consensus binding site and deregulates IRF-3 expression at mRNA level.[152] |
| CLOCK ratio BM AL1, DBP, HLF, E4BP4, CREB, RORa, HSF1, STAT3, SP1 and HNF-4a, PAX-4, C/EBP, EVI-1, IRF, E2F, AP-1, HIF-1 and NF-Y, Clock Controlled Genes (CCGs) | Circadian Rhythm - metabolism, endocrine regulation, pharmacokinetics | Listed TFs and SP1, EGR, ZF5, AP-2, WT1, NRF-1, MEF-2, HMG1Y, HNF-1, OCT-1 are all part of the regulatory network for Circadian Rhythm. E.box, E2F, Spi, AP-2, STAT1, HIF-1 were found to have binding regions in promoter sequences.[41] |
| ISRE, E2F3a, STAT | Epidermal growth factor receptor (EGFR) | Proliferation | Ovarian cancer cells | E2F3a was selectively upregulated following EGFR activation and EGFR-E2F3a axis was found to be STAT1,3 dependent and the ratio of IFN-regulatory factor (IRF)-1 to IRF-2 is determinative for E2F3a control. [153] |
| ISRE, STAT1,3, NFkb, MAPK | IRF1, SOCS1, IL-17 | Graft v host disease, inflammation pathway | Mice: Bone Marrow | GVHD-specific STAT1,3 activation preceded NFkb and MAPK activation and was associated with STAT1,3 dependent inflammatory genes (listed). GVHD was associated with significant expansion of phospho[p]-STAT1 and p-STAT3 expressing CD4(+) T cells and CD8(+) T cells. [154] |
| ISRE-like, STAT1,2, IRF-1, C.EBPb | Interferon regulatory factor-9 (IRF-9) | 2fTGH stably transfected with mutant IFNAR-2c (FF) |  | In mutant cells, IFN-1 led to transcription without activating STAT. IRF-9 activated ISRE-like, IRF-1, and the C.EBP part of the promoter complex. [155] |
| E.Box, NFI, ZEB/AREB6, REB, MRF4 | Na(+) channel gene expression | NA+ channel expression | C2C12 muscle cells | E.Box represses gene expression with ZEB/AREB6, REB. NFI binds up and downstream of E.box preventing the repression. [156] |
| E.Box, NFI, CYP1A2, HNF-1, USF1,2 | Cytochrome P450 1A2 (CYP1A2) regulation | Primary hepatocytes, HepG2 cell |  | On CYP1A2 promoter E.box binds USF1,2, NFI binds DNA-protein complexes. E.box mutations lead to substantial reduction of promoter activity. [157] |
| Ebox, NFI, SP1,3, CRE | Mouse Ren-1(c) gene | Kidney tumor-derived As4.1 cells |  | Renin enhancer contains CRE and E.box, SP1,3, and NFI sites, mutations of the 4 NFI-binding sites resulted in an 89% decrease in expression. [158] |
| C-Myc, p73, E.box | Y-box-binding protein-1 (cold shock domain) | Signaling DNA damage and cell proliferation | Human cancer cell line KB | p73 stimulates the transcription of the YB-1 promoter by enhancing recruitment of the c-Myc-Max complex to the E-box binding site. [159] |
| E.box, NFI | Rpe65 gene | Trans- to 11-cis-retinoid isomerization | Human RPE cell line D407, and non-RPE cell lines Hela, HepG2, and HS27 | Cis-acting elements of Rpe65 gene are an octamer, an NFI site, and two E.box sites. E.box plays a critical role in transcriptional regulation of the Rpe65. [160] |</p>
<table>
<thead>
<tr>
<th>E.box, NF-E2/AP.1, SP1, CREB,</th>
<th>Human HO-1 (heme oxygenase-1) gene</th>
<th>Anti-inflammatory</th>
<th>Renal cells</th>
<th>Sp1 regulates chromatin looping between an intronic enhancer and distal promoter. Mutations of cyclic AMP, NF-E2/AP1 and/or E.box sequences (all downstream) resulted in loss of enhancer activity.[161]</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.box, NF-E2/AP.1, cAMP, CREB, CRE</td>
<td>Human HO-1 (heme oxygenase-1) gene</td>
<td>Anti-inflammatory</td>
<td>HAEC (human aortic endothelial cells), HEK-293 (human embryonic kidney 293) cells, and TF deficient MEF (mouse embryonic fibroblasts)</td>
<td>Regulation of human HO-1 expression by LNO2 (nitrolinoleic acid) requires synergy between CRE, AP-1 and E-box sequences and involves CREB-1.[162]</td>
</tr>
<tr>
<td>E.Box, AP.1, NF-kb, C.EBP, CRE-1,2</td>
<td>Model for COX-2 gene expression during inflammation</td>
<td>Inflammation</td>
<td>RAW264.7 murine macrophages</td>
<td>NFkb, C.EBP, CRE-1 are cis-acting elements in the COX-2 promoter. Ebox mediates transcriptional repression, while CRE2 and AP1 are activating (all cis-acting).[163]</td>
</tr>
<tr>
<td>E.box, E2F3a, MyoD</td>
<td>Cdc6: gives chromatin the capability of replicating DNA</td>
<td>differentiation of muscle stem cells (satellite cells)</td>
<td>C2Cl2 and primary mouse myoblasts</td>
<td>MyoD occupies an E-box binding site. This association, along with E2F3a is required for regulating Cdc6.[164]</td>
</tr>
<tr>
<td>E.box, E2F, NFkb, c-Myb, PAX5</td>
<td>Activation-induced cytidine deaminase (AID) (encoded by Acida)</td>
<td>Antibody memory</td>
<td>Class switch-inducible CH12F3-2 cells</td>
<td>AID can be activated by NFkb, cytokines or other activators, but is has the ubiquitous silencers c-Myb, E2F. Activating sites are Pax5 and Ebox, but AID activation is a balance between them. AID is essential for the generation of antibody memory but also targets oncogenes (tumorogenesis).[165]</td>
</tr>
<tr>
<td>E.box, E2F, Myc</td>
<td>Pole3 (subunit of DNA polymerase e)</td>
<td>Cell cycle regulation</td>
<td>Pole 3 promoter is linked bi-directionally to a gene of unknown function. This promoter is regulated by E2F and Myc.[166]</td>
<td></td>
</tr>
<tr>
<td>TF3 = E.box factor, E2F</td>
<td>p68 subunit gene of DNA polymerase alpha</td>
<td>DNA replication, cell cycle progression, and cell fate determination</td>
<td>Combinatorial gene control (few TFs regulating many genes) involving E2F and E.Box family members: E2F target genes share a common promoter architecture.[167]</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>TF3 = E.box factor, E2F</td>
<td>p53 gene</td>
<td>HBV-mediated hepatocellular carcinogenesis (HCC)</td>
<td>Interaction of TF3 and E2F facilitates transcriptional activation of the p68 gene and provides strong evidence for the specificity of E2F function.[168]</td>
<td></td>
</tr>
<tr>
<td>E.box, E2F1, HBx</td>
<td>p53 gene</td>
<td>HepG2 and HeLa cells</td>
<td>E2F1 overcomes the repressive effect of hepatitis B viral X protein (HBx) on the p53 promoter to activate it. HBx repressed the p53 promoter through E-box.[169]</td>
<td></td>
</tr>
<tr>
<td>E.box, E2F, GATA-1</td>
<td>Follicle-stimulating hormone receptor (FSHR) gene</td>
<td>Sertoli (m) and granulosa (f) cells</td>
<td>There are E.box, GATA, and E2F sites in the promoter of the FSHR gene. Mutations to any of the sites eliminated nearly all activity.[170]</td>
<td></td>
</tr>
<tr>
<td>E.box, STAT3, C.EBP, BMP/SMAD4</td>
<td>Hepcidin</td>
<td>HuH7 hepatoma cells</td>
<td>STAT3, C.EBP and E-BOX were mutated to look at levels of Hepcidin. Hypoxia represses hepcidin expression through inhibition of BMP/SMAD signaling.[171]</td>
<td></td>
</tr>
<tr>
<td>E.box, STAT3, nescent helix-loop-helix 2 (Nhlih2)</td>
<td>Pro-hormone convertase 1/3 (PC1/3)</td>
<td>Janus kinase/STAT signaling</td>
<td>N2KO mice, and hypothalamic cell line</td>
<td>Pro-opiominocortin is cleaved by pro-hormone convertase 1/3 (PC1/3) to produce peptides that regulate the body's response to energy availability. Nhlih2 binds to E-box motifs found adjacent to STAT3 sites in the PC1/3 promoter. Nhlih2 is an integral element of the Janus kinase/STAT signaling pathway and the first to demonstrate coordinated control of PC1/3 transcription by Nhlih2 and STAT3 after leptin stimulation.[172]</td>
</tr>
<tr>
<td>NFI, AP.1, SP.1</td>
<td>Human alpha5 integrin gene</td>
<td>Cellular repair</td>
<td>Corneal epithelium</td>
<td>AP-1 and Sp1 activated expression is directed by the alpha5 promoter. NFI functioned as a potent repressor.[173]</td>
</tr>
<tr>
<td>NFI, AP.1, c-jun (TAM67)</td>
<td>Alpha1-antichymotrypsin (ACT)</td>
<td>Astrocyte-specific expression</td>
<td>Astrocytes and glioma cells</td>
<td>NFI-X cooperates with AP-1 by an unknown mechanism in astrocytes, which results in the expression of a subset of astrocyte-specific genes.[174]</td>
</tr>
<tr>
<td>NFI, E2F, c-Jun</td>
<td>Op18 (Oncoprotein 18, Statamin)</td>
<td>Mitotic regulation</td>
<td>Metastatic prostate carcinoma cell line PC-3-M</td>
<td>Op18 is highly expressed in many cancers. Regulation is through 4 E2F sites - sites 2 and 3 have FNI binding sites. Mutations of these sites show they are responsible for 80% of Op18 expression.[175]</td>
</tr>
<tr>
<td>Wt1, SP1, CTF, VP16, p53, E2F1</td>
<td>Wilm's tumor suppressor gene (wt1)</td>
<td>Activation domains</td>
<td>Wt1 represses p53 and E2F1, which stimulate both initiation and elongation via RNA Polymerase II transcription (type IIb).[176]</td>
<td></td>
</tr>
<tr>
<td>NFI, E2F, SP1</td>
<td>B-subunit of human DNA polymerase epsilon (POLE2)</td>
<td>DNA replication</td>
<td>Human DNA</td>
<td>POLE2 is regulated by two E2F-pocket protein complexes, one associated with Sp1 and the other with NFI.[177]</td>
</tr>
<tr>
<td>AP.1, DR4,5-TRAIL receptors</td>
<td></td>
<td>Apoptosis</td>
<td>LNCaP prostate cancer cells</td>
<td>Induction of AP.1 by TPA promotes a change in phenotype of some TRAIL-resistant prostate cancer cells to a TRAIL-sensitive phenotype. AP.1 promotes tumor necrosis factor (TNF)-related apoptosis which induces ligand (TRAIL)-induced apoptosis by repressing the anti-apoptotic molecule c-FLIP(L).[178]</td>
</tr>
<tr>
<td>AP.1, DR5, NFkb</td>
<td>Apoptotic cell death, pro-inflammatory and angiogenic signals in human brain tumors</td>
<td>Two human glioma cell lines, CRT-MG and U87-MG</td>
<td>DRS ligation on human glioma cells leads to apoptosis and the activation of AP-1 and NFkB which leads to the induction of IL-8 expression (responses dependant on caspase activation).[179]</td>
<td></td>
</tr>
<tr>
<td>AP.1, DR4,5, NFkb</td>
<td>p53 signaling</td>
<td>Apoptosis following drug-induced DNA damage</td>
<td>Chemotherapeutic agents activate NFkB and AP.1, which then directly transactivate Fasl. Death receptors for Fasl (Fas) and TRAIL (DR4,5) are emerging as important regulators of drug-induced apoptosis in human cancers, mediated by caspase activation.[180]</td>
<td></td>
</tr>
<tr>
<td>C.EBP, AP.1, E.box, HNF3, GRE, NFI, IRE</td>
<td>Rat cytosolic aspartate aminotransferase gene (cAspAT)</td>
<td>Insulin response</td>
<td>H4IEC3 hepatoma cells</td>
<td>Insulin effect is due to a complex interplay of factors requiring the synergistic contribution of at least two IRE sites which underlines the contribution of HNF-3 and NFI-like proteins. Stimulation by glucocorticoids and repression by insulin in the liver.[181]</td>
</tr>
<tr>
<td>AP.1, E2F, NFkb, STAT1</td>
<td>Graft arterial disease (GAD) formation</td>
<td>Inflammation (organ rejection)</td>
<td>Rodent and non-human primate Heart Transplant</td>
<td>E2F decays (bind to the TFs to prevent functioning) blocked cell cycle regulatory elements. NFkb decay attenuates chronic (GAD) and acute rejections in cardiac allografts. STAT1 and AP.1 regulate vascular adhesion molecule expression.[182]</td>
</tr>
<tr>
<td>AP.1, E2F1,2</td>
<td>Cell cycle genes (cyclins E, A, B and D3, G2, and I)</td>
<td>Cell Cycle</td>
<td>MCF-7 breast cancer cells</td>
<td>AP.1 regulates the expression of cyclin D and E2F (the latter regulates E2F-downstream genes), leading to cell cycle progression and breast cancer cell proliferation.[183]</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------</td>
<td>------------</td>
<td>--------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>AP.1, E2F</td>
<td>CAGE (cancer/testis antigen), Cyclin D1, E</td>
<td>Cell Cycle</td>
<td>HeLa transfectant clones with the tetracycline-inducible CAGE gene</td>
<td>CAGE stimulates cell proliferation by upregulating cyclins D1 and E in an AP.1 and E2F dependant manner.[184]</td>
</tr>
<tr>
<td>AP1,2, E2F1, MTF1, SREBP, SP1, PPAR, NFκb, HNF4</td>
<td>TGFA, which protects surfactant-associated protein B (SFTPB)</td>
<td>Cell damage/protection</td>
<td>Mouse lung epithelial (MLE-15) cells</td>
<td>Nickel exposure increased expression of genes with promoter binding sites for E2F1, AP.2, SREBP, SP1, PPAR, NFκb, and HNF4. Genes with AP.1 sites had decreased transcripts. Maintenance of SFTPB expression is critical to survival during acute lung injury.[185]</td>
</tr>
<tr>
<td>AP.1, E2F, SP3, NF1, c-Myb, B-Myb</td>
<td>Proximal promoter of transcription factor Sp3</td>
<td>SP3 regulation</td>
<td>HELa cells</td>
<td>SP3 regulation depends in part on the relative amounts of Sp1,3, NF-Y,1, c-Myb, B-Myb, AP.1, and E2F proteins in the cell.[186]</td>
</tr>
<tr>
<td>AP.1, STAT3, NFκb</td>
<td>Inflammation in cancer</td>
<td></td>
<td>NFKB, AP-1 and STAT3 and their gene products provide the molecular basis for the role of inflammation in cancer.[47]</td>
<td></td>
</tr>
<tr>
<td>AP.1, STAT3, p65/NFκb</td>
<td>Thyroid hormone receptors (TRs)</td>
<td>Proliferation and inflammation</td>
<td>Thyroid hormone binding isoforms TRα1 and TRβ (KO mice)</td>
<td>TRs act as endogenous inhibitors of skin inflammation, most likely due to interference with AP-1, NF-kB and STAT3 activation.[187]</td>
</tr>
<tr>
<td>AP.1, STAT3, NFκb, C.EBP</td>
<td>Nitric oxide synthase (NOS) and cyclooxygenase-2 (COX-2)</td>
<td>Inflammation</td>
<td>(LPS)-stimulated macrophages (RAW264.7 cells)</td>
<td>Andrographolide downregulates inflammatory iNOS and COX-2 gene expression by inhibiting the activation of NFKB and STAT3 by interfering with the expression of SOCS1 and SOCS3 signaling.[188]</td>
</tr>
<tr>
<td>AP.1, STAT3</td>
<td>Matrix metalloproteinase-1 (MMP-1)</td>
<td>Carcinogenesis</td>
<td>Colon Carcinoma Cells</td>
<td>MMP-1 promoter has necessary binding sites for STAT3 and AP.1. Aberrant STAT3, AP.1 activity drives MMP-1 expression and tissue invasion in colorectal cancer.[189]</td>
</tr>
<tr>
<td>AP.1, STAT3/Jak2, NFκb</td>
<td>IL6 autocrine production</td>
<td>Carcinogenesis</td>
<td>Lung adenocarcinoma PC14PE6/AS2 (AS2) cells</td>
<td>IL6 can be regulated through downstream pathways by NFκb and AP.1, but autocrine production of IL6 occurs through a positive feedback loop with STAT3.[190]</td>
</tr>
<tr>
<td>AP.1, STAT3, c-Jun</td>
<td>Catecholamine stimulation on MMP-7 expression</td>
<td>Metastasis</td>
<td>Gastric Cancer Cells</td>
<td>Up-regulation of MMP-7 expression through β2-AR-mediated signaling pathway is involved in invasion and metastasis of gastric cancer. MMP-7 promoter has STAT3 and AP.1 sites.[191]</td>
</tr>
<tr>
<td>CMV, DRS, YY1</td>
<td>Raf-1 kinase inhibitor protein (RKIP)</td>
<td>TRAIL induced Apoptosis</td>
<td>TRAIL-resistant prostate carcinoma PC-3 and melanoma M202 cell lines</td>
<td>RKIP over-expression regulates sensitivity to TRAIL via inhibition of YY1 and upregulation of DR5 (low RKIP expression = pro-oncogenic). NFκb regulates tumor resistance against apoptosis by the TNFα family via inactivation of the transcription repressor Yin Yang 1 (YY1).[192]</td>
</tr>
<tr>
<td>CMV, p-STAT3, VEGF, TGF-beta, viral IE1 and pp65</td>
<td>CMV IL10</td>
<td>CMV mediated response</td>
<td>Human monocytes and Glioblastoma multiform (GBM) tumors</td>
<td>CMV IL10 induces viral IE1 expression. CMV IL-10-treated monocytes produced angiogenic VEGF, immunosuppressive TGF-beta, and enhanced migration of gCSCs. CMV triggers a feed-forward mechanism of gliomagenesis by inducing tumor-supportive monocytes.[193]</td>
</tr>
<tr>
<td>CMV, STAT3</td>
<td>CMV IL10, SOCS3</td>
<td>Immune response</td>
<td>HeLa Cells</td>
<td>IL-10R1 variants differentially reduce the signaling activity of CMVIL-10 and thereby may affect CMV's ability to escape from the host's immune surveillance. CMVIL-10 production resulted in increased STAT3 production.[194]</td>
</tr>
<tr>
<td>CMVirus, STAT3</td>
<td>hIL10 production</td>
<td>Immune response</td>
<td></td>
<td>CMVIL-10 stimulated autocrine production of hIL-10 by B lymphocytes and led to activation of the latent transcription factor Stat3.[195]</td>
</tr>
<tr>
<td>E2F1, STAT3, p53</td>
<td>Suppressor of cytokine signalling (SOCS)</td>
<td>Apoptosis</td>
<td>Socs3(-/-) mammary epithelial cells</td>
<td>c-myc regulates apoptosis during involution by augmented Stat3 activation: likely targets are E2F-1, Bax and p53 genes. [196]</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------------------------</td>
<td>-----------</td>
<td>-------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>E2F, STAT3</td>
<td>Cyclin D1</td>
<td>Oncogenesis</td>
<td>Primary breast tumors and breast cancer-derived cell lines, rat/mouse/human derived cell lines</td>
<td>Cyclin D1 is transcriptionally regulated and required for transformation by STAT3. Constitutively activated STAT3 is oncogenic. [197]</td>
</tr>
<tr>
<td>E2F, STAT3,5A, NFkb, Cdk9, IL-1b,6,10, YY1, MMP-7,12,14, AP1</td>
<td>HIV replication</td>
<td>Neonatal (cord) blood mononuclear cells (CBMC) compared with adult blood cells (PBMC)</td>
<td>TFs: NFkb, E2F, Cdk9. Signal transducers: STAT3,5A. And cytokines: IL-1b,6,10 were more upregulated in CBMC than PBMC; which is known to influence HIV-1 replication. YY1 was further downregulated in CBMC than PBMC. MMB7,12,14 was further upregulated in HIV-1 infected CBMC than PBMC. [198]</td>
<td></td>
</tr>
<tr>
<td>E2F-DP1, STAT3, SP1, Rb, p53, NFkb</td>
<td>Genes regulated by SP1 promoter</td>
<td>Cell Cycle</td>
<td>HT1080 cells</td>
<td>SP1 affects and is affected by multiple TFs: CDK4, SKP2, Rad51, BRCA2 and p21, E2F-DP1, Cyclin D1, STAT3, Rb. NFkb and p53 inhibit Sp1. Sp1 may be a key mediator of cell cycle associated changes in gene expression. [199]</td>
</tr>
<tr>
<td>SP1, p53, Hypoxia inducible factor (HIF)</td>
<td>p21, p27</td>
<td>Cell cycle</td>
<td>Normoxic HIFa knockout cells</td>
<td>HIFα deletion leads to p21 induction via, at least in part, p53-independent but SP1-dependent mechanisms. Knockdown of HIF-1α results in a significant increase in cells in the G1 phase of the cell cycle. [200]</td>
</tr>
<tr>
<td>SP1, p53, NF-Y</td>
<td>G2/M cell cycle progression/tumorigenesis</td>
<td>Cell cycle</td>
<td>HT1080 cells</td>
<td>Cell cycle proteins Cdc25B phosphatases are downregulated by p53 and require SP1/SP3 and NF-Y binding sites. Over-expression of Cdc25B results in bypass of the G2/M checkpoint and illegitimate entry into mitosis, and also causes replicative stress, leading to genomic instability. [201]</td>
</tr>
<tr>
<td>SP1, p53, STAT3</td>
<td>DNA damage</td>
<td>Co-treated HepG2 cells with a rat liver S9 fraction as metabolic activation system (MAS)</td>
<td>Genotoxicity determined through differential regulation of STAT3, p53, SP1. TFs play roles in stress and DNA damage response.[202]</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>SP1, p53, c-Myc, HNF4A, ESR1</td>
<td>Limb regeneration</td>
<td></td>
<td>c-Myc, SP1 and their target genes could potentially play a central role in limb regeneration. Other TFs known to be involved in epigenetic reprogramming such as Klf4, Oct4, and Lin28 are also connected to c-Myc and SP1.[203]</td>
<td></td>
</tr>
<tr>
<td>SP1,2, PPRE, CP, Xp</td>
<td>Hepatitis B virus (HBV) gene</td>
<td>HCC, HeLa, 293T, HepA2 cells</td>
<td>HE-145 suppressed HBV expression and replication in HCC cells (liver cells) via suppression of SP2 and CP. Ectopic expression of PPRE and HNF4 reverse the repression.[204]</td>
<td></td>
</tr>
<tr>
<td>Ahr, PXR, CAR, Nrf2, PPARy, C.EBPb, NFkb, EGR1</td>
<td>Glutathione Transferase (GST) genes</td>
<td>Xenobiotic response, inflammation</td>
<td>Family of compounds up-regulate GSTs through CAR, PXR, Ahr, Nrf2, PPARy, C.EBPb. Induction of some genes (e.g. TNFα) is mediated by C.EBP, NFkb, EGR.1.[205]</td>
<td></td>
</tr>
<tr>
<td>Myb, GLI, CCND1, MDM2, CDK4, ESR1, and AIB1</td>
<td>Oncogenes</td>
<td>Pleomorphic liposarcomas</td>
<td>CCND1, MDM2, GLI, CDK4, MYB, ESR1, and AIB1 are oncogenes.[206]</td>
<td></td>
</tr>
<tr>
<td>c-myb, GLI, GATA3, ETS, Erg</td>
<td>T cell development</td>
<td></td>
<td>T Cell development involves GATA3, c-Myb, TCF.1, Ets, Gli, among others.[207]</td>
<td></td>
</tr>
<tr>
<td>c-Myb, ETS, GR, C.EBP</td>
<td>697 pre-B-ALL and CEM-C7 T-ALL cells</td>
<td></td>
<td>Glucocorticoid response units bind c-Myb and Ets.[208]</td>
<td></td>
</tr>
<tr>
<td>Myb, ETS (Etv4,5)</td>
<td>Cxcr4, Myb, Met and Mmp14</td>
<td>Kidney development</td>
<td>Expression of Cxcr4, Myb, Met, and Mmp14 depends on Etv4,5. Etv4,5 is positively regulated by Ret signaling. Etv4 and Etv5 are key components of a gene network downstream of Ret that promotes and controls renal branching morphogenesis.[209]</td>
<td></td>
</tr>
<tr>
<td>RTU/TF Associations Indicated by RTU-RTU Correlation Represented in Current Literature</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myb, ETS, AML1, GATA, Sry</td>
<td>PSF1</td>
<td>Spermatogenesis</td>
<td>Mouse testis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c-Myb, ETS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucocorticoid (GC) hormones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hormone-sensitive ALL cell lines</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucocorticoid receptor units have recognition sequences for c-Myb, GR, and Ets. There is a spectrum of TFs present in different types of ALL which might allow for tailored therapy. [211]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c-myb, ETS-1, SP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilms' tumor 1 gene (WT1), histone deacetylases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene regulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histone acetyltransferase p300 was able to counteract the HDAC4/HDAC5-mediated repression. p300/CREB synergized with transcription factors Sp1, c-Myb, and Ets-1 in activation of the WT1 reporter. [212]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v-Myb, ETS-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v-Myb oncogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macrophage differentiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v-Myb has been intensely studied as a transcriptional activator but v-Myb can repress biologically relevant genes such as Ets-2. Ets-2 promotes macrophage differentiation. [213]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c-Myb, c-ETS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human glucocorticoid receptor (hGR) 1A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lymphoblast cells (T(CEM-C7+Jurkat)+B(ML-9))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 DNA elements in the hGR-1A promoter (footprints 11 and 12) regulate the gene via a molecular switch model as the direction of auto-regulation is determined by which TF binds footprint 12 and when hGR binds footprint 11. [214]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c-Myb, ETS, Notch, GATA3, E2A/HEB, TCF-1, Runx, Ikaros</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALDH3A1, NGN3, etc</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene expression pattern</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local and segmental differences in the expression of liver-enriched transcription factors in the human intestine which impact epithelial cell biology of the gut. [215]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNF6,4a, GATA4, FOXA2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clusterin (stress response protein)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell proliferation, tissue differentiation, inflammation, lipid transport</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clusterin gene has glucose response element (GiRE) consisting of 2 E.box motifs activated by SREBP1c. Glucose induces SREBP1c recruitment. This metabolically regulates clusterin. [217]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPAR, LXRa, SREBP1c, Wnt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcoholic and non-alcoholic steatohepatitis (ASH and NASH)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model: Mouse intragastric (hepatic stellate cells)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enhanced lipogenic regulation is frequently encountered as characterized by induction of lipogenic or adipogenic transcription factors (PPAR, LXRa, SREBP1c). Wnt signaling is implicated in 'anti-adipogenic' stellate cell transdifferentiation in liver fibrogenesis. [218]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RTU/TF Associations Indicated by RTU-RTU Correlation Represented in Current Literature
References

8. Krupczak-Hollis K Fau - Wang, X., et al., The mouse Forkhead Box m1 transcription factor is essential for hepatoblast mitosis and development of intrahepatic bile ducts and vessels during liver morphogenesis. (0012-1606 (Print)).


110. Hendricks-Taylor, L.R. and G.J. Darlington, Inhibition of cell proliferation by C/EBP alpha occurs in many cell types, does not require the presence of p53 or Rb, and is not affected by large T-antigen. Nucleic Acids Research, 1995. 23(22): p. 4726-4733.


118. Hirot a, T., et al., Transcriptional repressor TIEG1 regulates Bmal1 gene through GC box and controls circadian clockwork. Genes to cells : devoted to molecular & cellular mechanisms, 2010(Journal Article).


189. Zugowski, C., et al., STAT3 controls matrix metalloproteinase-1 expression in colon carcinoma cells by both direct and AP-1-mediated interaction with the MMP-1 promoter. Biological chemistry, 2011(Journal Article).


