

Electronic Thesis and Dissertation Repository

3-12-2021 11:00 AM

Characterizing the Role of TDG in FXR-dependent Signaling

Oladapo A. Onabote, *The University of Western Ontario*

Supervisor: Torchia, Joseph, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biochemistry

© Oladapo A. Onabote 2021

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Medical Biochemistry Commons](#)

Recommended Citation

Onabote, Oladapo A., "Characterizing the Role of TDG in FXR-dependent Signaling" (2021). *Electronic Thesis and Dissertation Repository*. 7739.

<https://ir.lib.uwo.ca/etd/7739>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Thymine DNA Glycosylase (TDG) plays a key role in active demethylation by excising intermediates of 5-methylcytosine. The function of TDG is required for embryonic development, as *Tdg*-null embryos die at E11.5. To bypass this embryonic lethality, our lab generated conditional *Tdg* knockout (TDG_{CKO}) mice. These mice develop late-onset hepatocellular carcinoma (HCC), partly due to impaired Farnesoid X Receptor (FXR) signaling. Interestingly, *Fxr*-knockout mice display a similar phenotype and transcriptional profile to TDG_{CKO} mice, prompting us to investigate a role for TDG in FXR signaling. To this end, we generated *Tdg/Fxr* double-knockout (DKO) mice. We also generated a novel *Fxr*-null mouse model using CRISPR/Cas9, which facilitated the knockout of FXR through a 47-bp deletion event. We demonstrated that 3-week-old *Fxr*-null mice display impaired bile acid and glucose metabolism. Moreover, we demonstrated a novel interaction between TDG and FXR *in vivo*. Collectively, these findings implicate TDG as a coactivator of FXR signaling.

Keywords

Cancer, Epigenetics, Thymine DNA Glycosylase, Knockout Mice, Farnesoid X Receptor, CRISPR

Summary for Lay Audience

DNA can be modified by a process known as methylation. This modification can be reversed by a counteracting process known as active demethylation. A key protein involved in active demethylation is Thymine DNA Glycosylase (TDG). Mouse studies have demonstrated that TDG is required for embryonic development. When TDG was deleted from birth in mouse embryos, these embryos died twelve days post-conception.

Since the deletion of TDG from birth is lethal, our lab deleted TDG eight weeks after birth to bypass this obstacle. This is known as a ‘conditional’ deletion; hence these mice are called conditional TDG-knockout (TDG_{CKO}) mice. Our lab found that TDG_{CKO} mice develop late-onset liver cancer, partly due to an increase in bile acids (BAs). Excessive amounts of BAs can cause damage to the liver. Consequently, BAs are tightly regulated. The main protein involved in regulating BAs is Farnesoid X Receptor (FXR). Interestingly, FXR-knockout (FXR_{KO}) mice develop late-onset liver cancer, which is similar to TDG_{CKO} mice. To this end, I aimed to generate TDG/FXR double-knockout (DKO) mice by breeding FXR_{KO} and TDG_{CKO} mice together, predicting that DKO mice will develop a more accelerated form of liver cancer. However, to generate DKO mice in this manner, the genes for TDG and FXR would need to be on separate chromosomes. Incidentally, TDG and FXR are on the same chromosome. Consequently, we used a gene-editing technique that allowed us to bypass this hindrance. With this technique, we generated a new FXR_{KO} model which enabled us to subsequently generate DKO mice. This technique functions by introducing various mutations into a gene of interest. I showed that the specific mutation that occurred in our FXR_{KO} mice was a deletion of 47 base pairs. I found that our FXR_{KO} mice have significantly more BAs in the liver compared to normal mice. Collectively, the preliminary data generated from our FXR_{KO} model is consistent with the published data from previous FXR_{KO} models.

Finally, I investigated whether TDG plays a role in FXR function. I demonstrated that TDG and FXR interact in mouse liver. Altogether, my results suggest that TDG plays a coactivating role in FXR function.

Co-Authorship Statement

The generation and RNASeq analysis of TDG_{CKO} mice was done by Haider Hassan and Majdina Iovic. Mouse breeding and tissue isolation were performed by Majdina Iovic. Co-immunoprecipitation experiments were performed by Haider Hassan.

Acknowledgments

Firstly, I would like to thank Dr. Joseph Torchia for this research opportunity and all for all the support that he has provided me during this time. I would also like to thank my committee members Dr. Trevor Shepherd and Dr. Fred Dick for their valuable input on this project. Special thanks to Dr. Fred Dick for collaborating with our lab in the generation of the *Fxr*-null mice. As for the lab, I'm truly grateful to Haider Hassan and Majdina Iovic for helping me every step of the way in my research experience. I could not have gotten this far without your help and I really appreciate the time that you have contributed for my sake.

Table of Contents

Abstract.....	ii
Summary for Lay Audience.....	iii
Co-Authorship Statement.....	iv
Acknowledgments.....	v
Table of Contents.....	vi
List of Tables.....	viii
List of Figures.....	ix
List of Abbreviations.....	xi
1 Introduction.....	1
1.1 Mammalian gene regulation.....	1
1.1.1 Gene expression and regulation.....	1
1.1.2 Epigenetic regulation of gene expression.....	3
1.1.3 DNA methylation.....	4
1.2 DNA demethylation.....	8
1.2.1 Active demethylation.....	8
1.2.2 Thymine DNA Glycosylase.....	16
1.2.3 Phenotypic effects of TDG deletion.....	20
1.3 Hepatic metabolism.....	21
1.3.1 Liver metabolism.....	21
1.3.2 FXR.....	24
1.3.3 Hepatocellular carcinoma.....	31
1.4 Rationale and hypothesis.....	33
2 Methods.....	37

2.1	Generation of TDG _{CKO} , <i>Fxr</i> -null, and <i>Tdg/Fxr</i> DKO mice	37
2.2	RT-PCR/RT-qPCR	38
2.3	Protein extraction and western blot.....	38
2.4	Immunoprecipitation.....	39
2.5	RNASeq analysis	39
2.6	CRISPR/Cas9 genome editing.....	40
2.7	Total bile acid analysis.....	40
2.8	Primer sequences	41
3	Results	43
3.1	Determining the molecular role of TDG in FXR signaling	43
3.2	CRISPR/CAS9-mediated generation of mutant <i>Fxr</i> allele	51
3.3	Generation of <i>Fxr</i> -null mice	59
3.4	Generation of <i>Tdg/Fxr</i> DKO mice.....	66
4	Discussion	71
4.1	Overview.....	71
4.2	TDG's role as a coactivator of FXR signaling.....	71
4.3	Alteration of the <i>Fxr</i> allele using CRISPR/Cas9	72
4.4	Comparison of previous <i>Fxr</i> -null mouse models.....	73
4.5	Cre-mediated knockout of TDG in <i>Tdg/Fxr</i> DKO mice.....	74
4.6	Summary and future directions.....	74
	Bibliography	76
	Curriculum Vitae	84

List of Tables

Table 2.1 Primers for RT-PCR	41
Table 2.2 Primers for RT-qPCR	41
Table 2.3 Primers for genotyping	41
Table 2.4 gRNA oligos	42

List of Figures

Figure 1.1. Transcriptional regulatory elements	6
Figure 1.2. Active demethylation pathways	12
Figure 1.3. Functions of oxidized 5mC derivatives.....	14
Figure 1.4. TDG structural domains	18
Figure 1.5. Bile acid synthesis and transport	22
Figure 1.6 FXR structural domains and isoforms.....	26
Figure 1.7. FXR signaling pathways.....	28
Figure 1.8 Experimental approach behind generation of <i>Fxr</i> -null mice.....	35
Figure 3.1 Transcriptional overlap between TDG _{CKO} and FXR _{KO} mice	45
Figure 3.2 qPCR analysis of TDG _{CKO} mice.....	47
Figure 3.3 Co-immunoprecipitation of TDG and FXR	49
Figure 3.4 gRNA-mediated targeting of <i>Fxr</i>	53
Figure 3.5 Generation of five founder mice.....	55
Figure 3.6 A 47-bp deletion event at exon 5 of FXR.....	57
Figure 3.7 Western blot analysis of <i>Fxr</i> -null mice	60
Figure 3.8 qPCR analysis of <i>Fxr</i> -null mice	62
Figure 3.9 Hepatic bile acid levels of <i>Fxr</i> -null mice	64
Figure 3.10 Breeding scheme for <i>Tdg/Fxr</i> double knockout (<i>Fxr</i> ^{-/-} <i>Tdg</i> ^{fl/fl} <i>UBC-cre/ERT2</i> ^{+/-}) mice.....	67

Figure 3.11 Generation of *Tdg/Fxr* double knockout mice 69

List of Abbreviations

2OG	2-oxoglutarate
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxymethyluracil
5mC	5-methylcytosine
AID/APOBEC	Activation-induced deaminase/apolipoprotein B RNA-editing catalytic component
AP	Apurinic/aprimidinic
APE1	AP endonuclease 1
AR	Androgen receptor
ASBT	Apical sodium-dependent bile acid transporter
BA	Bile acid
BER	Base excision repair
BS ³	Bissulfosuccinimidyl suberate
BSEP	Bile salt export pump
CBP/p300	CREB-binding protein
CDCA	Chenodeoxycholic acid
CGI	CpG island
CpG	Cytosine-guanine
CRE	cAMP response element
CREB	cAMP response element binding protein
CRISPR	Clustered regularly interspaced short palindromic repeats
CYP7A1	Cholesterol 7 α -hydroxylase
DKO	Double knockout
DNMT	DNA Methyltransferase
DSBH	Double-stranded β -helix
EDTA	Ethylenediaminetetraacetic acid
ER α	Estrogen receptor α
ESC	Embryonic stem cell
FBP1	Fructose 1,6-bis phosphatase
FGF15/FGF19	Fibroblast growth factor 15/19
FGFR4	Fibroblast growth factor receptor 4
FXR	Farnesoid X receptor
FXRE	FXR response element
FXR _{KO}	FXR-knockout
G6Pase	Glucose-6-phosphatase
gRNA	guide RNA
GTF	General transcription factor
H3K14ac	Histone 3 lysine 14 acetylation
H3K27me3	Histone 3 lysine 27 trimethylation
H3K4me3	Histone 3 lysine 4 trimethylation
H3K9me3	Histone 3 lysine 9 trimethylation
HAT	Histone acetyltransferase

HCC	Hepatocellular carcinoma
HDAC	Histone deacetylase
HNF4	Hepatocyte nuclear factor
IFN γ	Interferon- γ
IGF	Insulin-like growth factor
INR	Initiator element
IVF	In-vitro fertilization
KMT2A	Lysine methyltransferase 2A
LCR	Locus control regions
LRH-1	Liver receptor homologue-1
MBD4	Methyl CpG binding domain 4
MBP	mCpG binding protein
mCpG	methyl CpG
MEF	Mouse embryonic fibroblast
MUG	Mismatch Uracil DNA Glycosylase
NCOA	Nuclear receptor coactivator
NHEJ	Non-homologous end joining
NR	Nuclear receptor
NTCP	Sodium taurocholate cotransporting polypeptide
OATP	Organic anion transporting polypeptide
OST α/β	Organic solute transporters
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEPCK	Phosphoenol-pyruvate carboxykinase
PIC	Pre-initiation complex
PKC α	Protein kinase C α
PVDF	Polyvinylidene difluoride
qPCR	quantitative PCR
RA	Retinoic acid
RAR	Retinoic acid receptor
RARE	RA response element
RNAP	RNA polymerase
RT-PCR	Reverse transcription PCR
RXR	Retinoid X receptor
SAM	S-adenosylmethionine
SBM	SUMO binding motif
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHP	Small heterodimer partner
SMUG1	Single-Strand-Selective Monofunctional Uracil-DNA Glycosylase 1
SRC	Steroid receptor coactivator
SUMO	Small Ubiquitin-like Modifiers
T2D	Type 2 diabetes
TAM	Tamoxifen
TBP	TATA binding protein
TDG	Thymine DNA glycosylase

TDG _{CKO}	Conditional TDG knockout
TET	Ten-eleven translocation
TGF β 1	Transforming growth factor β 1
TSS	Transcription start site
TTF1	Thyroid transcription factor 1
UBC	Ubiquitin-C
UDG	Uracil DNA glycosylase
UHRF1	Ubiquitin-like containing PHD and RING finger domains 1
UNG 1/2	Uracil N-glycosylase 1/2
VEGF	Vascular endothelial growth factor

1 Introduction

1.1 Mammalian gene regulation

Cells have mechanisms to control the expression of genes. At the transcriptional level, there are many regulatory elements in DNA that influence gene expression (e.g. promoters, enhancers, and silencers). At the epigenetic level, gene regulation is achieved through chromatin remodelling. Two well-established mechanisms involved in epigenetic gene regulation are histone modifications and DNA methylation.

1.1.1 Gene expression and regulation

Gene expression is the process by which the DNA sequence of a gene is converted into a functional gene product, which is normally a protein. This process is further subdivided into two stages known as transcription and translation. During transcription, the DNA sequence is used as a template to create a premature RNA transcript through the function of RNA Polymerases (RNAPs) (Cramer, 2019). This transcript is then processed into a mature RNA transcript (e.g. mRNA, tRNAs, rRNAs, etc.) that functions according to its genetic sequence. If the transcript is an mRNA, it will proceed to the next stage of gene expression which is translation. During translation, the mRNA sequence is decoded by the ribosome to produce a protein. Considering that the human genome contains approximately 20,000 protein-coding genes, this process requires strict regulation in order to conserve energy, ensuring that genes are active in a time and spatially dependent fashion (Guo, 2014).

At the transcriptional level, genes contain regulatory elements which help to modulate their expression (Figure 1.1). A major regulatory element in a gene is the promoter, which is the region that often signifies the start of the gene (Cramer, 2019). The promoter region usually contains a core promoter, a short DNA sequence (~100 bp) where the core transcriptional machinery (e.g. RNAP) binds to initiate transcription (Andersson et al., 2015; Haberle and Stark, 2018; Andersson and Sandelin, 2020). The core promoter possesses key elements for transcription initiation such as the transcription start site (TSS), the TATA box, and the initiator element (INR) (Andersson et al., 2015;

Haberle and Stark, 2018; Andersson and Sandelin, 2020). Other regulatory elements are classified as ‘proximal’ or ‘distal’, relative to the core promoter. Distal elements—which can be up to 1 million bp away from the core promoter— include enhancers (elements that promote transcription), silencers (elements that repress transcription), insulators (elements that prevent enhancer activation of neighbouring genes), and locus control regions (LCR; elements that enhance expression of linked genes) (Maston et al., 2006; Andersson and Sandelin, 2020). Proximal promoter elements —typically within 200 bp upstream of the TSS— consist of response elements, which are consensus sequences that act as binding sites for specific transcription factors (Haberle and Stark, 2018).

Regulatory elements function through the binding of transcription factors, which can be broadly classified into two categories: general transcription factors (GTFs) and specific transcription factors. GTFs (e.g. TFIIA, TFIIB, TFIID) bind to regulatory elements within the core promoter and form a pre-initiation complex (PIC) with RNAPII (Andersson et al., 2015; Haberle and Stark, 2018; Andersson and Sandelin, 2020). For example, the TATA-binding protein (TBP), a subunit of TFIID, is a general transcription factor that binds to the TATA box (Haberle and Stark, 2018). Specific transcription factors bind to regulatory elements often found outside the core promoter (Andersson and Sandelin, 2020). For example, the cAMP response element binding protein (CREB) is a specific transcription factor that recognizes cAMP response elements (CREs). The binding of specific transcription factors to regulatory elements causes recruitment of coregulators. Generally, these coregulators do not bind to DNA specifically; rather, they interact with transcription factors and can either function as coactivators or corepressors (Maston et al., 2006). Transcription factors bound to enhancer elements recruit coactivators; whereas transcription factors bound to repressor elements recruit corepressors. For example, the coactivator CREB-binding protein (CBP/p300) interacts with CREB and promotes the assembly of the PIC, thus enhancing the rate of transcription at that locus. With over 3000 different transcription factors expressed in the human genome, this heterogeneity contributes to the tissue-specific —and in some cases, cell-type specific— gene expression patterns observed in humans, especially during early development (Babu et al., 2004).

1.1.2 Epigenetic regulation of gene expression

Gene regulation also occurs at the epigenetic level. Epigenetics refers to heritable changes in gene expression not caused by changes in DNA sequence (Fincham, 1997). From a broad perspective, epigenetic regulation of gene expression occurs through alterations in chromatin structure. Chromatin is organized as a multilayered structure with varying degrees of compaction. At the most fundamental level, chromatin is composed of repeating units known as nucleosomes (Kornberg, 1974). Nucleosomes are comprised of 147 bp of DNA wrapped around a histone octamer containing two copies each of histone H2A, H2B, H3, and H4 (Kornberg, 1974; Quina et al., 2006). Each nucleosome is connected by a segment of linker DNA bound to histone H1, which helps to condense the chromatin into higher-order structures (Kornberg, 1974; Quina et al., 2006; Imhof, 2006). The electrostatic interactions between the negatively charged DNA and the positively charged histones facilitates the packaging of DNA into compact units.

There is approximately 3 billion bp of DNA that is highly condensed and localized in the nucleus. However, for gene regulation to occur, the DNA must still be accessible to the binding of various transcription factors and enzymes. This balance is achieved in part through an epigenetic process known as chromatin remodelling. A key mechanism involved in chromatin remodelling is histone modification, which can promote gene activation or gene repression. Histones have an unstructured, flexible tail domain that extends beyond the nucleosome core (Imhof, 2006). This tail consists of residues that can undergo many forms of post-translational modifications including acetylation, phosphorylation, methylation, and ubiquitination (Quina et al., 2006).

Generally, the histone modifications involved in gene activation recruit effector complexes with histone acetyltransferase (HAT) activity and ATP-dependent chromatin remodelling activity, which work in tandem to restructure the nucleosome and make the DNA more accessible (Barrero et al., 2010). Moreover, acetylation of histones also causes repulsion with DNA due to increased negative charge, leading to increased accessibility of DNA (Barrero et al., 2010). Chromatin that is less condensed and thus more accessible is known as euchromatin; in contrast, chromatin that is highly condensed is known as heterochromatin. Correspondingly, euchromatin is found in genomic regions

that contain actively transcribed genes; whereas, heterochromatin is found in genomic regions that are transcriptionally inactive (Quina et al., 2006; Gibb et al., 2011). Certain histone marks are characteristic of euchromatin, such as trimethylation of H3 at lysine 4 (H3K4me3) and acetylation of H3 at lysine 14 (H3K14ac) (Zhang and Pradhan, 2014). Histone marks for heterochromatin include trimethylation of H3 at lysine 9 (H3K9me3) and lysine 27 (H3K27me3) (Zhang and Pradhan, 2014).

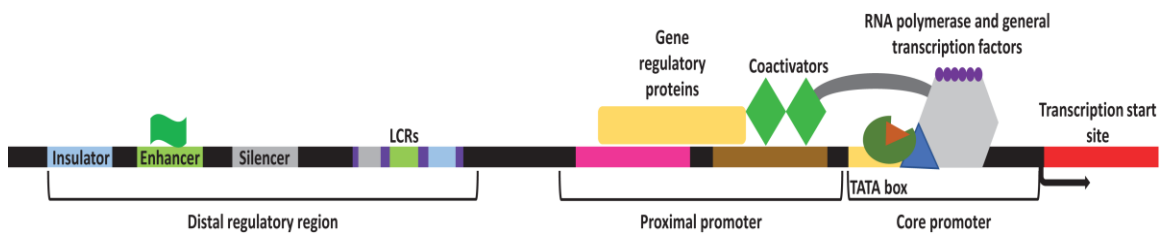
1.1.3 DNA methylation

DNA methylation is the second major mechanism involved in epigenetic regulation. DNA methylation is the covalent addition of a methyl group to the C5 position of cytosine that generates 5-methylcytosine (5mC) (Illum et al., 2018). This modification is catalyzed by DNA methyltransferase (DNMT) enzymes, which transfer the methyl group from the universal methyl-donor S-adenosylmethionine (SAM) onto cytosine (Bhattacharjee et al., 2016; Illum et al., 2018). DNA methylation occurs primarily at cytosine-guanine (CpG) sites, which are unevenly distributed throughout the genome (Illum et al., 2018). Genomic regions that contain a high density of CpGs (>550 bp region with an observed-to-expected CpG ratio of 0.65 or greater) are referred to as CpG islands (CGIs) (Illum et al., 2018). CGIs are highly prevalent in the promoter regions of most genes and they are also present in gene bodies (Vaissiere et al., 2008; Illum et al., 2018). Although 80% of CpGs in the genome are methylated, promoter CGIs are typically unmethylated under normal physiological conditions (Bhattacharjee et al., 2016; Illum et al., 2018). Generally, DNA methylation represses gene expression, especially in the context of promoter regions, as the steric hinderance caused by auxiliary methyl groups restricts the binding of transcription factors to the promoter (Vaissiere et al., 2008; Smallwood and Kelsey, 2012). Moreover, methylated CpG sites (mCpGs) are recognized by mCpG binding proteins (MBPs) which in turn recruit effector complexes with histone deacetylase (HDAC) activity, promoting heterochromatin formation (Rottach et al., 2009; Illum et al., 2018).

DNA methylation impacts many biological processes that are essential to mammalian development, including X-chromosome inactivation, genomic imprinting, transposon silencing, and cell differentiation (Illum et al., 2018). Methylation patterns are

established early during embryonic development by the *de novo* methyltransferases DNMT3a/DNMT3b (Vaissiere et al., 2008). These methylation patterns are maintained through subsequent cell divisions by the function of DNMT1 with its associated cofactor Ubiquitin-like containing PHD and RING finger domains 1 (UHRF1) (Rottach et al., 2009). Alterations in 5mC patterns can lead to abnormal gene expression and genomic instability. These aberrant methylation patterns are frequently observed in most cancers, where promoter regions that are normally unmethylated (e.g. tumor suppressor gene promoters) become hypermethylated, and regions that are normally hypermethylated (e.g. transposable elements) become hypomethylated (Vaissiere et al., 2008; Zhang and Pradhan, 2014). Deletion of any of the three major *Dnmts* (*Dnmt1*, *Dnmt3a*, *Dnmt3b*) in mice resulted in severe genomic hypomethylation and lethality, either during or immediately after embryonic development (Li et al., 1992; Lei et al., 1996; Jaenisch and Bird, 2003).

Figure 1.1. Transcriptional regulatory elements. Transcription is initiated by binding of the pre-initiation complex (consisting of RNAP and GTFs) at the core promoter. The rate of transcription is influenced by the following regulatory elements, which are typically distal relative to the promoter: enhancers, silencers, insulators, and LCRs. Specific transcription factors bind to these elements and recruit coregulators that either enhance or repress transcription. Figure is from Hawkins et al. (2018).



1.2 DNA demethylation

Removal of 5mC occurs through passive or active mechanisms. Passive demethylation involves the replication-dependent dilution of 5mC; whereas, active demethylation is a replication-independent mechanism that requires modification of 5mC followed by base excision repair. Active demethylation requires Thymine DNA Glycosylase (TDG), which can excise 5fC and 5caC. The function of TDG is essential for embryonic development.

1.2.1 Active demethylation

Once thought to be a static epigenetic mark, many studies have since established that DNA methylation at many regions is highly dynamic (Wu and Zhang, 2011; Bhutani et al., 2011; Kohli and Zhang, 2013; An et al., 2017). Removal of 5mC can occur through two mechanisms: passive demethylation or active demethylation. Passive demethylation involves an absence or a reduction of DNMT1 activity following successive rounds of DNA replication (An et al., 2017). This occurs immediately after fertilization, when the maternal genome is passively demethylated due to nuclear exclusion of DNMT1 or UHRF1 (Wu and Zhang, 2014). In contrast, active demethylation, which occurs independent of DNA replication, is the enzymatic process that results in the removal of the methyl group from 5mC in part through base excision repair (BER) (Nabel et al., 2012a). This also occurs immediately after fertilization— before DNA replication takes place— when the paternal genome undergoes a rapid, global demethylation event that generates several 5mC metabolites (Guo et al., 2011; Williams et al., 2012).

For many years, a demethylation pathway involving deamination of 5mC was thought to be the predominant model for active demethylation in mammals. In this model, 5mC is deaminated by the activation-induced deaminase/apolipoprotein B RNA-editing catalytic component (AID/APOBEC) enzyme family, generating a thymine base (Bochtler et al., 2017a). This thymine, now part of a G:T mismatch, is then excised by Thymine DNA Glycosylase (TDG) or its functional family member Methyl-CpG Binding Domain 4 (MBD4), generating an apurinic/apyrimidinic (AP) site (Teperek-Tkacz et al., 2011). This AP site is further processed by AP endonuclease 1 (APE1) before DNA

Polymerase β incorporates an unmodified cytosine and the nick sealed by the XRCC1-DNA Ligase III α complex (Figure 1.2) (Teperek-Tkacz et al., 2011). Spontaneous deamination of 5mC does occur quite often *in vivo*, at a rate that is 2.2-fold higher than deamination of cytosine to uracil (Shen et al., 1994). If left unrepaired, these deamination events can result in G:T/G:U point mutations. Consequently, mCpG sites are hotspots for point mutations in the genome. In fact, approximately one third of all oncogenic point mutations can be attributed to spontaneous deamination at mCpG sites (Shen et al., 1994). Nevertheless, the rate at which spontaneous deamination of cytosine/5mC occurs cannot fully explain the vast number of oncogenic mutations that occur at CpGs (Franchini et al., 2012). Alternatively, some researchers proposed that 5-hydroxymethylcytosine (5hmC), a 5mC metabolite, could be deaminated by AID/APOBEC to generate 5-hydroxymethyluracil (5hmU), which is excised by TDG (Figure 1.2) (Wu and Zhang, 2011). However, several studies have since refuted this deamination model pertaining to active demethylation, providing evidence for the current model which involves succedent oxidation of 5mC (Figure 1.2) (Popp et al., 2010; Bhutani et al., 2011; Teperek-Tkacz et al., 2011; Nabel et al., 2012b). Firstly, AID/APOBEC enzymes showed minimal deamination activity on 5mC or 5hmC *in vitro*; in fact, these enzymes favour unmodified cytosines as their substrate over modified cytosines (Nabel et al., 2012b). Secondly, if this deamination pathway was the main pathway for active demethylation, then a severe (if not lethal) phenotype would be expected in AID/APOBEC-null mice, which are in fact viable and fertile (Popp et al., 2010). Taken together, this evidence suggests that deamination-induced demethylation is unlikely to be the major mechanism for active demethylation.

The discovery of 5hmC in mammalian DNA largely contributed to the widely accepted active demethylation model in mammals. In this model, 5mC is oxidized by a member of the ten-eleven translocation (TET) enzyme family, generating 5hmC. 5hmC is further oxidized by TETs into 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), both of which are recognized and excised by TDG (Wu and Zhang, 2011). This excision creates an AP site which is replaced with an unmethylated cytosine via BER (Figure 1.2) (Wu and Zhang, 2011). The TET family —TET1, TET2, and TET3— are Fe²⁺ and 2-oxoglutarate (2OG)-dependent dioxygenases that are capable of converting 5mC and

5hmC *in vitro* and *in vivo* (Williams et al., 2012). Each member possesses a conserved, C-terminal catalytic core comprised of a double-stranded β -helix (DSBH) domain which contains key residues that interact with Fe^{2+} and 2OG, and a cysteine-rich domain which wraps around the DSBH to stabilize the overall structure and TET-DNA interaction (An et al., 2017; Wu and Zhang, 2017). In addition to the catalytic domain, TET1 and TET3 have an N-terminal CXXC domain which is composed of two zinc finger motifs that can bind DNA (An et al., 2017). This CXXC domain is not present in TET2, likely due to a chromosomal rearrangement in the *Tet2* gene during evolution which formed the *Idax/Cxxc4* gene that contains the CXXC domain (Zhu et al., 2020). Consequently, TET2 interacts with DNA indirectly through its partner IDAX to carry out its role in active demethylation (Wu and Zhang, 2017). The CXXC domain binds preferentially to unmethylated CpG-rich regions (An et al., 2017). TET1/3 are highly enriched at promoter CGIs; whereas TET2 is mostly enriched at gene bodies and enhancer regions (An et al., 2017). Interestingly, TET enzymes display tissue-specific differential expression patterns (Williams et al., 2012). TET1 is predominantly expressed in embryonic stem cells (ESCs), and its expression is gradually downregulated during differentiation; whereas TET2 and TET3 levels increase or remain constant during differentiation (Rasmussen and Helin, 2016; Melamed et al., 2018). TET2 is also expressed in ESCs, albeit at lower levels, however it is robustly expressed in many adult tissues (Rasmussen and Helin, 2016; Melamed et al., 2018). TET2 works in conjunction with TET1 to ensure accurate differentiation during ESC lineage specification, while playing a unique role in hematopoietic stem cell differentiation (Melamed et al., 2018). TET3 is the predominantly expressed TET in oocytes and zygotes, suggesting that TET3 is largely responsible for the global demethylation of the paternal genome that occurs after fertilization (Melamed et al., 2018).

The oxidized 5mC derivatives (5hmC, 5fC and 5caC) generated by TETs act as intermediaries for active demethylation and can also accumulate at specific sites and function as epigenetic marks for gene regulation (Figure 1.3). In mouse ESCs, 5hmC is much more abundant than 5fC/5caC (>10-fold), as TETs convert 10% of 5mC to 5hmC, and only a subset (1-10%) of 5hmC is converted to 5fC/5caC (An et al., 2017). This implies that TET/TDG-mediated active demethylation usually stops at the 5hmC step

because the presence of 5hmC is a more stable epigenetic mark compared to 5fC and 5caC, which are readily excised by TDG (An et al., 2017). Like TET1 expression levels, 5hmC levels are most abundant in ESCs, then gradually decrease during differentiation (Guo et al., 2011). 5hmC levels in gene bodies positively correlate with gene expression (Wu and Zhang, 2017). Counterintuitively, 5hmC levels are low at promoters with high CpG density (e.g. promoters of highly expressed genes), even though TET1/3 are enriched at these regions (Wu and Zhang, 2017). Rather, 5hmC is enriched at: (1) promoters that have low CpG density and/or associated with bivalent domains, which are regions that contain activating and repressive histone marks, typically found in developmental genes that are repressed in ESCs but activated during differentiation, (2) gene bodies of actively transcribed genes, and (3) distal regulatory elements including enhancers, insulators, and regions flanking transcription factor binding sites (Wu and Zhang, 2017). 5fC/5caC are mostly found in gene bodies of actively transcribed genes, at bivalent promoters, and distal enhancers/insulators (Wu and Zhang, 2017). Through mass spectrometry-based proteomics, specific reader proteins have been identified for 5hmC, 5fC, and 5caC, suggesting that each metabolite has a unique biological function (Spruijt et al., 2013). Additionally, *in vitro* studies suggest that the presence of 5fC/5caC in DNA causes stalling of RNAPII during transcription elongation, as RNAPII specifically recognizes 5fC and 5caC and forms hydrogen bonds with the 5-carbonyl and 5-carboxyl group, respectively (Kellinger et al., 2012; An et al., 2017). However, the exact function of 5fC/5caC in gene regulation is not entirely clear.

Figure 1.2. Active demethylation pathways. Several pathways have been proposed for active demethylation such as the deamination pathway, the hydroxylation-deamination pathway, and the current model which is the deamination-independent pathway. In the current model, 5mC is successively oxidized by TETs to generate 5hmC, 5fC, and 5caC.. The 5fC and 5caC bases are excised by TDG leading to an abasic site that is repaired through BER. The alternative pathways involve deamination of either 5mC or 5hmC to thymine or 5hmU (respectively) prior to excision by TDG or other DNA glycosylases such as MBD4 and SMUG1. Figure is from Dalton and Bellacosa (2012).

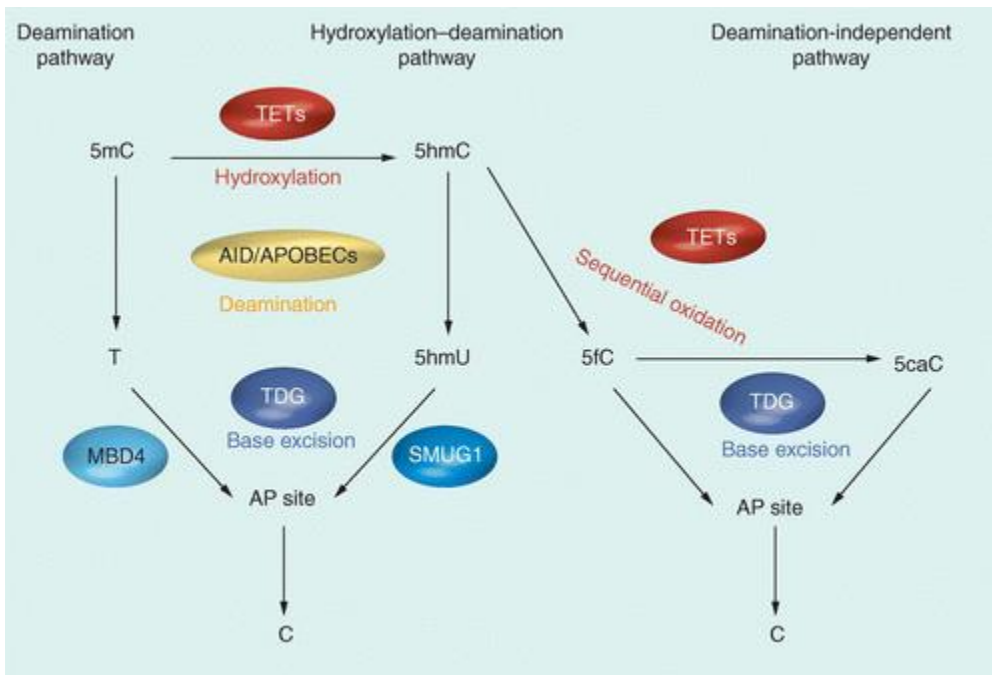


Figure 1.3. Functions of oxidized 5mC derivatives. The process of active demethylation generates several oxidized derivatives of 5mC including: 5hmC, 5fC, and 5caC. Each derivative is bound to specific reader proteins that may recruit chromatin modifying enzymes. Although the exact function of 5fC/5caC in gene regulation is not yet elucidated, 5fC/5caC may impact gene regulation by altering the conformation of DNA and/or reducing the rate of elongation of RNAPII. Figure is from Wu and Zhang (2015).



1.2.2 Thymine DNA Glycosylase

TDG, a 410 amino acid enzyme, was initially identified as a BER enzyme belonging to the mammalian uracil DNA glycosylase (UDG) superfamily which includes Single-Strand-Selective Monofunctional Uracil-DNA Glycosylase 1 (SMUG1), MBD4, and Uracil N-glycosylase 1/2 (UNG1/2) (Neddermann et al., 1996; Cortázar et al., 2007; Sjolund et al., 2013). Within this superfamily, TDG belongs to the Mismatch Uracil DNA Glycosylase (MUG) subfamily which all share a common α/β structural fold, a conserved catalytic core, and variable N- and C-terminal tail regions (Figure 1.4) (Cortázar et al., 2007; Sjolund et al., 2013). These tail regions undergo various post-translational modifications that modulate the enzymatic activity of TDG (Xu et al., 2016; Koliadenko and Wilanowski, 2020). The N-terminal tail domain (aa 1-122) contains a lysine-rich region that undergoes acetylation by CBP, and adjacent serine residues that undergo phosphorylation by protein kinase C alpha (PKC α) (Koliadenko and Wilanowski, 2020). This domain allows TDG to bind to DNA, preferably at AP sites, which prevents enzymatic degradation at these sites (Xu et al., 2016; Koliadenko and Wilanowski, 2020). The C-terminal tail domain (aa 301-410) contains a K330 residue that undergoes sumoylation by Small Ubiquitin-like Modifiers 1-3 (SUMO 1-3) (Sjolund et al., 2013). TDG's catalytic core (aa 123-300) contains the glycosylase domain responsible for excising thymines and uracils from G:T/G:U mismatches (Sjolund et al., 2013). The resulting AP site interacts with residues K246 and K248 within the catalytic core, contributing to TDG's high affinity for AP sites (Maiti et al., 2008; Popov et al., 2020; Koliadenko and Wilanowski, 2020). Accordingly, dissociation of TDG from the AP site is the rate-limiting step of the glycosylase reaction (Sjolund et al., 2013). SUMO conjugation to the C-terminus of TDG promotes dissociation of TDG from the AP site through steric hinderance and recruitment of APE1, which cleaves the AP site (Xu et al., 2016).

Outside of G:T/G:U mismatch repair, TDG has many functional roles relating to transcriptional regulation. As mentioned previously, TDG participates in transcriptional regulation through its role in active demethylation. In fact, TDG is the only member of the UDG superfamily that can excise 5fC and 5caC from double-stranded DNA (Dalton

and Bellacosa, 2012). TDG also contributes to passive demethylation through its interaction with DNMT3a, which inhibits the methylation activity of DNMT3a while enhancing the glycosylase activity of TDG (Sjolund et al., 2013). Moreover, TDG is known to interact with various members of the nuclear receptor (NR) superfamily, including retinoic acid receptor (RAR), retinoid X receptor (RXR), estrogen receptor α (ER α), etc. (Um et al., 1998; Chen et al., 2003). Generally, TDG's interaction with these NRs contributes to transcriptional coactivation. For example, TDG's interacts with the RAR/RXR heterodimer, enhancing its binding to retinoic acid response elements (RAREs) in a ligand-independent manner (Um et al., 1998; Xu et al., 2016). When the RAR/RXR heterodimer is activated by retinoic acid (RA), TDG recruits CBP and forms a ternary complex that induces expression of RA target genes (Xu et al., 2016). Also, TDG can interact in a ligand-dependent manner with ER α to stimulate its activity (Chen et al., 2003; Xu et al., 2016). This interaction causes recruitment of a nuclear receptor coactivator (NCOA) complex including steroid receptor coactivator 1 (SRC1) and SRC3, which induces expression of ER target genes (Xu et al., 2016).

Figure 1.4. TDG structural domains. TDG possesses a core catalytic domain (CAT) flanked by two disordered tail regions. Contained within its N-terminal tail region is a regulatory domain that undergoes various post-translation modifications such as acetylation and phosphorylation. TDG possesses two SUMO-binding motifs (SBM1/2) and undergoes sumoylation at K330. Figure is from Smet-Nocca et al. (2011).



1.2.3 Phenotypic effects of TDG deletion

TDG is highly and ubiquitously expressed in developing mouse embryos, particularly between days E7.5-E13.5 (Cortázar et al., 2007). Afterwards, its expression is enriched in specific tissues including the intestine, lungs, kidney, thymus, and liver (Cortázar et al., 2007). TDG is involved in mesenchymal-to-epithelial transition, which occurs in processes such as somatic cell reprogramming and organogenesis (Bochtler et al., 2017b). TDG is the only DNA glycosylase that is essential for embryonic development, as knockout of *Tdg* in mice resulted in embryonic lethality by day E12.5 (Cortázar et al., 2011). Interestingly, the absence of TDG in these embryos did not affect the mutation frequency associated with BER deficiency, suggesting that the embryonic lethality caused by *Tdg* deletion is likely due to epigenetic abnormalities rather than an accumulation of mutations. Epigenetic abnormalities displayed in *Tdg*-null ESCs include aberrant DNA methylation and irregular histone modifications in the promoter regions of developmental genes (Cortázar et al., 2011). TDG is required for recruiting coactivators like CBP and lysine methyltransferase 2A (KMT2A) to these promoters (Cortázar et al., 2011). Embryonic deletion of *Tdg* resulted in a general loss of activating histone marks and an increase of repressive histone marks at many genes in mouse embryonic fibroblasts (MEFs) (Cortázar et al., 2011). Furthermore, *Tdg*-null ESCs showed a 5-10 fold increase in genomic 5fC and 5caC levels (Shen et al., 2014). Interestingly, the embryonic lethality seen in *Tdg*-null mice was largely attributed to liver hemorrhage, highlighting that TDG plays a role in proper liver development in the embryo. Overall, these knockout studies support the essential role for TDG in maintaining epigenetic stability *in vivo*.

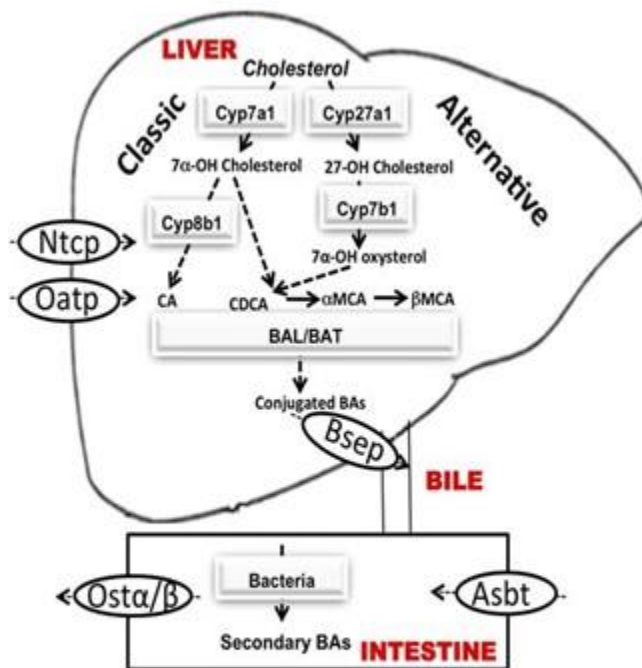
1.3 Hepatic metabolism

The liver performs many metabolic functions in the body including bile acid synthesis. The Farnesoid X Receptor (FXR) is the master regulator of bile acid homeostasis. FXR acts as a tumor suppressor *in vivo* by protecting against the development of hepatocellular carcinoma.

1.3.1 Liver metabolism

The distinguishing feature of the liver is its regenerative capacity. Upon liver injury, hepatocytes —the main parenchymal cell type of the liver (~80% by mass) — transition from their typically quiescent state and undergo substantial proliferation to replace necrotic tissue (Stanger, 2015). Hepatic injury can occur through bile acid (BA) accumulation in the liver as a result of impaired bile flow (i.e. cholestasis), or dysregulation of BA synthesis. BAs are amphipathic sterols with detergent-like properties that can generate cytotoxic effects when present in high concentrations in the liver (Li and Chiang, 2014). Accordingly, the synthesis and transport of BAs is highly regulated. Primary BAs are synthesized from cholesterol primarily through the classical pathway (Li and Chiang, 2014). This pathway is initiated by cholesterol 7 α -hydroxylase (CYP7A1) which performs the rate-limiting step of this pathway by oxidizing cholesterol. Primary BAs are conjugated with glycine or taurine to form bile salts which are more soluble (Li and Chiang, 2014). Bile salts are then secreted into bile through the apical membrane of hepatocytes by the bile salt export pump (BSEP), where they are stored in the gall bladder as mixed micelles to prevent damage to the bile duct (Li and Chiang, 2014). Following postprandial stimulation, bile salts are secreted into the intestine, where they become deconjugated and converted into secondary bile acids to facilitate the absorption of lipids and vitamins (Li and Chiang, 2014). Most BAs (~95%) are reabsorbed by ileal enterocytes via the apical sodium-dependent bile acid transporter (ASBT), and then release into portal circulation via organic solute transporters α/β (OST α/β). Finally, BAs are taken up at the basolateral membrane of hepatocytes via the sodium taurocholate cotransporting polypeptide (NTCP) to undergo enterohepatic circulation (Figure 1.5) (Li and Chiang, 2014).

Figure 1.5. Bile acid synthesis and transport. Primary bile acids (BAs) are synthesized from cholesterol in the liver, primarily through a classical pathway and also through an alternative pathway. BAs are then conjugated to glycine/taurine and secreted into bile via BSEP. ~95% of BAs are reabsorbed in the intestine via ASBT where they are converted to secondary BAs. BAs are then released into portal circulation via OST α/β . BAs in the portal blood are then reabsorbed into hepatocytes by NTCP or OATP to undergo enterohepatic circulation. Figure is from Cheng et al. (2014).



1.3.2 FXR

The farnesoid X receptor (FXR, also referred to as FXR α or *Nr1h4*)—a bile acid receptor that is highly expressed in the liver and intestine—is the master regulator of bile acid homeostasis (Zhu et al., 2011). There are two known FXR genes, FXR α (*Nr1h4*) and FXR β (*Nr1h5*) (Zhang and Edwards, 2008). While FXR β represents a functional receptor in mice, it constitutes a pseudogene in human and primates (Zhang and Edwards, 2008). FXR belongs to the nuclear receptor (NR) superfamily, which are ligand-activated transcription factors that regulate the expression of target genes (Zhu et al., 2011). Generally, NRs share a common structure: (1) a highly conserved N-terminal DNA-binding domain composed of two zinc-finger motifs that recognize specific response elements, (2) a variable N-terminal transactivation (AF-1) domain responsible for modulating transcriptional activity through interactions with cofactors in a ligand-independent manner, (3) a C-terminal ligand-binding domain involved in dimerization of nuclear receptors, (4) a flexible hinge region that links that DNA-binding domain with the ligand-binding domain, and (5) a C-terminal transactivation (AF-2) domain that modulates transcriptional activity in a ligand-dependent manner (Zhu et al., 2011; Li and Chiang, 2014) (Figure 1.6). In the absence of BAs, FXR is bound to FXR response elements (FXREs) as a heterodimer with RXR in association with a host of corepressors (Zhu et al., 2011). Binding of chenodeoxycholic acid (CDCA)—the most efficacious ligand of FXR—to the ligand-binding domain triggers a conformational change in FXR which releases the corepressors and recruits coactivators, promoting the expression of the target gene (Zhu et al., 2011; Ding et al., 2015). The predominant FXR target gene relating to bile acid homeostasis is small heterodimer partner (SHP), which belongs to the ‘orphan’ subfamily of NRs because it lacks a known endogenous ligand (Zhang and Edwards, 2008). Furthermore, SHP cannot bind to DNA because it lacks a DNA-binding domain, hence SHP regulates gene expression by acting as a corepressor (Zhang and Edwards, 2008). FXR-mediated induction of SHP results in downregulation of CYP7A1, as SHP inactivates the NRs that contribute to basal expression of CYP7A1, which are liver receptor homologue-1 (LRH-1) and hepatocyte nuclear factor 4 (HNF4) (Ding et al., 2015). Ultimately, the BA-induced activation of FXR results in a decrease in BA synthesis in the liver through downregulation of CYP7A1, which is the rate-limiting

enzyme in the classical BA synthesis pathway (Figure 1.7). This negative feedback response is one of the mechanisms involved in maintaining BA homeostasis. FXR also functions in the intestine to regulate BA synthesis (Figure 1.7). FXR activation in the small intestine induces expression of fibroblast growth factor 15/19 (FGF15/FGF19), which is secreted to the liver where it binds fibroblast growth factor receptor 4 (FGFR4) found on hepatocytes, initiating a signaling transduction cascade that represses CYP7A1 expression (Zhu et al., 2011). In addition to regulating BA synthesis, FXR regulates BA efflux by inducing expression of BSEP and OST α/β , which promote enterohepatic circulation of BAs (Zhu et al., 2011). Outside of BA homeostasis, FXR also regulates hepatic glucose homeostasis. FXR regulates the expression of several genes involved in the gluconeogenic pathway, including phosphoenol-pyruvate carboxykinase (PEPCK), fructose 1,6-bis phosphatase (FBP1) and glucose-6-phosphatase (G6Pase) (Jiao et al., 2015). The role of FXR activation on hepatic gluconeogenesis is not yet clear (Jiao et al., 2015). Some studies demonstrate that FXR activation inhibits gluconeogenesis through downregulation of PEPCK and G6Pase in human hepatoma cells; however, other studies demonstrate that FXR activation promotes gluconeogenesis through PEPCK upregulation in primary hepatocytes (De Fabiani et al., 2003; Cariou et al., 2005; Stayrook et al., 2005; Zieve et al., 2007).

The murine *Fxr α* gene (*Nr1h4*) contains 11 exons and two different promoters that initiate transcription from either exon 1 or exon 3 (Huber et al., 2002; Jiao et al., 2015). Four isoforms (FXR α 1, FXR α 2, FXR α 3, FXR α 4) are generated through differential promoter usage and alternative splicing between exons 5 and 6, which produces a four amino acid (MYTG) insert immediately adjacent to the DNA-binding domain (Huber et al., 2002; Jiao et al., 2015). The full-length FXR protein (FXR α 3) contains only 9 of the 11 potential exons (Figure 1.6) (Huber et al., 2002; Jiao et al., 2015).

Figure 1.6 FXR structural domains and isoforms. The FXR gene is composed of 11 exons and 10 introns. FXR encodes four isoforms (FXR α 1-4) through the use of two different promoters and alternative splicing between exons 5 and 6 which produces an MYTG insert in FXR α 1 and FXR α 3. FXR's structural domains include the AF1/AF2 activation domains, the DNA-binding domain (DBD), the hinge domain, and the ligand-binding domain (LBD). Figure is from Modica et al. (2010).

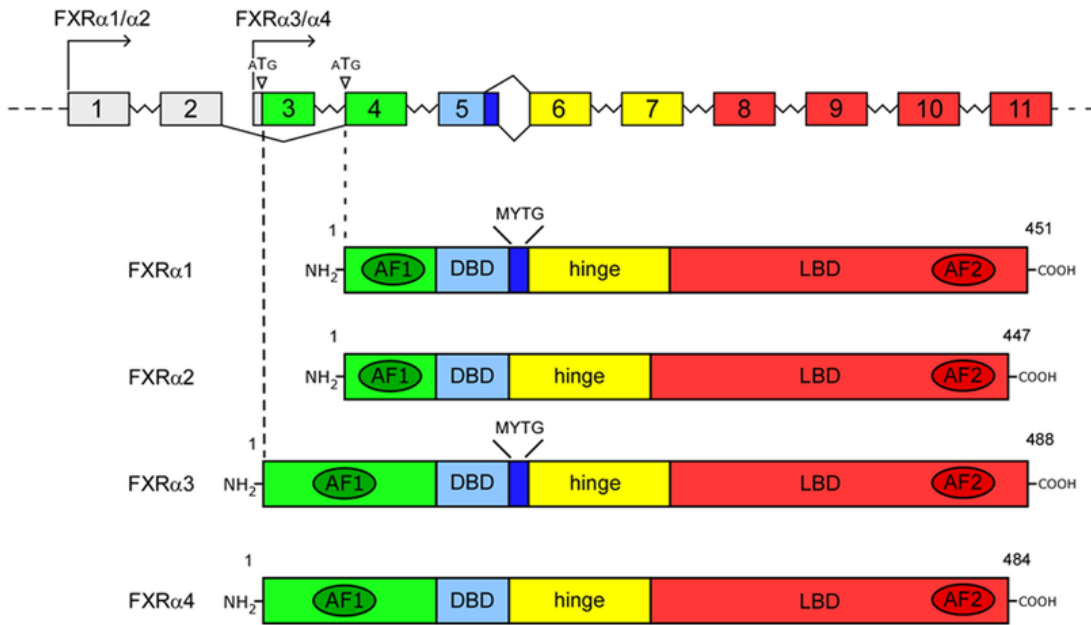
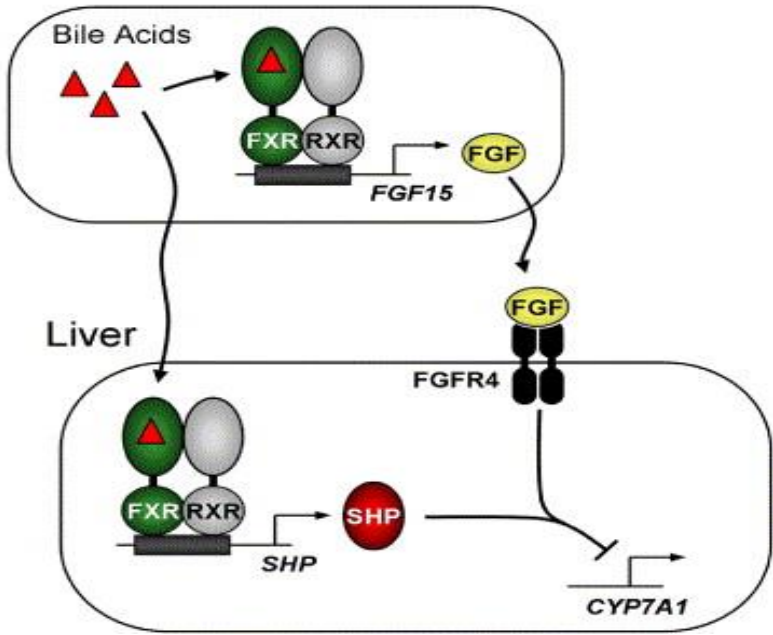


Figure 1.7. FXR signaling pathways. FXR is bound to FXREs as a heterodimer with RXR in association with corepressors. FXR is activated by the binding of BAs. Activation of FXR in the liver induces the expression of SHP which represses CYP7A1 expression, preventing BA synthesis in the liver. Activation of FXR in the intestine induces the expression of FGF15, which is secreted into the liver where it binds to its receptor FGFR4 to repress CYP7A1 expression. Figure is from Inagaki et al. (2005).

Intestine



To our knowledge, three *Fxr* knockout (FXR_{KO}) models have been previously described in the literature (Sinal et al., 2000) (Kok et al., 2003) (Bjursell et al., 2013). Previous FXR_{KO} models have been generated through deletion of either the ligand-binding domain or the DNA-binding domain. One model was generated through Cre-mediated deletion of exon 9 of the full-length FXR protein, which encodes a large portion of the ligand-binding domain, in addition to deletion of the 3'UTR (Sinal et al., 2000). In contrast, another model was generated through targeted deletion of exon 2 of the full length FXR-protein, which encodes a segment of DNA-binding domain (Kok et al., 2003). Lastly, a third model was generated through Cre-mediated deletion of exon 9 of the full-length FXR protein without disrupting the 3'UTR (Bjursell et al., 2013).

Several FXR_{KO} studies have been performed in mice establishing FXR's role as a tumour suppressor *in vivo* (Kim et al., 2007; Yang et al., 2007). FXR protects the liver from cancer development caused by BA overload. FXR_{KO} mice spontaneously develop a late onset hepatocellular carcinoma (HCC) between 12 and 15 months of age and have significantly higher levels of hepatic BAs (Kim et al., 2007; Yang et al., 2007). Moreover, FXR_{KO} mice show elevated fasting glucose levels, insulin resistance, and glucose intolerance (Zhang et al., 2006). In humans, HCC incidence is correlated with downregulation of FXR (Matsubara et al., 2013). HCC incidence in FXR_{KO} mice can be accelerated by deletion of interferon- γ (*IFN γ*), a pro-inflammatory cytokine that modulates *Fxr* expression, and *Shp* (Anakk et al., 2011; Meng et al., 2012; Kim et al., 2017). Since FXR and SHP function linearly in the BA synthesis pathway, it is expected that *Fxr* should be epistatic to *Shp*, such that the phenotype observed in *Fxr*-null mice should be similar to that of *Fxr/Shp* double-knockout (DKO) mice. However, the phenotype displayed in *Fxr/Shp* DKO mice was much more severe than that of *Fxr*-null or *Shp*-null mice (Anakk et al., 2011). The combined deletion of *Fxr* and *Shp* in mice caused juvenile-onset cholestasis, resulting in a significantly higher accumulation of BAs in the liver and serum compared to either knockout alone (Anakk et al., 2011; Kim et al., 2017). This exacerbated phenotype in *Fxr/Shp* DKO mice demonstrates that *Fxr* and *Shp* have nonoverlapping functions pertaining to BA homeostasis, despite their involvement in a common molecular pathway.

Interestingly FXR_{KO} mice share a similar phenotype with conditional *Tdg* knockout (TDG_{CKO}) mice generated by Hassan et al. (2020) to bypass the embryonic lethality of a constitutive *Tdg* knockout. In this mouse model, *Tdg* is deleted eight weeks post-partum by a tamoxifen-inducible Cre-ERT2 under the control of the ubiquitin C (UBC) promoter. Similar to FXR_{KO} mice, TDG_{CKO} mice develop a late onset HCC and display symptoms associated with obesity and type 2 diabetes (T2D), such as increased body weight, glucose intolerance, and the accumulation of primary bile acids with age (Hassan et al., 2020). Through transcriptomic analysis, the metabolic abnormalities observed in TDG_{CKO} mice were attributed to a disruption in FXR signaling, as metabolism and the primary BA synthetic pathway were highly dysregulated following *Tdg* deletion (Hassan et al., 2020).

1.3.3 Hepatocellular carcinoma

Primary liver cancer is the second leading cause of cancer-related deaths worldwide (Sia et al., 2017). HCC, the most common primary liver cancer, accounts for 90% of all primary liver cancers (Sia et al., 2017). HCC typically develops from a background of cirrhosis due to chronic liver injury (Sanyal et al., 2010). Chronic liver injury causes impairment of hepatocyte proliferation and subsequent activation of hepatic stellate cells (HSCs), which are nonparenchymal, progenitor cells located in the space of Disse (Gordillo et al., 2015). HSCs not only proliferate upon activation, but they also differentiate into myofibroblasts and secrete collagen fibers and growth factors which contribute to inflammation and scarring of the liver (i.e. fibrosis/cirrhosis) (Gordillo et al., 2015). This scarring is exacerbated by the activation of Kupffer cells—specialized immune cells resident within the sinusoid—which secrete cytokines that induce a pro-inflammatory immune response at the site of injury (Manco et al., 2018). Prolonged liver regeneration during hepatocarcinogenesis leads to activation of oncogenic signaling pathways such as insulin-like growth factor (IGF) signaling and Hippo signaling (Moeini et al., 2012).

Established risk factors of HCC include alcoholism, hepatitis B, or hepatitis C infection (Sanyal et al., 2010). More recently, T2D has been described as another risk factor for HCC (Sanyal et al., 2010). T2D is characterized by insulin resistance,

hyperglycemia, hyperinsulinemia, and inflammation, all of which have been implicated in the progression of many cancers (Baffy et al., 2012; Yu et al., 2013; Allaire and Nault, 2016). T2D is associated with a 2 to 3-fold increase in the risk for HCC occurrence, with this incidence being significantly higher in male diabetics compared to female diabetics (Gao et al., 2013; Allaire and Nault, 2016). However, the molecular details implicating HCC and T2D are not yet clear.

1.4 Rationale and hypothesis

Our lab has demonstrated that TDG_{CKO} mice spontaneously develop a late-onset HCC associated with T2D symptoms, including increased body weight, glucose intolerance, and BA overload (Hassan et al., 2020). This hepatocarcinogenic phenotype was largely attributed to impaired FXR signalling, suggesting that the loss of TDG negatively impacts FXR signaling (Hassan et al., 2020). This notion is supported by the phenotypic similarities between FXR_{KO} and TDG_{CKO} mice (Zhang et al., 2006; Kim et al., 2007; Yang et al., 2007). Previous studies have demonstrated that TDG acts a coactivator of various nuclear receptor signaling pathways, such as RAR and ER α signaling (Um et al., 1998; Chen et al., 2003). However, a putative role for TDG in FXR signaling has not been previously investigated. These observations prompted the following hypothesis and aims for this study:

Hypothesis: TDG is an essential co-activator for FXR signalling in the liver and the loss of *Tdg* in *Fxr*-null mice will cause epigenetic defects that accelerate the development of HCC.

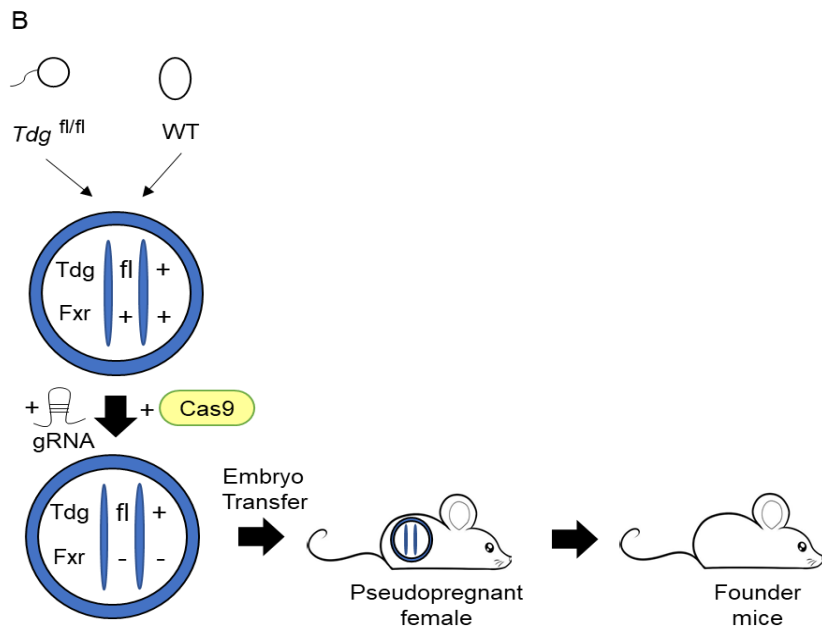
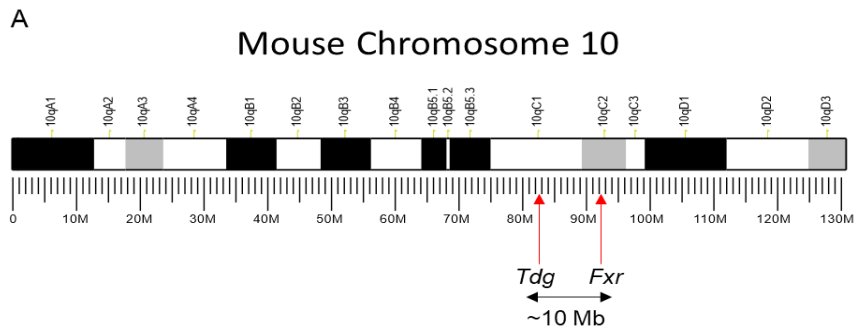
Aim 1. Characterize the molecular role of TDG in FXR signalling

Aim 2. Generate *Tdg/Fxr* DKO mice (*Fxr*^{-/-} *Tdg*^{fl/fl} *CreERT2*^{+/-}) and characterize the preliminary phenotype

For my first aim, I will attempt to characterize a novel role for TDG in FXR signaling by testing a potential interaction between TDG and FXR in the liver. For my second aim, I will utilize CRISPR/Cas9 genome editing to generate *Tdg/Fxr* DKO mice, with the intent to generate a more accelerated HCC mouse model. Ideally, generation of the *Tdg/Fxr* DKO mouse model would entail breeding of commercially available FXR_{KO} mice with our TDG_{CKO} mice. However, *Tdg* and *Fxr* genes are linked on mouse chromosome 10, separated by approximately 10 million bp (Figure 1.8A). Based on the proximity of these genes, the probability of a recombination event at that locus is unlikely (~5%). For this reason, we determined that our desired mouse model should be generated using a genome-editing technique, as opposed to traditional breeding methods (Figure 1.8B). Considering the various genome-editing techniques that are presently available,

we decided to use the CRISPR/Cas9 method to generate our mouse model because of the convenience in experimental design and the knockout efficiency. This genome-editing technique ensures that the *Fxr*-null allele and the floxed *Tdg* allele are present within the same chromatid, resulting in co-segregation of *Tdg* and *Fxr* alleles during genetic recombination. Moreover, this strategy provides us the added benefit of generating a novel *Fxr*-null mouse model in a floxed *Tdg* background.

Figure 1.8 Experimental approach behind generation of *Fxr*-null mice. A) *Tdg* and *Fxr* are located on mouse chromosome 10 separated by ~10 Mb. B) *Tdg*^{flox/flox} sperm were used to fertilize oocytes from wildtype C57BL6 mice. Single-celled embryos were injected with Cas9 and the corresponding gRNA which targets exon 5 of *Fxr*. These embryos were implanted into pseudopregnant CD-1 mice to generate founder mice with the following genotype: *Fxr*^{-/-}*Tdg*^{flox/+}.



2 Methods

2.1 Generation of TDG_{CKO}, *Fxr*-null, and *Tdg/Fxr* DKO mice

All mouse experiments were done in compliance with the Institutional Animal Care and Use Committee guidelines at the London Regional Cancer Center and at Western University (Mouse Protocol Number: 2018-051). For TDG_{CKO} mice (*B6-Tdg^{tm1(cre/ERT2)Torchia}*), *Tdg^{flox/flox}* mice (*B6-Tdg^{tm1Geno}*) were bred with *UBC-cre/ERT2^{+/-}* mice (*B6.Cg-Tg(UBC-cre/ERT2)^{1Ejb/J}*) (Jackson Laboratory, Bar Harbor, ME, USA) to generate *Tdg^{flox/flox} CreERT2^{+/-}* mice. These mice were then bred with *Tdg^{flox/-}* mice to generate the *Tdg^{flox/-} CreERT2^{+/-}* experimental genotype (TDG_{CKO}). In this model, exon 2 of *Tdg* is targeted for deletion, which has previously been shown to generate an efficient knockdown of the TDG protein. For TAM injections, TAM (Sigma, St. Louis, MO, USA, T5648-1G) was dissolved in corn oil (Sigma, St. Louis, MO, USA, C8267) overnight at a concentration of 20 mg/mL and stored at 4°C. Adult (8 weeks old) *Tdg^{flox/-} CreERT2^{+/-}* and age/sex matched *Tdg^{flox/flox}* controls were injected intraperitoneally with 3 mg TAM daily for 5 days to facilitate efficient TDG deletion. TAM-treated mice were then monitored during the course of aging. All mice lines were maintained in a pure C57BL6 background.

Fxr-null mice (*B6-Nr1h4^{em1Torchia}Tdg^{em1Torchia}*) were generated using the CRISPR/Cas9 genome editing method. Briefly, *in-vitro* fertilization (IVF) was performed by 12-hour incubation of oocytes harvested from superovulated wild-type C57BL6 females and sperm harvested from *Tdg^{flox/flox}* mice. Following IVF, all fertilized one-cell zygotes were microinjected with 5 ng of Cas9 mRNA (TriLink Biotechnologies, San Diego, CA, USA) and 2.5 ng of the corresponding gRNA into the male pronucleus. All injected zygotes were incubated overnight at 37°C, and all embryos that developed to the 2-cell stage were implanted into 0.5 dpc pseudopregant CD-1 females the following morning. Founders were bred with wildtype C57BL6 mice to generate heterozygous mice (*Fxr^{+/-} Tdg^{flox/+}*). Heterozygous mice were intercrossed to generate *Fxr*-null mice (*Fxr^{-/-} Tdg^{flox/flox}*). Livers were harvested from 3-week-old *Fxr*-null mice and wildtype mice for gene expression analysis and protein expression analysis.

For *Tdg/Fxr* DKO mice (*B6.Cg-Tg(UBCcre/ERT2)^{1Ejb}Nr1h4^{em1}TorchiaTdg^{em1}Torchia*), *Fxr*-null mice were bred with *UBC-cre/ERT2^{+/-}* mice, and the resulting offspring was backcrossed to generate our working model (*Fxr^{-/-}Tdg^{fl/fl}UBC-cre/ERT2^{+/-}*). 4-week old *Fxr^{-/-}Tdg^{fl/fl}UBC-cre/ERT2^{+/-}* mice and age/sex matched *Fxr^{-/-}Tdg^{flox/flox}* controls were injected intraperitoneally with 1.3 mg TAM daily for 5 days to facilitate efficient TDG deletion. Mice were sacrificed two weeks after the last injection, and then livers and colons were harvested for protein expression analysis.

2.2 RT-PCR/RT-qPCR

Total RNA from liver tissues was extracted using an RNeasy solution (Sigma, St. Louis, MO, USA) as per the manufacturer's protocol. The cDNA was synthesized using the Applied Biosystems Reverse Transcription Kit as per manufacturer's protocol. For RT-PCR, cDNA was PCR amplified using primers from Table 2.1. Samples were loaded onto a 2% agarose gel and electrophoresed for 30 minutes at 180 V. After electrophoresis, DNA fragments were visualized using ethidium bromide staining and imaged using the ChemiDoc XRS imaging system (Bio-Rad, Hercules, CA, USA). Quantitative PCR (qPCR) was performed using a SYBR Green-detection system (Applied Biosystems, Foster City, CA, USA) using primers outlined in Table 2.2. Transcript abundance was normalized to *Gapdh* mRNA.

2.3 Protein extraction and western blot

For protein extraction, tissues were homogenized in 1 ml of ice-cold RIPA lysis buffer (150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 1% NP-40, 50 mM Tris-HCl, pH 8) containing 1x Halt Protease Inhibitor Cocktail (Thermo Scientific, Rockford, IL, USA). Lysates were incubated on ice for 15 minutes, centrifuged at maximum speed (13,000 x g) at 4°C for 15 minutes and the supernatant was retained. Protein concentrations were determined using the Bradford assay. For western blot, 50 µg protein samples were loaded onto a 4-12% gradient Bis-Tris NuPAGE gel (Invitrogen, Carlsbad, CA, USA), subjected to SDS-PAGE, and transferred onto a PVDF membrane. PVDF membranes were incubated in blocking buffer consisting of 5% skim milk in PBS for one hour and hybridized overnight with the appropriate primary antibody at the indicated

dilution. After five ten-minute washes in blocking buffer, membranes were hybridized for one hour with the appropriate secondary antibodies. The membranes were then washed 5 times in blocking buffer and the blots were developed using the Clarity Western ECL Substrate (Bio-Rad, Hercules, CA, USA) and imaged using the ChemiDoc XRS imaging system (Bio-Rad, Hercules, CA, USA). The following primary antibodies were used: rabbit monoclonal anti-TDG (Active Motif, Carlsbad, CA, USA, 1:1000 dilution), rabbit polyclonal anti-FXR (Abcam, Cambridge, MA, USA, ab85606, 1:1000 dilution), mouse monoclonal anti-Vinculin (Abcam, Cambridge, MA, USA, ab130007, 1:20,000 dilution).

2.4 Immunoprecipitation

Liver tissue (~30 mg) was homogenized in 1 ml of lysis buffer (20 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 0.5% NP40, 100 mM NaCl, 1 mM MgCl₂). Then 150 U of the Benzonase nuclease (Sigma, St. Louis, MO, USA) was added to the lysate, and the mixture was rocked at 4°C for 40 minutes. The lysate was then cleared by centrifugation at maximum speed (13,000 x g) for 30 minutes at 4°C, and the supernatant was collected. 5 mg of FXR or IgG antibody was pre-mixed with 50 µl of protein A/G Dynabeads, and subsequently crosslinked using Bissulfosuccinimidyl suberate (BS³) (Thermo Scientific, Rockford, IL, USA) for 40 minutes at room temperature. The BS³ was quenched using 20 mM Tris-HCl, pH 8.0 and then 1 mg of protein from the cleared lysate was diluted ten-fold in buffer A (20 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 0.05% NP40, 150 mM NaCl, 1 mM MgCl₂) and incubated with 50 µl of the crosslinked Dynabeads overnight. The Dynabeads were then washed 10 times using buffer A, eluted using the LDS running buffer and analyzed by western blot. 5% of the cleared lysate (input) was used to normalize loading.

2.5 RNASeq analysis

RNA sample quality was assessed using the Agilent 2100 Bioanalyzer. RNA samples were then prepped following the standard protocol for the NEBnext Ultra ii Stranded mRNA (New England Biolabs, Beverly, MA, USA) at the University of British Columbia. Sequencing was performed on the Illumina NextSeq 500 with Paired End 42 bp × 42 bp reads. The raw data was aligned to the mm10 mouse genome using the STAR

aligner and the gene list was generated using cufflinks (Trapnell et al., 2010). A list of differentially expressed genes was generated using $q < 0.05$ as the cutoff for significance. This RNASeq data was compared with published FXR_{KO} RNASeq data from Anakk et al. (2011) to generate an overlap of the downregulated and upregulated genes between TDG_{CKO} and FXR_{KO} mice.

2.6 CRISPR/Cas9 genome editing

gRNAs were designed and synthesized using the EnGen sgRNA synthesis kit (New England Biolabs, Beverly, MA, USA), as per the manufacturer's protocol (Table 2.4). gRNAs were purified using the Monarch RNA Cleanup Kit (New England Biolabs, Beverly, MA, USA), as per the manufacturer's protocol. For the gRNA targeting assay, gRNAs (2.5 ng) were microinjected into wildtype, single-cell embryos along with Cas9 mRNA (5 ng). DNA was then extracted at the blastocyst stage using the Monarch Genomic DNA Purification kit (New England Biolabs, Beverly, MA, USA), as per the manufacturer's protocol. DNA was PCR amplified and then incubated with *Bsu36I* or *HindIII* (New England Biolabs, Beverly, MA, USA) restriction enzymes for 90 minutes at 37°C (Table 2.3). Samples were loaded onto a 2% agarose gel and electrophoresed for 1 hour at 100 V. For genotyping, tail snips were collected from mouse pups and digested using the DirectPCR Lysis reagent (Viagen, Los Angeles, CA, USA) for DNA analysis, as per the manufacturer's protocol. For sequencing of the mutant *Fxr* allele, DNA from founder mice was PCR amplified and cloned using TA cloning kit (Invitrogen, Carlsbad, CA, USA) as per the manufacturer's protocol. Clones were subjected to *Bsu36I* digest prior to sequencing at the Robarts Sequencing Facility.

2.7 Total bile acid analysis

Liver tissue (~30 mg) was homogenized in 1 ml of 70% ethanol and then incubated at 50°C for 2 hours. The homogenate was centrifuged at 8,000 x g and the supernatant was vacuum-dried, resuspended in 200 mL of water, aliquoted and kept at 80°C until analysis. For BA analysis, the Total Bile Acid Assay Enzyme Cycling Method Kit (Diazyme, San Diego, CA, USA) was used with modifications. Briefly, prior to analysis the calibration curve was made using serial dilutions of 50 mM cholic acid (5

mM – 50 mM). The BioTek Synergy H4 Hybrid reader was used to analyze the samples at 37°C over a 4-minute period, with readings taken every 30 s at 405 nm. The calibration curve was generated by taking the difference between OD405 readings from 30 s and 4 minutes and correlating to concentration of standard used. Then, 4 mL of the liver ethanolic extract was used for analysis and the concentration of BAs was determined using the standard calibration curve.

2.8 Primer sequences

Table 2.1 Primers for RT-PCR

Gene	Forward (5' – 3')	Reverse (5' – 3')
FXR exon 1	GTGTGAAGCCAGCTAAAGGTATGC	TGTGGCTGAACTTGAGGAAACGG
FXR exon 5	GCTGATCAGACAGCTAATGAGG	GTGATTTCTGAGGCATTC
FXR exon 9	CCTCTCTCCAGACAGAC	GGTTCTCAGGCTGGTACATCTTGC

Table 2.2 Primers for RT-qPCR

Gene	Forward (5' – 3')	Reverse (5' – 3')
NR4A1	GTGGGCATGGTGAAGGAAGTTG	AGGGAAGTGAGAAGATTGGTAGGG
IGFBP2	GCCGGTACAACCTTAAGCAGTG	TGCTGCTCGTTGTAGAAGAGATGG
IRS1	ACTATGCCAGCATCAGCTTCCAG	TCTGCTGTGATGTCCAGTTACGC
FXR	GCTGATCAGACAGCTAATGAGG	GTGATTTCTGAGGCATTC
SHP	GCCTGAGACCTTGGTGCCTG	CTGCCACTGCCTGGATGC
CYP7A1	TCAATACCATGCTTTTGTCTGC	GACCTGCACAGCATCCACT
CYP8B1	GCCCTTACTCCAAATCCTACCA	TCGCACACATGGCTCGAT
HSD3B7	TGGTGGATGTGTTTGGGAAGGC	ATTCTGTGTGCCCTGCACGTTG
HSDB3B	ACCAGCTGCGATCCAGAAACCTTC	TACGTGACACTGGAGATGGTCAGC
ABCC2	TCTGTCCAACGCCCTCAACATC	TCTGACGTCATCCTCACTAGCC
ABCC3	TGAGATCGTCATTGATGGGC	AGCTGAGAGCGCAGGTCG
ABCC4	TTAGATGGGCCTCTGGTTCT	GCCACAATTCCAACCTTT
NTCP	CACCATGGAGTTCAGCAAGA	AGCACTGAGGGGCATGATAC
BSEP	GGTTGGTGGACATTAACAGCG	CCTAGGATAAGGACAGCCACACC
PEPCK	GGCCACAGCTGCTGCAG	GGTCGCATGGCAAAGGG
G6PASE	CAGTGGTCGGAGACTGGTTC	TATAGGCACGGAGCTGTTGC
GAPDH	CCAGAACATCATCCCTGC	CTTGGCAGGTTTCTCCAGGC

Table 2.3 Primers for genotyping

FXR forward	ATATGCCTTTGACCGCCCTC
FXR reverse	GGCACACTTTACATATTTCAAGAAC

FXR reverse (47bp deletion)	CACATTTACATATAAATCCCACC
-----------------------------	-------------------------

Table 2.4 gRNA oligos

FXR gRNA 1	TTCTAATACGACTCACTATAGCAACAAACAGAGAATGCCTCGTTTTAGAGCTAGA
FXR gRNA 2	TTCTAATACGACTCACTATAGAATTCACAAAAAAGCTTCCGTTTTAGAGCTAGA

3 Results

3.1 Determining the molecular role of TDG in FXR signaling

Our lab generated conditional *Tdg* knockout (TDG_{CKO}) mice using the Cre-ERT2 mouse model as described in Section 1.3.2. In this model, adult (8-week old) *Tdg*^{flox/-} *CreERT2*^{+/-} mice were intraperitoneally injected with 15 mg TAM over a 5-day period and monitored during the course of aging. We observed that these mice develop a late onset HCC as early as 17-months post-TAM, and this phenotype was more prevalent in males compared to females (Hassan et al., 2020). Moreover, this hepatocarcinogenic phenotype was associated with increased body weight, glucose intolerance, and BA overload (Hassan et al., 2020).

To investigate the transcriptional profile of TDG_{CKO} mice, RNAseq analysis was performed using the pre-tumour livers and hepatocarcinogenic livers of male TDG_{CKO} mice and age/sex matched controls. Transcriptomic analysis revealed that metabolism and the primary BA synthetic pathway were highly dysregulated following *Tdg* deletion (Hassan et al., 2020). This finding, along with the phenotypic similarities between TDG_{CKO} mice and FXR_{KO} mice, prompted us to compare the transcriptional profiles of FXR_{KO} and TDG_{CKO} mice. Interestingly, comparison of these transcriptional profiles showed a significant overlap (~25%) of dysregulated genes, suggesting that TDG and FXR are involved in common molecular pathways (Figure 3.1). To validate the transcriptomic analysis of TDG_{CKO} mice, the expression of several upregulated and downregulated genes were analyzed using qPCR, including genes involved in insulin signaling, FXR signaling, BA synthesis, and BA transport (Figure 3.2). We found that this qPCR analysis was consistent with the RNAseq analysis of TDG_{CKO} mice.

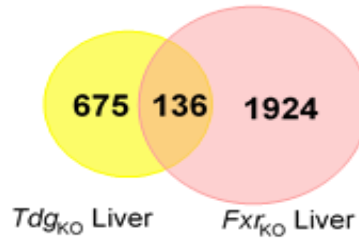
Based on the comparison of the transcriptional profiles of FXR_{KO} and TDG_{CKO} mice, we speculated that TDG participates in the FXR signaling pathway. Furthermore, it is well-established that TDG interacts with various nuclear receptors as a transcriptional coactivator although direct associations between TDG and FXR have not been previously investigated. To explore the potential interaction between TDG and FXR, co-immunoprecipitation experiments were performed using wildtype mouse livers. We

demonstrated a novel interaction between TDG and FXR *in vivo* (Figure 3.3). Altogether, these results suggest that TDG may function as a co-activator of FXR signaling.

Figure 3.1 Transcriptional overlap between TDG_{CKO} and FXR_{KO} mice. Overlap of the downregulated and upregulated genes between age/sex matched TDG_{CKO} and FXR_{KO} livers. Pathway analysis of the upregulated and downregulated genes was performed using the KEGG and REACTOME databases. RNAseq data from male FXR_{KO} livers were obtained from Anakk et al. (2011).

Downregulated genes

$P < 2.95 \times 10^{-14}$



Upregulated genes

$P < 3.88 \times 10^{-19}$

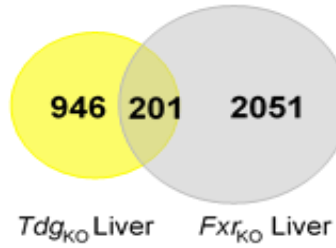


Figure 3.2 qPCR analysis of TDG_{CKO} mice. mRNA was isolated from the livers of control mice (Tdg^{fl/fl}) and 4-month old (post-tamoxifen) TDG_{CKO} mice. qPCR analysis (n=3) was performed on genes involved in: A) insulin signaling B) FXR signaling C) primary BA synthesis and D) BA transport. T-test p-values: *p<0.05, **p<0.01.

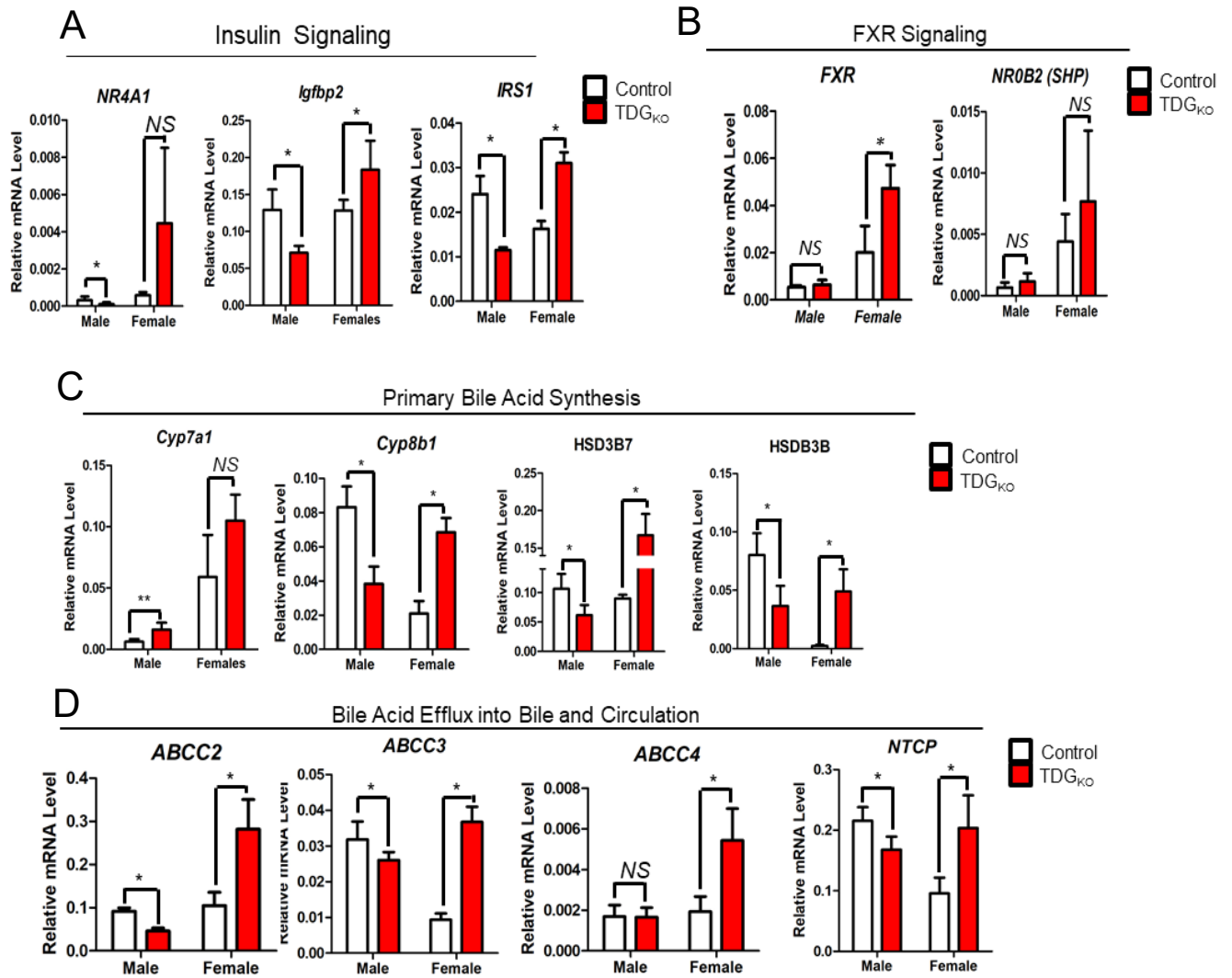
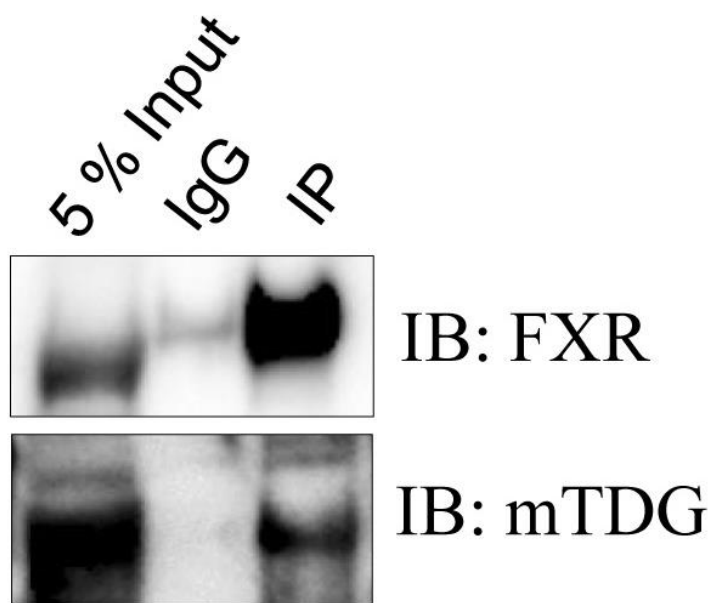


Figure 3.3 Co-immunoprecipitation of TDG and FXR. Immunoprecipitation was performed in the livers of wildtype mice using an FXR antibody. Lysates were pretreated with benzonase nuclease prior to immunoprecipitation. The immunoprecipitated material was then analyzed by western blot using FXR and TDG antibodies. 5% of the cleared lysate (input) was used to normalize loading.



3.2 CRISPR/CAS9-mediated generation of mutant *Fxr* allele

My goal is to generate *Tdg/Fxr* DKO mice, with the hope of producing a more accelerated HCC mouse model. Based on the proximity of the *Tdg* and *Fxr* genes as described in Section 1.4, we used the CRISPR/Cas9 method to generate this mouse model. For the gRNA design, we targeted an exon of FXR which is common to all isoforms to ensure that the gRNA-mediated deletion event resulted in an *Fxr*-null allele. To validate our experimental approach, we designed two guide RNAs (gRNAs): gRNA 1 targets exon 5 of the full-length FXR; whereas gRNA 2 targets exon 6 of the full-length FXR. These gRNAs were synthesized *in vitro* and purified, yielding concentrations of 27 ng/μl (gRNA 1) and 6 ng/μl (gRNA 2) (Figure 3.4A). Exons 5 and 6 contain a *Bsu36I* restriction site and a *HindIII* restriction site, respectively. If the gRNA is successfully targeted to its respective site, Cas9 will induce a double stranded break in the DNA ~3 bp upstream of the PAM site (NGG), which is present within each restriction site. This double stranded break is then repaired through the error-prone process of nonhomologous end joining (NHEJ), resulting in the formation of indels and subsequent alteration of the restriction site (Sander and Joung, 2014). Hence, a successful gRNA-targeting event is identified as a non-cleavage event following restriction enzyme treatment (Figure 3.4B). gRNAs 1 and 2 were injected into 14 embryos and 29 embryos respectively, and the success rates observed for *Fxr*-targeting using gRNA 1 and gRNA 2 were 62% and 20%, respectively. Based on the higher targeting efficiency of gRNA 1, we used this gRNA to generate our *Fxr*-null mice.

After validating the experimental approach through successful gRNA targeting, *in vitro* fertilization (IVF) was performed using oocytes from wildtype C57BL6 mice and *Tdg*^{flox/flox} sperm. Single-cell embryos were injected with gRNA 1 and implanted into pseudopregnant CD-1 mice to generate 5 *Fxr*^{-/-}*Tdg*^{flox/+} founder mice (Figure 3.5). We bred these founders with wildtype C57BL6 mice to generate *Fxr*^{+/-}*Tdg*^{flox/+} offspring, confirming co-segregation of *Tdg* and *Fxr* alleles. To identify the mutation caused by gRNA-mediated deletion of *Fxr* in each of the 5 founders, we PCR-amplified and subcloned the 631-bp fragment corresponding to exon 5 of the full-length FXR, which

encodes a small segment of the hinge domain and the ligand-binding domain. We generated six clones for each of the five mutant alleles, amounting to 30 clones in total. Prior to sequencing, DNA from each clone was subjected to *Bsu36I* treatment to confirm that the restriction site was altered due to an indel mutation. Sequencing analysis of clones resistant to *Bsu36I* digest revealed that the gRNA-mediated mutation that occurred was a 47-bp deletion (Figure 3.6A). This specific deletion event occurred in four of the five mutant alleles. Based on this sequencing analysis, I predicted that this deletion event should produce a premature stop codon within exon 5. To validate these sequencing results, I designed primers to target this 47-bp deletion. Using agarose gel electrophoresis, I confirmed that the mutant *Fxr* allele is 47 bp shorter in length than the wildtype *Fxr* allele (Figure 3.6B). Advantageously, the use of these primers allowed us to bypass the need for *Bsu36I* treatment to confirm the presence of a mutation. Furthermore, in this 47-bp deletion event that occurred, 25 bp of exonic sequence were deleted within exon 5. To validate this, I performed RT-PCR using primers specific to this 25-bp deletion of exonic sequence (Figure 3.6C). Using agarose gel electrophoresis, I confirmed that these primers did not amplify the 25-bp sequence within exon 5 in *Fxr*^{-/-} mice. Collectively, these results demonstrate that gRNA-mediated targeting of *Fxr* generated a 47-bp deletion in exon 5.

Figure 3.4 gRNA-mediated targeting of *Fxr*. A) gRNAs 1 and 2 were synthesized *in vitro* and purified using the Monarch RNA Cleanup Kit, yielding concentrations of 27 ng/ μ l (gRNA 1) and 6 ng/ μ l (gRNA 2). A control gRNA was also synthesized *in vitro*. Samples were loaded onto a 10% TBE-Urea gel and electrophoresed for 1 hour at 180 V. B) Single-cell embryos were injected with Cas9 mRNA and one of two different guide RNAs. DNA was extracted from these embryos, PCR amplified, and then cleaved with *Bsu36I* or *HindIII*. Samples were loaded onto a 2% agarose gel and electrophoresed for 1 hour at 100 V. The 631-bp band corresponds to the undigested *Fxr* amplicon. The 503-bp and 128-bp bands represent *Bsu36I*-digested DNA fragments. The 539-bp and 62-bp bands represent *HindIII*-digested DNA fragments.

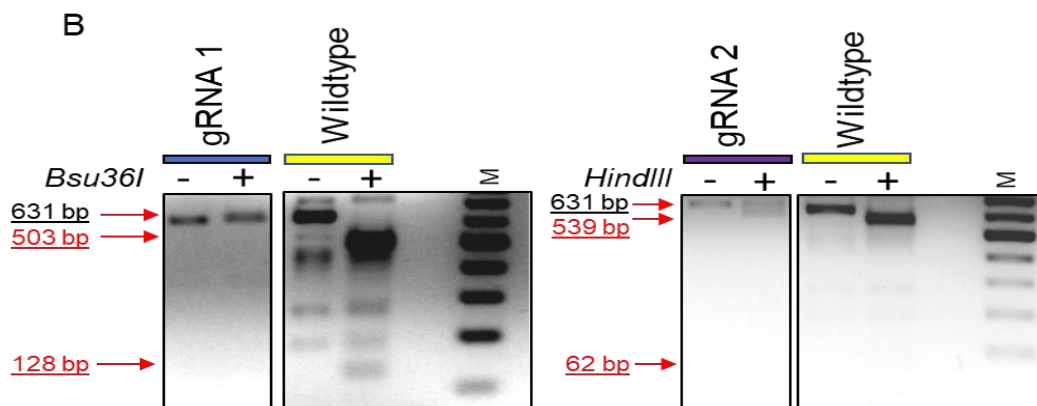
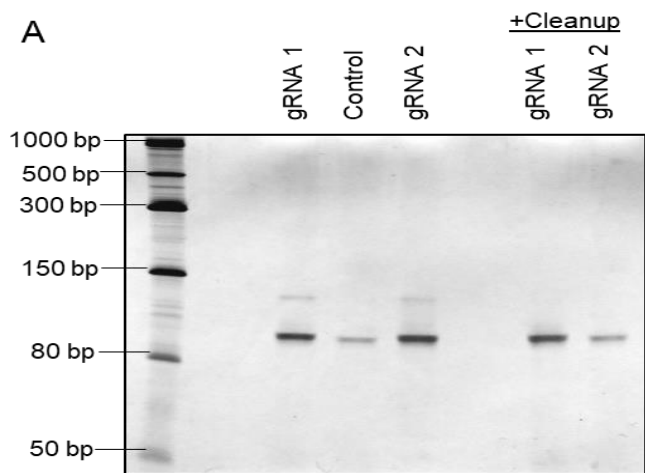


Figure 3.5 Generation of five founder mice. Single-cell embryos were injected with 5ng of Cas9 mRNA and 2.5ng of gRNA 1 which targets exon 5 of *Fxr*. Embryos were implanted into pseudopregnant CD-1 mice to generate 5 founder mice (M1-M5). DNA was extracted from newborn pups, PCR amplified, and then cleaved with *Bsu36I*. Samples were loaded onto a 2% agarose gel and electrophoresed for 1 hour at 100 V. The 631-bp band corresponds to the undigested *Fxr* amplicon. The 503-bp and 128-bp bands (in wildtype) represent *Bsu36I*-digested DNA fragments. DNA from a founder mouse (Pup) that did not survive was also PCR amplified and cleaved with *Bsu36I*.

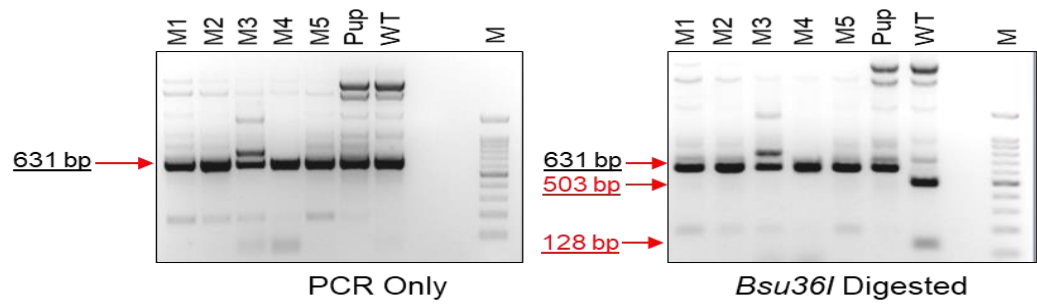
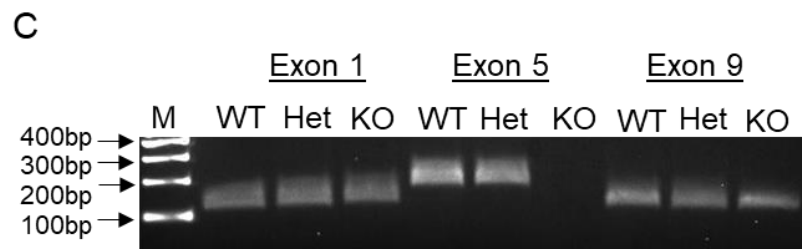
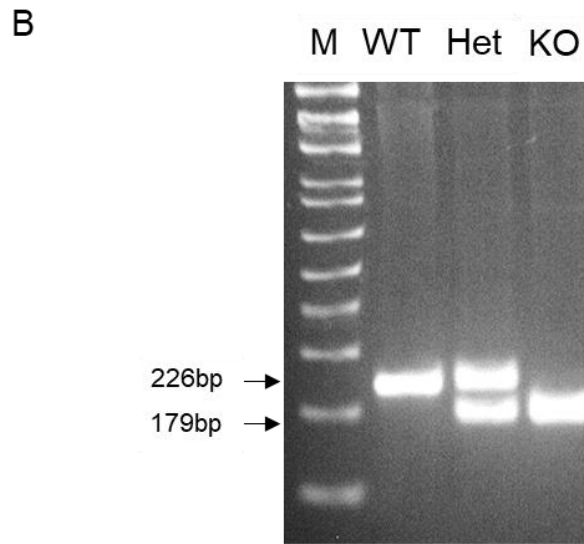
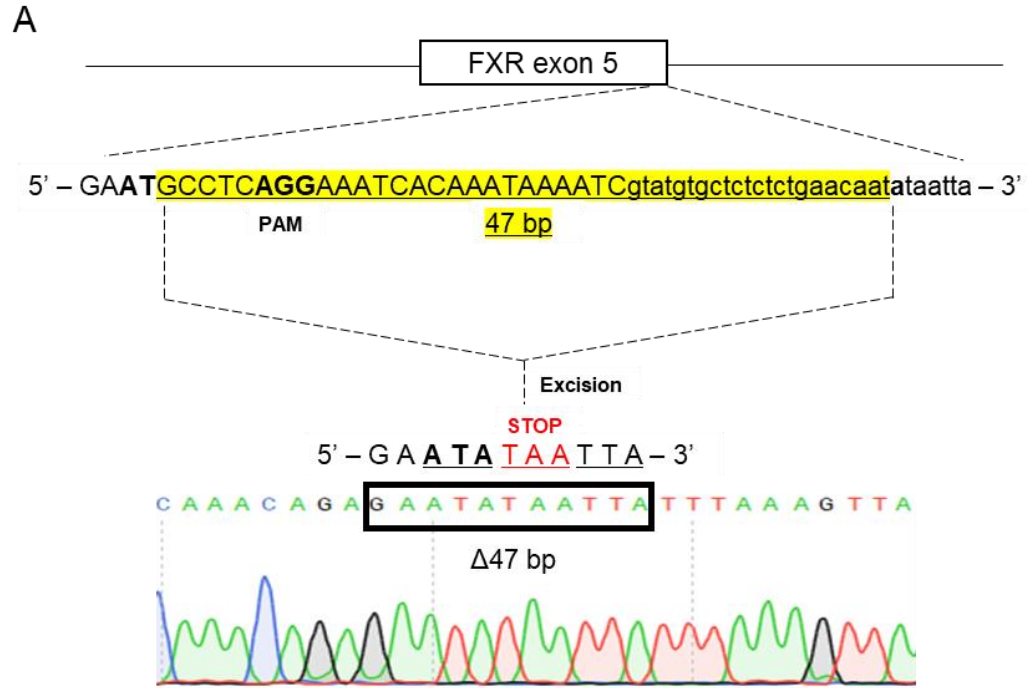


Figure 3.6 A 47-bp deletion event at exon 5 of FXR. A) Founder mice ($Fxr^{-/-} Tdg^{fl/+}$) were bred with wildtype C57BL6 mice to produce $Fxr^{+/-} Tdg^{fl/+}$ offspring. DNA was extracted from these mice and a 631-bp fragment of Fxr was PCR amplified and cloned. Clones resistant to $Bsu36I$ digestion were sequenced at the Roberts Sequencing Facility. B) DNA was extracted from wildtype mice (WT; $Fxr^{+/+} Tdg^{+/+}$), heterozygous mice (Het; $Fxr^{+/-} Tdg^{fl/+}$), and Fxr -null mice (KO; $Fxr^{-/-} Tdg^{fl/fl}$), and then PCR amplified. Samples were loaded onto a 2% agarose gel and electrophoresed for 30 minutes at 180 V. C) RNA from wildtype mice, heterozygous mice, and knockout mice was reverse transcribed. The resulting cDNA was PCR amplified, and samples were loaded onto a 2% agarose gel and electrophoresed for 30 minutes at 180 V.

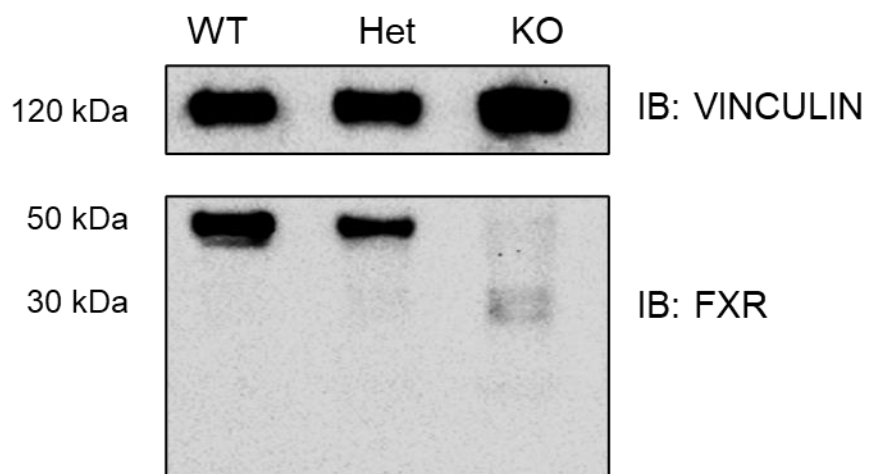


3.3 Generation of *Fxr*-null mice

The 25 bp of exonic sequence that was deleted following NHEJ encodes the start of the ligand-binding domain of FXR. This deletion event is predicted to generate a premature stop codon at the start of the ligand-binding domain. Therefore, any FXR protein produced in *Fxr*^{-/-} mice is expected to be truncated and/or degraded. To confirm this, we generated *Fxr*-null (*Fxr*^{-/-} *Tdg*^{flox/flox}) mice by breeding *Fxr*-het (*Fxr*^{+/-} *Tdg*^{flox/+}) mice together. Using an FXR antibody that specifically recognizes the amino-terminal region of FXR, I found that FXR was not detected in the livers of 3-week old *Fxr*-null mice by western blotting (Figure 3.7A). Interestingly, a faint band was detected at 30 kDa in *Fxr*-null mice, which corresponds with the predicted molecular weight of the putative truncated FXR protein (Figure 3.7A). The presence of this band suggests that a truncated FXR protein is formed in *Fxr*-null mice. As expected, TDG expression was detected in the livers of 3-week old *Fxr*-null mice (Figure 3.7B) and no notable difference in expression was detected compared to wildtype. To measure the transcriptional changes caused by FXR deficiency, I performed qPCR analysis on several FXR target genes involved in FXR signaling, glucose metabolism, BA synthesis, or BA transport. I found that *Shp* was downregulated 4-fold in the livers of 3-week old *Fxr*-null mice compared to age-matched controls (Figure 3.8A). Accordingly, *Cyp7a1* and *Cyp8b1*, both involved in BA synthesis, were upregulated 2-fold and 4-fold in *Fxr*-null mice, respectively, indicating severe dysregulation of BA synthesis (Figure 3.8B). Moreover, *Bsep*, a key BA transporter, was downregulated 1.5-fold in *Fxr*-null (Figure 3.8C). In terms of glucose metabolism, *Pepck* and *G6Pase* were downregulated 4-fold and 3-fold in *Fxr*-null mice, respectively (Figure 3.8D), indicating dysregulation of glucose metabolism. Based on the dysregulation of genes involved BA synthesis in *Fxr*-null mice, I measured the total BA levels in the livers of 3-week old *Fxr*-null mice. I found that the total hepatic BA levels were significantly higher in *Fxr*-null mice compared to age-matched controls (Figure 3.9). Collectively, these results demonstrate that *Fxr*-null mice display significant dysregulation of BA and glucose metabolism. These results are consistent with findings from *Fxr*-null mice generated by conventional methods.

Figure 3.7 Western blot analysis of *Fxr*-null mice. Protein extracts were collected from livers of 3-week old wildtype mice (WT; *Fxr*^{+/+} *Tdg*^{+/+}), heterozygous mice (Het; *Fxr*^{+/-} *Tdg*^{fl/+}), and *Fxr*-null mice (KO; *Fxr*^{-/-} *Tdg*^{fl/fl}). A) 50 μg extracts were analyzed by western blot using Fxr and Vinculin antibodies. B) 50 μg extracts were analyzed by western blot using Tdg and Vinculin antibodies.

A



B

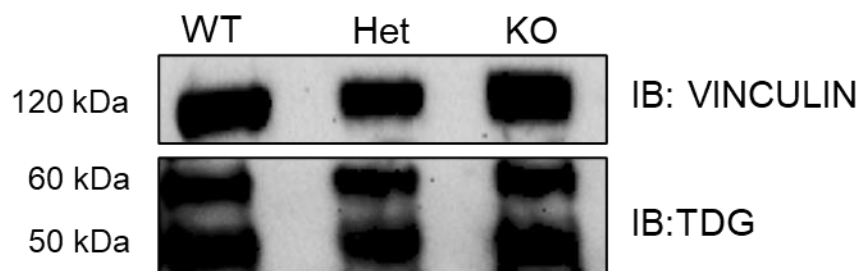


Figure 3.8 qPCR analysis of *Fxr*-null mice. mRNA was isolated from the livers of 3-week old wildtype (WT) mice and *Fxr*-null mice (KO). mRNA levels for the following genes were analyzed by qPCR (n=3): A) *Fxr* and *Shp* (FXR signaling) B) *Cyp7a1* and *Cyp8b1* (BA synthesis) C) *Ntcp*, *Abcc2*, and *Bsep* (BA efflux) D) *G6Pase* and *Pepck* (Glucose metabolism). T-test p-values: *p<0.05, **p<0.01, ***p<0.001.

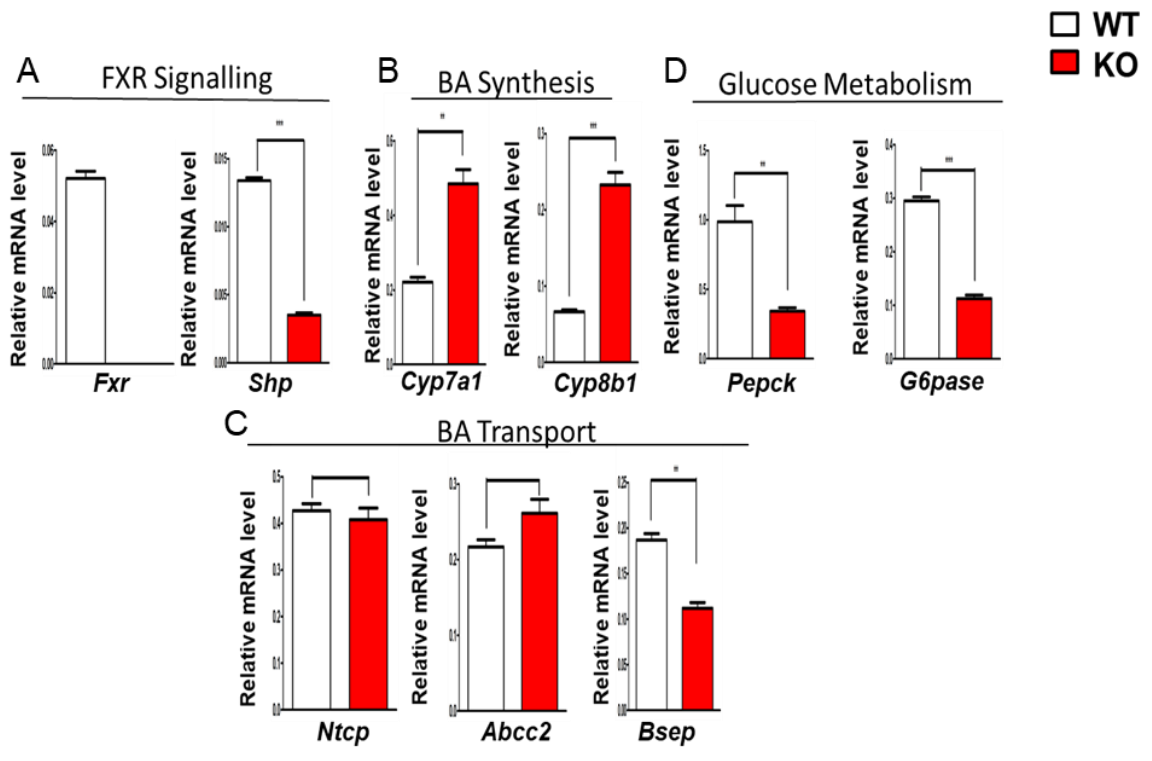
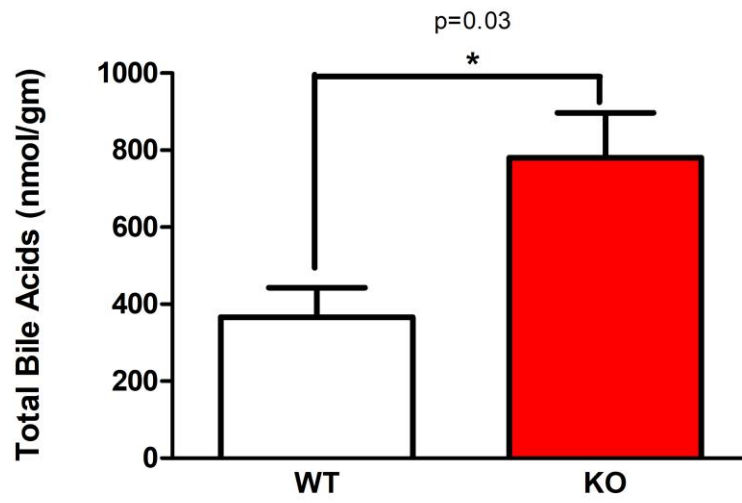


Figure 3.9 Hepatic bile acid levels of *Fxr*-null mice. Total bile acids were isolated from the livers of 3-week old wildtype (WT) mice and *Fxr*-null mice (KO). Bile acid levels were analyzed using the Total Bile Acid Assay Enzyme Cycling Method Kit (Diazyme) (n=2). T-test p-values: *p<0.05.

Hepatic Total Bile Acids



3.4 Generation of *Tdg/Fxr* DKO mice

We have demonstrated the successful generation of *Fxr*-null mice using CRISPR/Cas9. We have confirmed the co-segregation of *Tdg* and *Fxr* alleles in these mice, which will facilitate the generation of *Tdg/Fxr* DKO mice. The breeding strategy for generating *Tdg/Fxr* DKO mice is outlined in Figure 3.10. We bred *Fxr*-null mice with *Fxr*^{+/-} *Tdg*^{flox/+} *CreERT2*^{+/-} mice to generate our working model (*Fxr*^{-/-} *Tdg*^{flox/flox} *CreERT2*^{+/-}). 4-week old *Fxr*^{-/-} *Tdg*^{flox/flox} *CreERT2*^{+/-} mice and age/sex matched *Fxr*^{-/-} *Tdg*^{flox/flox} controls were intraperitoneally injected with 6.5 mg of TAM over a 5-day period (Figure 3.11A). After TAM injections, I did not observe any considerable differences in the overall health of the mice. 2 weeks post-TAM, the liver and colon of DKO mice were harvested for protein expression analysis. In the DKO mice, neither TDG nor FXR expression were detected based on western blotting. (Figure 3.11B, C). Altogether, these results demonstrate that Cre-ERT2 induction by TAM resulted in a conditional knockout of TDG in *Fxr*^{-/-} *Tdg*^{flox/flox} *CreERT2*^{+/-} mice.

Figure 3.10 Breeding scheme for *Tdg/Fxr* double knockout (*Fxr*^{-/-}*Tdg*^{fl/fl} *UBC-cre/ERT2*^{+/-}) mice. *Fxr*^{-/-}*Tdg*^{fl/+} founders were bred with wildtype C57BL6 mice to ensure that the *Fxr*-null allele is carried through the germline. Then, *Fxr*^{+/-}*Tdg*^{fl/+} mice were intercrossed to generate *Fxr*-null mice (*Fxr*^{-/-}*Tdg*^{fl/fl}). *Fxr*-null mice were bred with *UBC-cre/ERT2*^{+/-} mice, and the resulting offspring was backcrossed to generate our working model (*Fxr*^{-/-}*Tdg*^{fl/fl} *UBC-cre/ERT2*^{+/-}).

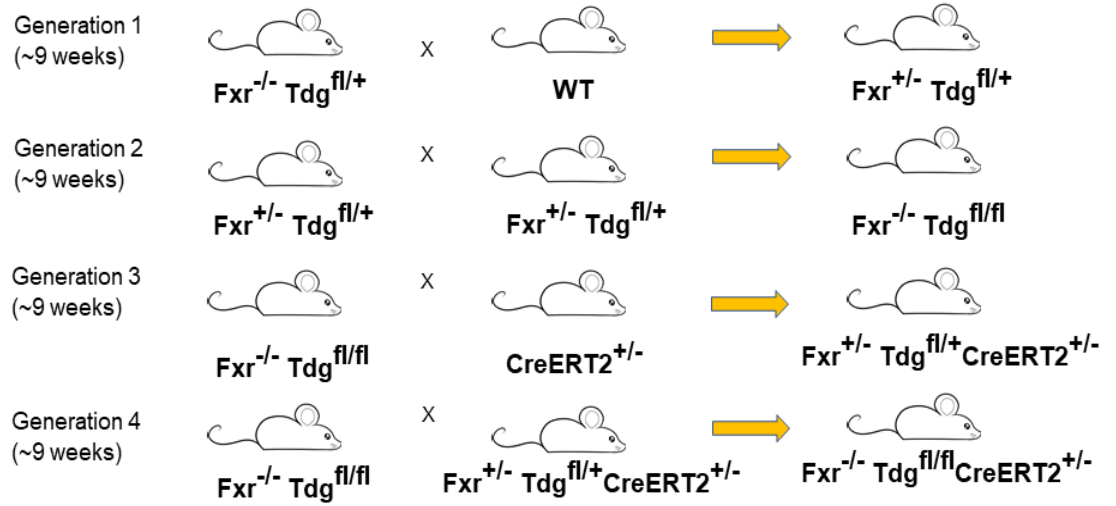
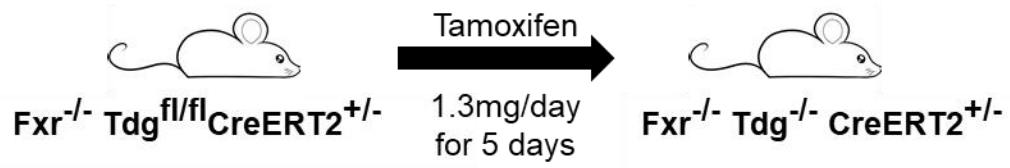
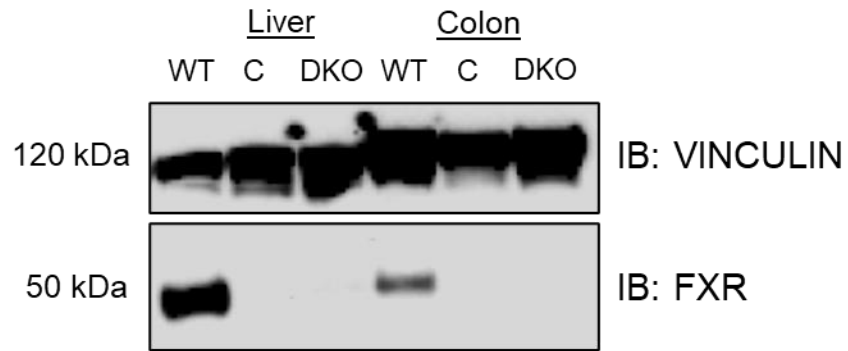


Figure 3.11 Generation of *Tdg/Fxr* double knockout mice. A) 4-week old *Fxr*^{-/-} *Tdg*^{flox/flox} *CreERT2*^{+/-} mice were injected with 6.5 mg of tamoxifen over a 5-day period. Tissues were harvested 2 weeks post-TAM for expression analysis. B) 50 µg protein extracts from wildtype mice (WT; *Fxr*^{+/+} *Tdg*^{+/+}), control mice (C; *Fxr*^{-/-} *Tdg*^{flox/flox}), and double-knockout mice (DKO; *Fxr*^{-/-} *Tdg*^{-/-} *UBC-cre/ERT2*^{+/-}) were analyzed by western blot using *Fxr* and Vinculin antibodies. C) 50 µg protein extracts from wildtype mice (WT; *Fxr*^{+/+} *Tdg*^{+/+}), control mice (C; *Fxr*^{-/-} *Tdg*^{flox/flox}), and double-knockout mice (DKO; *Fxr*^{-/-} *Tdg*^{-/-} *UBC-cre/ERT2*^{+/-}) were analyzed by western blot using *Tdg* and Vinculin antibodies.

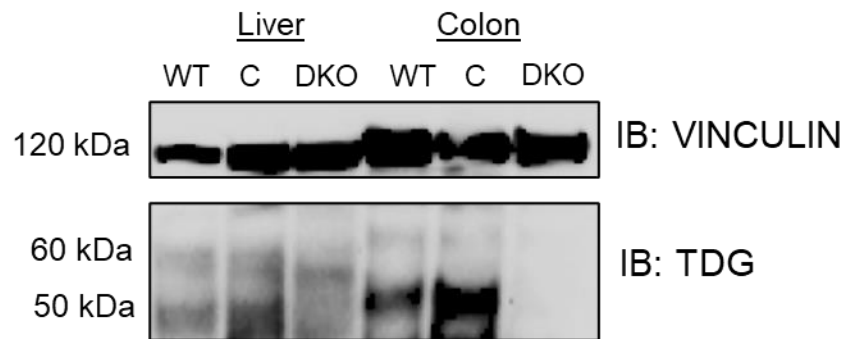
A



B



C



4 Discussion

4.1 Overview

In this study, we established a novel association between TDG and FXR. We also generated a novel *Fxr*-null mouse model using CRISPR/Cas9 genome editing. In this model, the *Fxr*-null allele is linked to a floxed *Tdg* allele, allowing for co-segregation of these alleles during genetic recombination. Additionally, this mouse model facilitated the subsequent generation of *Tdg/Fxr* DKO mice through the function of the tamoxifen-inducible Cre-ERT2. Through sequencing analysis, we showed that the CRISPR/Cas9-mediated mutation that occurred at the *Fxr*-null allele was a 47-bp deletion within exon 5 and its adjacent intron. This deletion event resulted in the formation of a premature stop codon in the ligand-binding domain of FXR, resulting in no FXR expression in the livers of *Fxr*-null mice based on western blotting. Moreover, we showed that *Fxr*-null mice display impaired BA and glucose metabolism as early as 3 weeks post-partum.

4.2 TDG's role as a coactivator of FXR signaling

TDG is known to function as a transcriptional coactivator through its interaction with nuclear receptors. In this study, we aimed to investigate the relationship between TDG and FXR at the molecular level. Initially, when we compared the phenotypes of TDG^{CKO} mice and *Fxr*-null mice, we found that there were many similarities, including late-onset HCC, glucose intolerance, and BA overload. Moreover, when we compared the transcriptional profiles of both mouse models, we found that there was a 25% overlap of dysregulated genes. Several genes contained in this overlap were involved in BA metabolism and glucose metabolism, indicating that TDG and FXR share a common function in these pathways. These findings suggested that TDG may be involved in FXR signaling as a transcriptional coactivator. In support of this notion, we showed that TDG and FXR co-immunoprecipitated using wildtype mouse livers, demonstrating a novel interaction between TDG and FXR *in vivo*. Taken together, these findings support my hypothesis that TDG is an essential coactivator of FXR signaling in the liver.

4.3 Alteration of the *Fxr* allele using CRISPR/Cas9

Another key aim of this study was to generate *Tdg/Fxr* DKO mice. The genetic linkage of *Tdg* and *Fxr* prevented us from generating DKO mice through traditional breeding methods. Therefore, we utilized CRISPR/Cas9 genome editing to generate our novel *Fxr*-null mouse model. Prior to the generation of this mouse model, we validated our CRISPR/Cas9 strategy by experimenting with two different gRNAs. Each gRNA targets a different region of *Fxr*, and we found that gRNA 1 had a higher targeting efficiency compared to gRNA 2. This observation can be attributed to the limitations in our targeting validation assay. The validation assay was designed such that a successful targeting event would be identified as a non-cleavage event following restriction enzyme treatment, as the presence of Cas9 and the gRNA should alter the restriction site through NHEJ. However, based on the difference in position of the PAM sequence in the *Bsu36I* restriction site compared to the *HindIII* restriction site, a larger portion of the restriction site would be altered using gRNA 1 compared to gRNA 2. Consequently, successful targeting events are less likely to be observed using gRNA 2 compared to gRNA 1. This limitation may have contributed to the higher targeting efficiency observed using gRNA 1.

Recognition of the target sequence by the gRNA should induce a double-stranded break by Cas9 ~3 bp upstream of the PAM sequence. Interestingly, through sequencing analysis, I observed that the Cas9 excision occurred 5 bp upstream of the PAM sequence in exon 5 of FXR. Moreover, I observed that this specific excision event resulted in the same DNA repair event (47-bp deletion) in four of the five mutant *Fxr* alleles. In fact, the unique *Fxr* mutant allele contained this 47-bp deletion in addition to a 43-bp insertion immediately upstream. Numerous studies have demonstrated that the CRISPR/Cas9 system displays high editing efficiency (Hwang et al., 2013; Bassett et al., 2013; Hu et al., 2018). Likewise, we demonstrated that the founder mice did not display mosaicism, as the mutant *Fxr* allele was carried through the germline during breeding. Altogether, we showed that the CRISPR/Cas9 system was highly efficient in editing the wildtype *Fxr* allele.

4.4 Comparison of previous *Fxr*-null mouse models

In our *Fxr*-null mouse model, 25 bp of exonic sequence were deleted within exon 5 along with 22 bp of intronic sequence. This deletion event caused a frameshift mutation that generated a premature stop codon in exon 5 of FXR. Through RT-PCR, I showed that the nascent *Fxr* transcript was produced in *Fxr*-null mice; however, the deleted region within exon 5 was not detected. In comparison with previous FXR_{KO} mouse models, only Kok *et al.* (2003) showed that no *Fxr* transcript was produced in their FXR_{KO} model. The premature stop codon generated in our *Fxr*-null mouse model is at position 272 of FXR, which encodes the start of the ligand-binding domain. The presence of this stop codon should result a truncated FXR protein that is nonfunctional, as it does not possess a ligand-binding domain. This truncated FXR protein was detected in *Fxr*-null mice, however its levels were noticeably lower compared to the FXR protein that was present in wildtype mice. These observations suggest that although a truncated FXR protein is expressed in *Fxr*-null mice, it is partially degraded.

Previous FXR_{KO} mouse models display BA overload between 8-12 weeks of age (Sinal *et al.*, 2000; Kok *et al.*, 2003). In this study, we demonstrated that *Fxr*-null mice display BA overload as early as 3 weeks of age. qPCR analysis of *Fxr*-null mice showed that the BA synthesis genes *Cyp7a1* and *Cyp8b1* were significantly upregulated in *Fxr*-null mice, supporting the hepatic BA overload displayed in our mice. This finding is in accordance with Sinal *et al.* (2000) and Kok *et al.* (2003), as both studies showed upregulation of *Cyp7a1* and *Cyp8b1*. Furthermore, I showed that *Shp*, a key regulator of BA homeostasis, was significantly downregulated in *Fxr*-null mice. Also, the hepatic BA transporter *Bsep* was significantly downregulated in *Fxr*-null mice. These findings are in accordance with Sinal *et al.* (2000) and Kok *et al.* (2003), as both studies showed significant downregulation of *Shp* and *Bsep*. The gluconeogenic genes *Pepck* and *G6Pase* were significantly downregulated in *Fxr*-null mice, indicating that FXR deficiency caused a downregulation of hepatic gluconeogenesis. Likewise, Zhang *et al.* (2006) demonstrated *Pepck* and *G6Pase* downregulation in FXR_{KO} mice. Altogether, the findings in our *Fxr*-null model are largely consistent with previous FXR_{KO} mouse models.

4.5 Cre-mediated knockout of TDG in *Tdg/Fxr* DKO mice

The conditional knockout of TDG in adult mice (8-week old) has been previously demonstrated in our lab using the TAM-inducible Cre-ERT2. In this study, we demonstrated the conditional knockout of TDG in 4-week old *Fxr*^{-/-} *Tdg*^{flx/flx} *CreERT2*^{+/-} mice. Because this model contains two floxed *Tdg* alleles as opposed to the one floxed allele in our TDG_{CKO} model, I performed the TAM injection at an earlier timepoint. This increases the probability of observing the hepatocarcinogenic phenotype observed in TDG_{CKO} mice. Although TDG is likely to be a coactivator of FXR signaling, we speculated that *Tdg* is not epistatic to *Fxr* because TDG functions in many pathways aside from FXR signaling. Furthermore, Anakk et al. (2011) demonstrated that *Fxr* is not epistatic to *Shp*, even though both genes function linearly in the BA synthetic pathway. Therefore, I expect that the phenotype displayed in *Tdg/Fxr* DKO mice will be more severe than that of either knockout alone.

4.6 Summary and future directions

Overall, the aims of this study were to characterize the molecular role of TDG in FXR signaling, and to generate and characterize *Tdg/Fxr* DKO mice. Through transcriptomic analysis and co-immunoprecipitation experiments, we implicated TDG as a coactivator of FXR signaling. In addition to generating *Tdg/Fxr* DKO mice, we generated a novel *Fxr*-null mouse model using CRISPR/Cas9. The mutant *Fxr* allele that was generated through CRISPR/Cas9 contained a 47-bp deletion that resulted in the formation of a premature stop codon within exon 5 of FXR. This deletion event resulted in an *Fxr*-null allele that was not detected at the protein level. Our *Fxr*-null model displayed phenotypic traits consistent with that of previous FXR_{KO} models, including dysregulation of glucose metabolism and BA overload.

In the future, we will be further characterizing the *Fxr*-null mice and the DKO mice through aging studies to monitor for the development of HCC. Also, the epigenetic status of DKO mice will be investigated to determine whether loss of *Tdg* in *Fxr*-null mice causes epigenetic defects that accelerate the development of HCC. These potential

findings will provide insight into the epigenetic abnormalities involved in HCC development.

Bibliography

- Allaire, M., and Nault, J.C. (2016). Type 2 diabetes–associated hepatocellular carcinoma: A molecular profile. *Clin. Liver Dis.* 8, 53–58.
- An, J., Rao, A., and Ko, M. (2017). TET family dioxygenases and DNA demethylation in stem cells and cancers. *Exp. Mol. Med.* 49, e323–e323.
- Anakk, S., Watanabe, M., Ochsner, S.A., McKenna, N.J., Finegold, M.J., and Moore, D.D. (2011). Combined deletion of Fxr and Shp in mice induces Cyp17a1 and results in juvenile onset cholestasis. *J. Clin. Invest.* 121, 86–95.
- Andersson, R., and Sandelin, A. (2020). Determinants of enhancer and promoter activities of regulatory elements. *Nat. Rev. Genet.* 21, 71–87.
- Andersson, R., Sandelin, A., and Danko, C.G. (2015). A unified architecture of transcriptional regulatory elements. *Trends Genet.* 31, 426–433.
- Babu, M.M., Luscombe, N.M., Aravind, L., Gerstein, M., and Teichmann, S.A. (2004). Structure and evolution of transcriptional regulatory networks. *Curr. Opin. Struct. Biol.* 14, 283–291.
- Baffy, G., Brunt, E.M., and Caldwell, S.H. (2012). Hepatocellular carcinoma in non-alcoholic fatty liver disease: an emerging menace. *J. Hepatol.* 56, 1384–1391.
- Barrero, M.J., Boué, S., and Izpisua Belmonte, J.C. (2010). Epigenetic Mechanisms that Regulate Cell Identity. *Cell Stem Cell* 7, 565–570.
- Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.-L. (2013). Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep.* 4, 220–228.
- Bhattacharjee, D., Shenoy, S., and Bairy, K.L. (2016). DNA Methylation and Chromatin Remodeling: The Blueprint of Cancer Epigenetics (Hindawi).
- Bhutani, N., Burns, D.M., and Blau, H.M. (2011). DNA Demethylation Dynamics. *Cell* 146, 866–872.
- Bjursell, M., Wedin, M., Admyre, T., Hermansson, M., Böttcher, G., Göransson, M., Lindén, D., Bamberg, K., Oscarsson, J., and Bohlooly-Y, M. (2013). Ageing Fxr Deficient Mice Develop Increased Energy Expenditure, Improved Glucose Control and Liver Damage Resembling NASH. *PLoS ONE* 8.
- Bochtler, M., Kolano, A., and Xu, G.-L. (2017a). DNA demethylation pathways: Additional players and regulators. *BioEssays* 39, e201600178.
- Bochtler, M., Kolano, A., and Xu, G.-L. (2017b). DNA demethylation pathways: Additional players and regulators. *BioEssays* 39, e201600178.

- Cariou, B., Duran-Sandoval, D., Kuipers, F., and Staels, B. (2005). Farnesoid X Receptor: A New Player in Glucose Metabolism? *Endocrinology* *146*, 981–983.
- Chen, D., Lucey, M.J., Phoenix, F., Lopez-Garcia, J., Hart, S.M., Losson, R., Buluwela, L., Coombes, R.C., Chambon, P., Schär, P., et al. (2003). T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor alpha. *J. Biol. Chem.* *278*, 38586–38592.
- Cheng, X., Zhang, Y., and Klaassen, C.D. (2014). Decreased Bile-Acid Synthesis in Livers of Hepatocyte-Conditional NADPH-Cytochrome P450 Reductase-Null Mice Results in Increased Bile Acids in Serum. *J. Pharmacol. Exp. Ther.* *351*, 105–113.
- Cortázar, D., Kunz, C., Saito, Y., Steinacher, R., and Schär, P. (2007). The enigmatic thymine DNA glycosylase. *DNA Repair* *6*, 489–504.
- Cortázar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A.L., Siegrist, F., et al. (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* *470*, 419–423.
- Cramer, P. (2019). Organization and regulation of gene transcription. *Nature* *573*, 45–54.
- Dalton, S.R., and Bellacosa, A. (2012). DNA Demethylation by TDG. *Epigenomics* *4*, 459–467.
- De Fabiani, E., Mitro, N., Gilardi, F., Caruso, D., Galli, G., and Crestani, M. (2003). Coordinated control of cholesterol catabolism to bile acids and of gluconeogenesis via a novel mechanism of transcription regulation linked to the fasted-to-fed cycle. *J. Biol. Chem.* *278*, 39124–39132.
- Ding, L., Yang, L., Wang, Z., and Huang, W. (2015). Bile acid nuclear receptor FXR and digestive system diseases. *Acta Pharm. Sin. B* *5*, 135–144.
- Fincham, J.R.S. (1997). *Epigenetic Mechanisms of Gene Regulation*. Edited by V. E. A. Russo, R. A. Martienssen and A. D. Riggs. Cold Spring Harbor Laboratory Press, 1996. 693+xii pages. Price \$125. ISBN 0 87969 490 4. *Genet. Res.* *69*, 159–162.
- Franchini, D.-M., Schmitz, K.-M., and Petersen-Mahrt, S.K. (2012). 5-Methylcytosine DNA Demethylation: More Than Losing a Methyl Group. *Annu. Rev. Genet.* *46*, 419–441.
- Gao, S., Li, A., Liu, F., Chen, F., Williams, M., Zhang, C., Kelley, Z., Wu, C.-L., Luo, R., and Xiao, H. (2013). NCOA5 Haplo-insufficiency Results in Glucose Intolerance and Subsequent Hepatocellular Carcinoma. *Cancer Cell* *24*, 725–737.
- Gibb, E.A., Brown, C.J., and Lam, W.L. (2011). The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer* *10*, 38.

- Gordillo, M., Evans, T., and Gouon-Evans, V. (2015). Orchestrating liver development. *Development* *142*, 2094–2108.
- Guo, J. (2014). Transcription: the epicenter of gene expression. *J. Zhejiang Univ. Sci. B* *15*, 409–411.
- Guo, J.U., Su, Y., Zhong, C., Ming, G., and Song, H. (2011). Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle* *10*, 2662–2668.
- Haberle, V., and Stark, A. (2018). Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol.* *19*, 621–637.
- Hassan, H.M., Isovici, M., Kolendowski, B., Bauer-Maison, N., Onabote, O., Cecchini, M., Haig, A., Maleki Vareki, S., Underhill, T.M., and Torchia, J. (2020). Loss of Thymine DNA Glycosylase Causes Dysregulation of Bile Acid Homeostasis and Hepatocellular Carcinoma. *Cell Rep.* *31*, 107475.
- Hawkins, L.J., Al-attar, R., and Storey, K.B. (2018). Transcriptional regulation of metabolism in disease: From transcription factors to epigenetics. *PeerJ* *6*, e5062.
- Hu, P., Zhao, X., Zhang, Q., Li, W., and Zu, Y. (2018). Comparison of Various Nuclear Localization Signal-Fused Cas9 Proteins and Cas9 mRNA for Genome Editing in Zebrafish. *G3 Genes Genomes Genet.* *8*, 823–831.
- Huber, R.M., Murphy, K., Miao, B., Link, J.R., Cunningham, M.R., Rupar, M.J., Gunyuzlu, P.L., Haws, T.F., Kassam, A., Powell, F., et al. (2002). Generation of multiple farnesoid-X-receptor isoforms through the use of alternative promoters. *Gene* *290*, 35–43.
- Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* *31*, 227–229.
- Illum, L.R.H., Bak, S.T., Lund, S., and Nielsen, A.L. (2018). DNA methylation in epigenetic inheritance of metabolic diseases through the male germ line. *J. Mol. Endocrinol.* *60*, R39–R56.
- Imhof, A. (2006). Epigenetic regulators and histone modification. *Brief. Funct. Genomic. Proteomic.* *5*, 222–227.
- Inagaki, T., Choi, M., Moschetta, A., Peng, L., Cummins, C.L., McDonald, J.G., Luo, G., Jones, S.A., Goodwin, B., Richardson, J.A., et al. (2005). Fibroblast growth factor 15 functions as an enterohepatic signal to regulate bile acid homeostasis. *Cell Metab.* *2*, 217–225.
- Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* *33*, 245–254.

- Jiao, Y., Lu, Y., and Li, X. (2015). Farnesoid X receptor: a master regulator of hepatic triglyceride and glucose homeostasis. *Acta Pharmacol. Sin.* *36*, 44–50.
- Kellinger, M.W., Song, C.-X., Chong, J., Lu, X.-Y., He, C., and Wang, D. (2012). 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* *19*, 831–833.
- Kim, I., Morimura, K., Shah, Y., Yang, Q., Ward, J.M., and Gonzalez, F.J. (2007). Spontaneous hepatocarcinogenesis in farnesoid X receptor-null mice. *Carcinogenesis* *28*, 940–946.
- Kim, K.H., Choi, S., Zhou, Y., Kim, E.Y., Lee, J.M., Saha, P.K., Anakk, S., and Moore, D.D. (2017). Hepatic FXR/SHP axis modulates systemic glucose and fatty acid homeostasis in aged mice. *Hepatology* *66*, 498–509.
- Kohli, R.M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* *502*, 472–479.
- Kok, T., Hulzebos, C.V., Wolters, H., Havinga, R., Agellon, L.B., Stellaard, F., Shan, B., Schwarz, M., and Kuipers, F. (2003). Enterohepatic circulation of bile salts in farnesoid X receptor-deficient mice: efficient intestinal bile salt absorption in the absence of ileal bile acid-binding protein. *J. Biol. Chem.* *278*, 41930–41937.
- Koliadenko, V., and Wilanowski, T. (2020). Additional functions of selected proteins involved in DNA repair. *Free Radic. Biol. Med.* *146*, 1–15.
- Kornberg, R.D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA. *Science* *184*, 868–871.
- Lei, H., Oh, S.P., Okano, M., Juttermann, R., Goss, K.A., Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* *122*, 3195–3205.
- Li, T., and Chiang, J.Y.L. (2014). Bile Acid Signaling in Metabolic Disease and Drug Therapy. *Pharmacol. Rev.* *66*, 948–983.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* *69*, 915–926.
- Maiti, A., Morgan, M.T., Pozharski, E., and Drohat, A.C. (2008). Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition. *Proc. Natl. Acad. Sci.* *105*, 8890–8895.
- Manco, R., Leclercq, I.A., and Clerbaux, L.-A. (2018). Liver Regeneration: Different Sub-Populations of Parenchymal Cells at Play Choreographed by an Injury-Specific Microenvironment. *Int. J. Mol. Sci.* *19*.

- Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional Regulatory Elements in the Human Genome. *Annu. Rev. Genomics Hum. Genet.* 7, 29–59.
- Matsubara, T., Li, F., and Gonzalez, F.J. (2013). FXR signaling in the enterohepatic system. *Mol. Cell. Endocrinol.* 368, 17–29.
- Melamed, P., Yosefzon, Y., David, C., Tsukerman, A., and Pnueli, L. (2018). Tet Enzymes, Variants, and Differential Effects on Function. *Front. Cell Dev. Biol.* 6.
- Meng, Z., Wang, X., Gan, Y., Zhang, Y., Zhou, H., Van Ness, C., Wu, J., Lou, G., Yu, H., He, C., et al. (2012). Deletion of IFN γ enhances hepatocarcinogenesis in FXR knockout mice. *J. Hepatol.* 57, 1004–1012.
- Modica, S., Gadaleta, R.M., and Moschetta, A. (2010). Deciphering the nuclear bile acid receptor FXR paradigm. *Nucl. Recept. Signal.* 8.
- Moeini, A., Cornellà, H., and Villanueva, A. (2012). Emerging Signaling Pathways in Hepatocellular Carcinoma. *Liver Cancer* 1, 83–93.
- Nabel, C.S., Manning, S.A., and Kohli, R.M. (2012a). The Curious Chemical Biology of Cytosine: Deamination, Methylation and Oxidation as Modulators of Genomic Potential. *ACS Chem. Biol.* 7, 20–30.
- Nabel, C.S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H.L., Stivers, J.T., Zhang, Y., and Kohli, R.M. (2012b). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat. Chem. Biol.* 8, 751–758.
- Neddermann, P., Gallinari, P., Lettieri, T., Schmid, D., Truong, O., Hsuan, J.J., Wiebauer, K., and Jiricny, J. (1996). Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J. Biol. Chem.* 271, 12767–12774.
- Popov, A.V., Grin, I.R., Dvornikova, A.P., Matkarimov, B.T., Groisman, R., Saparbaev, M., and Zharkov, D.O. (2020). Reading Targeted DNA Damage in the Active Demethylation Pathway: Role of Accessory Domains of Eukaryotic AP Endonucleases and Thymine-DNA Glycosylases. *J. Mol. Biol.* 432, 1747–1768.
- Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by Aid deficiency. *Nature* 463, 1101–1105.
- Quina, A.S., Buschbeck, M., and Di Croce, L. (2006). Chromatin structure and epigenetics. *Biochem. Pharmacol.* 72, 1563–1569.
- Rasmussen, K.D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* 30, 733–750.
- Rottach, A., Leonhardt, H., and Spada, F. (2009). DNA methylation-mediated epigenetic control. *J. Cell. Biochem.* 108, 43–51.

- Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347–355.
- Sanyal, A.J., Yoon, S.K., and Lencioni, R. (2010). The etiology of hepatocellular carcinoma and consequences for treatment. *The Oncologist* 15 *Suppl 4*, 14–22.
- Shen, J.C., Rideout, W.M., and Jones, P.A. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.* 22, 972–976.
- Shen, L., Song, C.-X., He, C., and Zhang, Y. (2014). Mechanism and Function of Oxidative Reversal of DNA and RNA Methylation. *Annu. Rev. Biochem.* 83, 585–614.
- Sia, D., Villanueva, A., Friedman, S.L., and Llovet, J.M. (2017). Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis. *Gastroenterology* 152, 745–761.
- Sinal, C.J., Tohkin, M., Miyata, M., Ward, J.M., Lambert, G., and Gonzalez, F.J. (2000). Targeted Disruption of the Nuclear Receptor FXR/BAR Impairs Bile Acid and Lipid Homeostasis. *Cell* 102, 731–744.
- Sjolund, A.B., Senejani, A.G., and Sweasy, J.B. (2013). MBD4 and TDG: Multifaceted DNA glycosylases with ever expanding biological roles. *Mutat. Res. Mol. Mech. Mutagen.* 743–744, 12–25.
- Smallwood, S.A., and Kelsey, G. (2012). De novo DNA methylation: a germ cell perspective. *Trends Genet.* 28, 33–42.
- Smet-Nocca, C., Wieruszeski, J.-M., Léger, H., Eilebrecht, S., and Benecke, A. (2011). SUMO-1 regulates the conformational dynamics of Thymine-DNA Glycosylase regulatory domain and competes with its DNA binding activity. *BMC Biochem.* 12, 4.
- Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W.T.C., Bauer, C., Münzel, M., Wagner, M., Müller, M., Khan, F., et al. (2013). Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* 152, 1146–1159.
- Stanger, B.Z. (2015). Cellular homeostasis and repair in the mammalian liver. *Annu. Rev. Physiol.* 77, 179–200.
- Stayrook, K.R., Bramlett, K.S., Savkur, R.S., Ficorilli, J., Cook, T., Christe, M.E., Michael, L.F., and Burris, T.P. (2005). Regulation of Carbohydrate Metabolism by the Farnesoid X Receptor. *Endocrinology* 146, 984–991.
- Teperek-Tkacz, M., Pasque, V., Gentsch, G., and Ferguson-Smith, A.C. (2011). Epigenetic reprogramming: is deamination key to active DNA demethylation? *Reproduction* 142, 621–632.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification

- by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515.
- Um, S., Harbers, M., Benecke, A., Pierrat, B., Losson, R., and Chambon, P. (1998). Retinoic Acid Receptors Interact Physically and Functionally with the T:G Mismatch-specific Thymine-DNA Glycosylase. *J. Biol. Chem.* 273, 20728–20736.
- Vaissiere, T., Sawan, C., and Herceg, Z. (2008). Epigenetic interplay between histone modifications and DNA methylation in gene silencing. *Mutat. Res. Mutat. Res.* 659, 40–48.
- Williams, K., Christensen, J., and Helin, K. (2012). DNA methylation: TET proteins—guardians of CpG islands? *EMBO Rep.* 13, 28–35.
- Wu, H., and Zhang, Y. (2011). Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev.* 25, 2436–2452.
- Wu, H., and Zhang, Y. (2014). Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. *Cell* 156, 45–68.
- Wu, H., and Zhang, Y. (2015). Charting oxidized methylcytosines at base resolution. *Nat. Struct. Mol. Biol.* 22, 656–661.
- Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.* 18, 517–534.
- Xu, X., Watt, D.S., and Liu, C. (2016). Multifaceted roles for thymine DNA glycosylase in embryonic development and human carcinogenesis. *Acta Biochim. Biophys. Sin.* 48, 82–89.
- Yang, F., Huang, X., Yi, T., Yen, Y., Moore, D.D., and Huang, W. (2007). Spontaneous development of liver tumors in the absence of the bile acid receptor farnesoid X receptor. *Cancer Res.* 67, 863–867.
- Yu, J., Shen, J., Sun, T.T., Zhang, X., and Wong, N. (2013). Obesity, insulin resistance, NASH and hepatocellular carcinoma. *Semin. Cancer Biol.* 23, 483–491.
- Zhang, G., and Pradhan, S. (2014). Mammalian epigenetic mechanisms. *IUBMB Life* 66, 240–256.
- Zhang, Y., and Edwards, P.A. (2008). FXR signaling in metabolic disease. *FEBS Lett.* 582, 10–18.
- Zhang, Y., Lee, F.Y., Barrera, G., Lee, H., Vales, C., Gonzalez, F.J., Willson, T.M., and Edwards, P.A. (2006). Activation of the nuclear receptor FXR improves hyperglycemia and hyperlipidemia in diabetic mice. *Proc. Natl. Acad. Sci. U. S. A.* 103, 1006–1011.

Zhu, T., Brown, A.P., and Ji, H. (2020). The Emerging Role of Ten-Eleven Translocation 1 in Epigenetic Responses to Environmental Exposures. *Epigenetics Insights* 13.

Zhu, Y., Li, F., and Guo, G.L. (2011). Tissue-specific function of farnesoid X receptor in liver and intestine. *Pharmacol. Res.* 63, 259–265.

Zieve, F.J., Kalin, M.F., Schwartz, S.L., Jones, M.R., and Bailey, W.L. (2007). Results of the glucose-lowering effect of WelChol study (GLOWS): a randomized, double-blind, placebo-controlled pilot study evaluating the effect of colesevelam hydrochloride on glycemic control in subjects with type 2 diabetes. *Clin. Ther.* 29, 74–83.

Curriculum Vitae

Name: Oladapo Onabote

Post-secondary Education and Degrees: The University of Western Ontario
London, Ontario, Canada
2013-2017 B. MSc., Honors Specialization in Biochemistry

The University of Western Ontario
London, Ontario, Canada
2018- M.Sc., Department of Biochemistry

Honours and Awards: Western Graduate Research Scholarship
2018-2020

Dean's Honor List
2017

Department Chair Recognition of Outstanding Performance
in Chemistry
2013

Western Scholarship of Distinction
2013

Related Work Experience Undergraduate Thesis Student
The University of Western Ontario
Supervisor: Dr. Michael Boffa
2017

Publications: Hassan HM, Isovici M, Kolendowski B, Bauer-Maison N, Onabote O, Cecchini M, Haig A, Vareki SM, Underhill MT, and Torchia J (2020). Loss of Thymine DNA Glycosylase Causes Dysregulation of Bile Acid Homeostasis and Hepatocellular Carcinoma. Cell Reports, 31, 107475.