4-26-2017

MicroRNA-152-3p Serves as an Indirect Epigenetic Modifier by Positively Regulating SETDB1 Protein Synthesis

Supriya K. Singh
University of Connecticut - Storrs, supriya.singh@uconn.edu

MicroRNA-152-3p Serves as an Indirect Epigenetic Modifier by Positively Regulating SETDB1 Protein Synthesis

Supriya Kumari Singh, PhD

University of Connecticut, 2017

ABSTRACT

MicroRNAs (miRNA, miR) are traditionally thought of as negative regulators of post-transcriptional gene expression. However, recent work shows that rarely, some miRNAs actually increase transcription or translation of their targets. MiR-152-3p is one such unique miRNA. We demonstrate miR-152-3p enhances protein expression of SET domain bifurcate 1 (SETDB1), a histone methyltransferase responsible for trimethylation of histone 3 lysine 9 (H3K9me3). Heterochromatin protein 1 (HP1) binds to H3K9me3 to assist in the formation and maintenance of transcriptionally inactive regions of DNA called constitutive heterochromatin, which in turn is associated with gene silencing. We conclude regulation of SETDB1 by miR-152-3p has functional consequences on H3K9me3 levels. Therefore, positive regulation of SETDB1 by miR-152-3p is the first known example of a miRNA which promotes the formation of constitutive heterochromatin by serving as an indirect epigenetic modifier. The miRanda database algorithmically predicts three target recognition sequences (TRS) in the 3'UTR of SETDB1. We inserted the SETDB1 3'UTR into a luciferase reporter vector and deleted each TRS to demonstrate the necessity of these TRS for full luciferase expression. We found that at least one of the TRS is required for full luciferase expression. We are the first to demonstrate a miRNA which positively regulates protein synthesis does so by targeting the 3'UTR, rather than the 5'UTR. In order to study the effects of miR-152-3p on SETDB1 mRNA and protein, we modulated levels of the miRNA with mimic or inhibitor (anti-miR). The mimic replicates the

functioning of endogenous miR-152-3p and the anti-miR prevents endogenous miR from binding to its targets. Using these tools, we found miR-152-3p mimic treatment increased SETDB1 protein levels, decreased H3K9me2 levels, and increased H3K9me3 levels. Conversely, inhibition of miR-152-3p decreased SETDB1 protein levels and increased H3K9me2 levels. No correlation was observed between miR-152-3p and SETDB1 mRNA, indicating the miRNA likely acts on translation machinery. Through this work, we advance our knowledge and understanding of miRNA roles in protein and epigenetic regulation. Our finding that miR-152-3p indirectly promotes the epigenetic modification associated with heterochromatin formation and gene silencing provides a novel perspective from which to approach research in the field of epigenetics.

MicroRNA-152-3p Serves as an Indirect Epigenetic Modifier by Positively Regulating SETDB1 Protein Synthesis

Supriya Kumari Singh

B.S., University of Georgia, **2011**

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2017

APPROVAL PAGE

Doctor of Philosophy Dissertation

MicroRNA-152-3p Serves as an Indirect Epigenetic Modifier by Positively Regulating SETDB1 Protein Synthesis

Presented by

Supriya Kumari Singh, B.S.

Major Advisor _____

Theodore P. Rasmussen

Associate Advisor_____

Charles Giardina

Associate Advisor_____

Xiaobo Zhong

Associate Advisor_____

José Manautou

Associate Advisor_____

Kyle Hadden

University of Connecticut

2017

DEDICATION

To my family for their never ending support, belief in me, encouragement, and inspiration.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my past and present mentors in research: Dr. and Mrs. Steve Stice, Dr. Ping Shen, and the graduate students and post-doctoral fellows in their labs at the University of Georgia, as well as my advisor and doctoral thesis committee, Dr. Theodore Rasmussen, Dr. Charles Giardina, Dr. Xiabo Zhong, Dr. José Manautou, and Dr. Kyle Hadden at the University of Connecticut.

Thank you to my loving family and friends for all their encouragement and support over the years: my father, Dr. Rakesh Singh, my mother, Mrs. Sunita Singh, my brother, Dr. Rahul Singh, my husband, Raj Ranade, my father-in-law, Dr. Gautam Ranade, my mother-in law, Mrs. Jyoti Ranade, my sister-in-law, Priya Ranade, M.P.H./M.B.A., Manisha Reddy, Dr. Erin Carner, Layal Kaba, J.D., and Kate Helmick, M.P.H..

Throughout my time at the University of Connecticut, I have fostered deep connections with colleagues. Their never ending advice and comradery helped me get through all the ups and downs of grad school. I am eternally grateful for all the hands on training and input my friends have given me over the years.

Thank you, Anthony Flamier, Dr. Borko Tanasijevic, Dr. Igor Gurevich, Dr. Swetha Rudraiah, and Phil Rohrer, M.S. for training me on a variety of molecular biology techniques and being available to answer all my questions.

Thank you to my amazing friends: Ziyan Zhao, Dr. Gregory Smith, Jennifer Pace, Dr. Kelly Teske, and Chad Pope. I appreciate all the time you all have spent double checking my work. Your kindness and support mean the world to me.

I especially want to thank Dr. Joseph Chichocki, the greatest friend, mentor, and colleague a person could ever hope for. Joe always encouraged me to work hard and keep at research, even if it wasn't going well. I can never thank him enough for his guidance throughout my time at UConn.

I also have the utmost appreciation for the following individuals for all of their advice, project assistance, and friendship: Dr. Upasana Banerjee, Dr. Albert DeBerardinis, Talya Mandelkern, Bindu Prabhakar, Stephanie Piekos, Rambon Shamilov, and Dr. Vincent Ramirez. Thank you to Heather White for initiating work on bisulfite sequencing for the COPZ2 promoter. Special thanks to Madeline Dorso for her dedication towards H3K9me2/3 western blots and HP1α western blots and immunofluorescence. I am extremely appreciative of all the hard work she put towards this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

CD	Chromodomain
CDS	Coding sequence
CSD	Chromoshadow domain
DNMT1	DNA methyltransferase 1
GDM	Global DNA methylation
H3K9me1	Mono-methylated histone 3 lysine 9
H3K9me2	Dimethylated histone 3 lysine 9
H3K9me3	Trimethylated histone 3 lysine 9
HCC	Hepatocellular carcinoma
HLC	Hepatocyte-like cell
HD	Hinge domain
HMT	Histone methyltransferase
HP1	Heterochromatin protein 1
iPSC	Induced pluripotent stem cell
KDM	Lysine demethylase
KID	KAP1-interacting domain
KMT	Lysine methyltransferase
MBD	Methyl binding domain
miR	microRNA
RISC	RNA-induced silencing complex
SAM	S-adenosyl-L-methionine
SETDB1	Su(var) 3-9 Enhancer Of Zeste, Trithorax Domain Bifurcate 1
TRS	Target recognition sequence
UTR	Untranslated region

CHAPTER 1

INTRODUCTION AND BACKGROUND

Introduction

Rationale

According to an algorithm developed by MicroRNA.org (miRanda), the miRNA miR-152-3p targets the histone lysine methyltransferase SETDB1 (su(var) 3-9 enhancer of zeste trithorax domain bifurcate 1) at the 3'UTR[1]. The algorithm makes this prediction by applying rules of miRNA-mRNA binding to the known sequences of miRNAs and mRNA targets. These rules are devised from patterns in experimentally determined and accepted miRNA-mRNA binding pairs. However, the binding of miR-152-3p to the SETDB1 3'UTR has not been experimentally confirmed. The only experimental evidence that miR-152-3p targets SETDB1 is from a CLASH (crosslinking, ligation, and sequencing of hybrids) binding assay, but CLASH showed that miR-152-3p binds to the coding sequence (CDS) instead of the miRanda predicted 3'UTR of SETDB1[2]. A potential reason for this discrepancy is that CLASH can include data on transient interactions, since the first step of the procedure is to crosslink. Our work aimed to confirm the miRanda predicted target recognition sites (TRS) and provided strong evidence that miR-152-3p binds to the 3'UTR.

Because miR-152-3p potentially targets SETDB1, we hypothesized that miR-152-3p regulates SETDB1, but there is no evidence in the literature to support this. Furthermore,

whether miR-152-3p promotes or represses SETDB1 expression is unknown. MiRNAs are often described as acting only in a repressive manner[1]. Yet, there are rare instances of miRNAs upregulating transcript or protein expression of their targets[3-5]. Therefore, it cannot be assumed that miRNAs exclusively destabilize mRNA or repress translation. For these reasons, our work aimed to determine if miR-152-3p affects SETDB1 at the mRNA or protein levels and to determine if miR-152-3p promotes or represses SETDB1 expression.

SETDB1 is a histone methyl transferase responsible for adding a methyl group on to dimethylated histone H3 lysine 9 (H3K9me2) residues in euchromatin, forming trimethyl histone H3 lysine 9 (H3K9me3)[6]. H3K9me3 initiates the binding of HP1, a structural component responsible for the maintenance and renewal of heterochromatin[7-9]. Therefore, if miR-152-3p targets and regulates SETDB1, it is predicted that miR-152-3p impacts the epigenetic modifications of H3K9 and, by extension, chromatin structure via HP1 recruitment. To address this prediction, our work aimed to demonstrate whether manipulation of miR-152-3p significantly affects H3K9me2/3 levels and HP1 localization.

Biological and Medical Significance

Our work has two primary sources of biological and medical significance. First, we provide biological information on miR-152-3p. Without our research, it was unknown whether miR-152-3p targeted the 3'UTR of SETDB1 *in vivo*. We were first to perform mutagenesis at the 3'UTR of SETDB1 and confirm any miRNA activity at the 3'UTR of SETDB1. With the present CLASH data, it is not known if miR-152-3p binding to SETDB1 represents a transient or stable, regulatory, interaction. Furthermore, the directionality (promoting or repressing) and level

(mRNA or protein) of SETDB1 regulation are unidentified in the literature. We concluded miR-152-3p positively regulates SETDB1 protein synthesis, and in doing so, contributed to the present knowledge.

Second, we provided the crucial foundation upon which disease-relevant work can be based. In addition to pioneering experiment-based research on miR-152-3p binding to the 3'UTR of SETDB1, we were also the first to demonstrate a corresponding impact on histone methylation status. Histone methylation is not presented in the literature as being susceptible to control by a miRNA. In looking at the indirect consequences of miR-152-3p regulation of SETDB1 on H3K9me2/3, a greater understanding of miRNA regulation of histone methylation is achieved.

With this novel understanding, new approaches in disease research can be taken since abnormalities in histone methylation status are heavily implicated in diseases such as cancer and Huntington's disease (HD). There is evidence that decreasing expression of SETDB1 improves disease outcome and increasing levels of miR-152 reduces cancer progression. One group found SETDB1 and H3K9me3 levels are coordinated in HD. When SETDB1 is chemically down-regulated, hypermethylation at H3K9 is reduced, and survival is extended by over 40% in HD model mice[10]. According to Dang, et. al., miR-152 is down-regulated in hepatocellular carcinoma (HCC), and this under expression of miR-152 is promotes hepatocarcinogenesis by dysregulating cell proliferation, cell motility, and apoptosis[11]. Interestingly, Wong, et. al. determined that SETDB1 expression is elevated in hepatocellular carcinoma[12]. Through our work, a disease pathway can be constructed to help put existing research into a new context. With further studies, our work may also contribute towards elucidating what part of the pathway is aberrant in diseases such as HD and HCC.

Background

MicroRNAs and miR-152-3p

MicroRNAs (miRNA, miR) are a class of small, single-stranded, non-coding RNAs that primarily post-transcriptionally regulate gene expression. MiRNAs can target regions of mRNA, resulting in gene silencing. This is accomplished through mRNA cleavage or translational repression[1]. Rarely, miRNAs may increase transcription or translation of their targets. In the case of miR-324-3p, RelA is transcriptionally activated upon promoter-miRNA binding in an Ago2-dependednt manner[5]. In regards to translation, miR-10a and miR-122 have both been found to stimulate translation of targets. MiR-10a enhances translation by binding to the 5'UTR of target ribosomal mRNA[3]. MiR-122 acts by increasing the association of ribosomes with hepatitis C virus (HCV) RNA during the initiation of translation, thereby increasing translation of HCV[4, 13]. MiR-152-3p is one miRNA that is thought to target the coding region of SETDB1[2]. The miRanda database further indicates that miR-152-3p may target the 3'UTR of SETDB1[14].

It is believed that most miRNA genes come from regions of the genome far from previously annotated genes. This may mean that those distant miRNA genes are derived from their own independent transcription units, according to a primary literature review on miRNA biogenesis by Bartel. However, similar to small nucleolar RNAs (snoRNAs), approximately one quarter of miRNA genes are processed from the introns of pre-mRNAs[1]. These miRNAs are called intronic miRNAs and are processed from unspliced intronic regions prior to splicing catalysis[15]. Since intronic miRNAs lack their own promoters, coordinated expression of miRNAs and proteins can be seen[1]. MiR-152 is an example of an intronic miRNA gene. It is 109 base

pairs in length and located between the first and second exons of the COPZ2 gene at locus 17q21.32[16].

Canonical miRNA biogenesis begins with Drosha, an RNase III endonuclease, cleaving the primary miRNA (pri-miRNA) in the nucleus. This releases an approximately 60-70 nucleotide stem-loop intermediate, called the precursor miRNA (pre-miRNA). Exportin 5 and ran-GTP then transport the pre-miRNA out of the nucleus. In the cytosol, Dicer, which is also an RNase III endonuclease, recognizes and cuts the double-stranded portion of the pre-miRNA. This forms a very short-lived miRNA-miRNA duplex. The mature miRNA forms a complex with RNA-induced silencing complex (RISC) based on thermodynamic instability. The weaker stand is degraded; however, sometimes both miRNAs are equally viable and can complex with RISC[1]. The miRNA-containing RISC (miRISC) in humans was first discovered under the name miRNA ribonucleoprotein complex (miRNP)[17] and the two terms are generally accepted as synonymous. The miRISC includes helicase proteins and eIF2C2, a human Argonaute protein homolog[17]. Interestingly, Ago2 can process pre-miRNAs into mature miRNAs independently of Dicer[18]. This capability represents one of the many bourgeoning non-canonical pathways for miRNA biogenesis.

MiRNAs are involved in the development and differentiation of various tissues. MiR-152-3p has been implicated in hepatocyte differentiation. In a 2011 study by Kim, et.al., miR-152 was found to be among the top ten miRNAs which increases in expression during the transition from definitive endoderm to hepatocyte cell[19]. Our own experimental results from the differentiation of human induced pluripotent stem cells to hepatocyte-like cells (HLC) supported these findings. We also found miR-152 expression increased during stage 4, HLC production. This result is discussed more fully in Chapter 6, Future Directions Part II. Examples of miRNAs

in development include: miR-27b, which has been implicated with myogenesis[20], and miRs-124, -221, and -222, which are involved in neurogenesis[21].

MiRNAs are also heavily studied as biomarkers for diseases. For example, miR-1 is under investigation as a marker for cardiac disease and increased miR-122 is associated with liver cancer[1, 4]. Currently, there are few studies focused exclusively on miR-152. Much of the information present comes from miRNA screens done in search of miRs that are highly expressed in disease tissues. Data from such miRNA screens implicates miR-152 in hepatic glycogenesis regulation[22] and points to miR-152 as a biomarker for the early detections of breast cancer, colorectal carcinoma, and lung cancer[23]. In addition to this disease-specific data on miR-152 expression, the remainder of existing research on miR-152 examines its mRNA targets. A list of experimentally validated targets for the 3 prime *homo sapien* miR-152 (hsa-miR-152-3p) is included in Appendix II: Supplementary Data[24].

Histone Modifications

Accurate epigenetic maintenance and inheritance is important for the proper development and functioning of cells. Broadly, there are two type of chromatin: euchromatin and heterochromatin. Euchromatin results from the loosening of DNA around histones so transcription can occur. Heterochromatin exists when DNA is tightly wound around histones such that it cannot be accessed by transcription machinery. There are two types of heterochromatin: constitutive and facultative. As implied by the nomenclature, constitutive heterochromatin is permanent while facultative tends to be temporary. The transitions among these types of chromatin are due to modifications made to histones.

Histones can undergo dynamic post-translational modifications such as acetylation, phosphorylation, methylation, and ubiquitination. Many of these modifications alter the interaction of DNA with histones and can result in conformational changes from heterochromatin to euchromatin, or *vice versa.* This change in chromatin organization ultimately influences gene expression by affecting access to genes for transcription factors, DNA repair enzymes, DNA replication machinery, or enzymes involved in recombination. Proper epigenetic regulation is critical for normal cellular function and development. Research has shown that some epigenetic changes can even have transgenerational consequences[25, 26]. Thus, the study of epigenetics is important in understanding cellular processes and disease.

The Brno nomenclature system is employed when histone modifications are discussed in the literature. This system, first published by Turner in 2005, was devised following a meeting of the Epigenome Network of Excellence in Brno, Czech Republic, from which it gains its namesake. The Brno nomenclature system condenses a description of histone modifications into an easily readable and understandable abbreviation. The first letter, H, denotes histone, and the first number following indicates which histone is being described. For example, H3 means Histone 3. The second letter is the single-letter amino acid code designation and is paired with a number to indicate which amino acid residue gains modification(s). For example, K9 means lysine number 9. Finally, a short letter abbreviation is used to describe the modification(s). Accepted abbreviations include “me” for “methylation”, “ac” for “acetylation”, and “ph” for “phosphorylation”. The last number shown, if any, denotes how many times that modification is present at that particular residue. For example, me2 means dimethylated. The letter “a” or “s” might be written at the end of the Brno name to indicate either an “asymmetrical” or “symmetrical” addition, respectively[27].

Histone Methylation

Histone methylation can occur on arginine (R) or lysine (K) residues. One key difference between these types of methylation is the stability of these modifications; lysine methylation is more stable than that of arginine[28]. This is due to the presence of a destabilizing imino- group seen in the guanidino methyl[28]. When R methylation occurs, it takes place on the η_1 nitrogen of the amino-specific residue, creating either a monomethyl-, symmetric dimethyl-, or an asymmetric dimethylarginine[29]. Conversely, K methylation takes place on the N_ε of the amino-specific residue and K methylation has a lower turnover than R methylation[28].

In regards to function, methylation of arginine regulates protein-protein interactions. More specifically, R methylation facilitates the interaction of guanine and arginine-rich and proline-, glycine-, methionine-, arginine-rich motifs with proteins containing Tudor domains[29]. Tudor domain protein binding can be favorable or negative, depending on the quantity and conformation of R methylation[29]. By contrast, lysine methylation results in changes in gene expression. Examples of events which proceed from lysine methylation include heterochromatin formation and X-chromosome inactivation. These changes can be either silencing or activating, depending on the quantity and lysine position being methylated. For example, methylation of H3 at K4, K36, or K79 is usually associated with gene activation, while methylation of H3 at K9 or K27 and of H4 at K20 generally results in gene silencing[30]. In addition to these positions, lysine residues on H1 can also be methylated at various positions[28]. Lysine methylation can occur up to three times, resulting in mono-, di-, or trimethylated versions of each lysine residue.

Lysine Methyltransferases (KMTs) and Demethylases (KDMs)

Lysine methylation is accomplished by a class of enzymes called lysine methyl transferases (KMTs). KMTs transfer up to three methyl groups from a single S-adenosyl-L-methionine (SAM or AdoMet) at a time[31]. SAM is synthesized from methionine using AdoMet synthetase[32]. Upon transfer of SAM's methyl group to lysine by methylase, S-adenosylhomocysteine (AdoHyc) is formed[32]. AdoHyc is hydrolyzed to adenosine and homocysteine, the latter of which can be synthesized to methionine and used for further SAM synthesis in the SAM cycle[32].

The discovery of a lysine demethylase (KDM) is relatively recent, with the earliest published findings in 2004[33]. In the *Cell* article, Shi, et. al. identified lysine-specific demethylase 1 (LSD1), an H3K4 demethylase[33]. Since then, additional KDMs have been identified and their functions continue to be characterized. One such example is KDM2A, which was initially identified in 2005 as an H3K36 demethylase[34]. In 2016, research was published demonstrating KDM2A binds to heterochromatin protein 1 alpha (HP1α)[35] Given the role of HP1α in the maintenance of heterochromatin formation via increased KMT recruitment[8], it would be expected that KDM2A-HP1α binding may result in increased H3K9 methylation[35]. This work provides an example in which a demethylase for one H3 lysine residue, K36, may be responsible for stimulating methylation of another H3 lysine residue, K9. This area of research is fairly new and merits continued study in order to fully understand the complex roles of KDMs.

Structurally, all KMTs possess a highly conserved Su(var)3-9, Enhancer of Zeste, Trithorax (SET) domain in which the enzymatic activity occurs[31]. The only known exception to this structure is KMT4, also called DOT1L, which lacks a SET domain[31]. The KMTs are divided into families which are sequentially numbered KMT1 through KMT8[30]. The KMT1 family

includes H3K9me2 KMTs: KMT1C (G9a) and KMT1D (GLP)[30]. The KMT1 family also consists of H3K9me3 KMTs: KMT1A (SUV39H1), KMT1B (SUV39H2), KMT1E (SETDB1), and KMT1F (SETDB2)[30]. KMTs are of great importance due to their roles in gene expression. KMT1E, commonly referred to as SETDB1, is of particular interest.

Su(var) 3-9 Enhancer of Zeste, Trithorax Domain Bifurcate 1 (SETDB1)

The diverse functions of SETDB1 are part of what make this KMT particularly fascinating. SETDB1 trimethylates histone H3 at lysine 9 resides (H3K9me3) in euchromatin regions[6]. It has also been found to mono- and dimethylate H3K9 during translation[36]. Research also shows that SETDB1 monomethylates at H3K9 when H3K9 is trimethylated by Suv39 at pericentric heterochromatin[37]. Due to these distinctive functions bearing different consequences, the regulation of SETDB1 is especially intriguing.

Much of what is known about the structure and function of SETDB1 originates from a groundbreaking study by Schultz, et. al. from 2002. They found the C-terminal region of SETDB1 consists of three domains: pre-SET, SET, and post-SET. They characterized the SET domain as being bifurcate. This means the SET domain is split in parts (therefore creating pre-SET and post-SET domains) due to the presence of a 347 amino acid insert (Figure 1.1). This amino acid insert is conserved across species, from human to *C. elegans* and *D. melanogaster*, and does not affect the catalytic activity of SETDB1. SETDB1 is described as having two other key features, a KAP-1 interacting domain (KID) and a methyl-binding domain (MBD) (Figure 1.1). The KID facilitates proper KMT activity. However, KAP-1 binding is not required for HMT activity. KAP-1 targets SETDB1 to gene promoters transcriptionally silenced by a KRAB repressor protein. The MBD, more specifically, is a CpG-DNA methyl binding domain of the

MeCP2 family[38]. While SETDB1 contains this MBD, it is unknown if it actually binds methylated DNA[39]. According to Ensembl, the SETDB1 gene is located in chromosome 1 at position 150,926,263-150,964,744[40].

Figure 1.1 The domain organization of SETDB1 determined and depicted by Schultz, et. al[38]. SETDB1 has an amine KID. SETD1 also contains a MBD, and C-terminus SET domain. The SET domain is bifurcate with a conserved amino acid insert that serves no known purpose.

SETDB1 can be regulated exogenously with inhibitors of SAM such as DZNep[41]. DZNep inhibits AdoHyc hydrolase[41], thus disrupting the SAM cycle. Endogenously, post-translational modification of SETDB1, such as ubiquitination, deamination, and phosphorylation, regulates its HMT activity[42]. For example, it has been shown that lysine 867 ubiquitination is necessary for full H3K9 methytransferase activity[42]. While direct ubiquitination of SETDB1 enhances its HMT activity, ubiquitination which targets the protein for degradation could downregulate HMT activity. Another manner in which SETDB1 is regulated within the cell is by direct targeting from miRNAs. In humans, miRanda predicts that miR-296-3p, -324-3p, 152-3p, 92a-3p, and -1260b target SETDB1[14]. CLASH confirms binding of these miRNAs to SETDB1, though not necessarily at the same sites predicted by miRanda[2].

SETDB1 is of particular interest due to its role in diseases and epigenetics. SETDB1 is upregulated in Huntington's Disease and the increased SETDB1 is linked to increased H3K9me3. When SETDB1 is knocked down in HD mice, H3K9me3 levels are also reduced and a decrease in neuronal degradation is observed[10]. Upregulation of SETDB1 is also implicated in the metastasis of hepatocellular carcinoma (HCC). Upon knock-down of SETDB1, HCC proliferation and cancer cell migration are reduced[12]. Thus, modulation of SETDB1 with a miRNA has disease implications and potential therapeutic applications.

Histone 3 Lysine 9 (H3K9) Methylation

KMT1C (G9a/EHMT2) is responsible for monomethylation of H3K9[43] (Figure 1.2). Within eukaryotes, the role of H3K9me1 as either gene silencing or activating is split. For example, H3K9me1 serves as a mark of gene silencing in *Chlamydomonas reinhardtii*, a type of green algae. However, H3K9me1 is associated with gene activation in humans. One study of over 12,000 genes found that H3K9me1 was enriched at active promoters surrounding the transcription start site[44]. Another study indicates that H3K9me1 is localized to euchromatin regions of silenced genes[45]. Furthermore, there is evidence that H3K9me1 acts cooperatively with H4K20me1 to designate distinct regions of silent chromatin[46].

Dimethylation at histone H3 lysine 9 (H3K9me2) arises from KMT1C (G9a/EHMT2) and KMT1D (GLP/EHMT1)[30] (Figure 1.2) and is enriched in silent genes[44]. H3K9me2 is characteristic of X chromosome inactivation[47]. H3K9me2, in combination with H3K27me3, recruits a chromatin-associated protein called Cdyl. Cdyl serves as a binding partner for G9a, which then propagates H3K9me2 in X chromosome inactivation[48]. H3K9me2 is also localized to

euchromatin regions of silenced genes[45]. H3K9me1/2 is reversed by KDM families 1, 3, and 7[30] (Figure 1.2).

Trimethylation of H3K9 (H3K9me3) is accomplished by five different enzymes, including KMT1E (SETDB1) (Figure 1.2)[30] and is enriched in heterochromatin[45]. SETDB1 is specifically responsible for converting H3K9me2 to H3K9me3. H3K9me3 levels are higher than H3K9me1 in a 10kb region of the transcription start site surrounding silent genes[44]. Previous work supports the assertion that H3K9me3 is associated with heterochromatin formation and gene silencing by binding to heterochromatin protein 1 (HP1), which assists in the self-maintenance of heterochromatin[49]. H3K9me2/3 is only reversed by the KDM4 family[30].

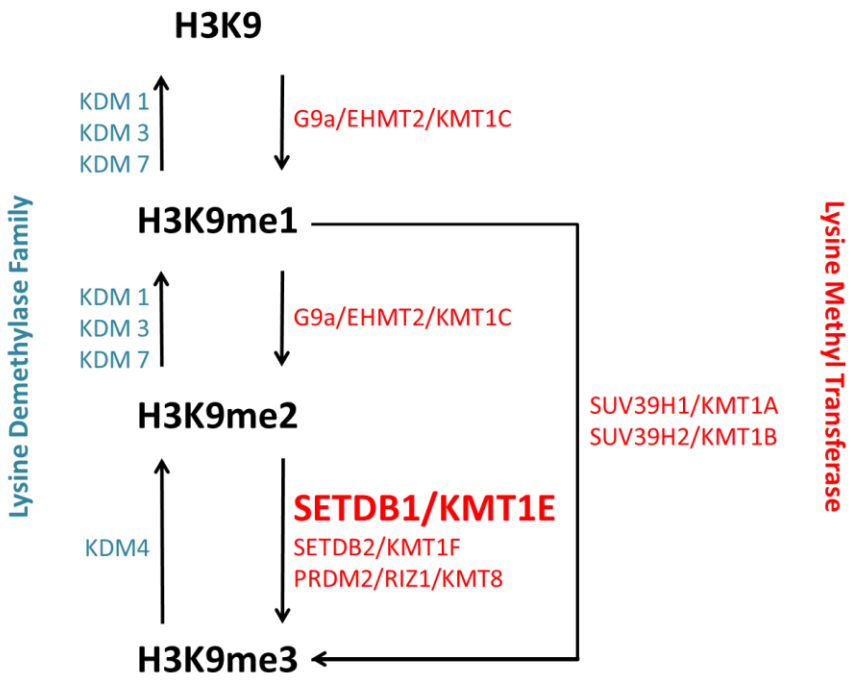


Figure 1.2 Summary of H3K9 methylation and demethylation enzymes. H3K9 mono- and dimethylation is accomplished only by G9a. H3K9me3 can result from any of the five KMTs listed. SETDB1 converts H3K9me2 to H3K9me3. Lysine demethylation enzyme families are also listed.

Heterochromatin Protein 1 (HP1)

Heterochromatin protein 1 (HP1) is a heterochromatin-associate protein involved in gene silencing. HP1 binds to methylated H3K9[49] and operates in a positive feedback loop. H3K9 methylation stimulates HP1 binding which then recruits more H3K9 methylation. This in turn forms massive amounts of heterochromatin which further propagates the self-renewal of HP1[50]. Structurally, HP1 proteins are characterized by a few conserved domains which each have different roles. The N-terminus chromodomain (CD) binds methylated H3K9, while the C-

terminus chromoshadow domain (CSD) is involved in interactions between proteins. Between the CD and CDS sits the hinge domain (HD), which binds DNA and linker histone. RNA binding at the HD is thought to target HP1 to heterochromatin[50]. In mammals, three isoforms of HP1 are found localized to various regions. HP1α is localized to pericentromeric heterochromatin, while HP1β and HP1ϒ are located at both pericentromeric heterochromatin and euchromatin, albeit to a lesser extent[8].

HP1β is associated with mono- and dimethylated H3K9, but not H3K9me3[51]. A knock out study found HP1β is critical for genome stability, as measured by significant increases in premature centromere divisions, polyploidy, diplochromosomes, and micronuclei detected through immunofluorescence[8, 52]. HP1ϒ, on the other hand, is important for the normal progression of mitosis[8]. As for HP1α, the localization has been found to be cell cycle dependent. As the nuclear envelop breaks down during mitosis, HP1α becomes diffuse, with little localization to centromeres. It moves between sister kinetochores and then to the spindle midzone. Specifically, HP1α relocated to centromeres on segregating chromosomes during anaphase. Once daughter nuclei are formed, HP1α distribution among heterochromatin is restored[53]. Because of these different roles, it is important to characterize the binding locations and functions of HP1 proteins.

CHAPTER 2

REGULATION OF SETDB1 EXPRESSION BY miR-152-3p

Introduction

In order to assess the effects of miR-152-3p on SETDB1, we chose to either inhibit or mimic the microRNA and study the resulting consequences at the transcript and protein levels in human cell lines. We introduced miR-152-3p mimic, a chemically synthesized RNA oligo which has a sequence identical to miR-152-3p, or anti- miR-152-3p, a chemically synthesized RNA oligo which has the complementary sequence to miR-152-3p, to cells. By increasing or inhibiting the miRNA, we can correlate the levels of miRNA to levels of predicted target protein with western blot analysis. If miR-152-3p represses translation, treating cells with anti-miR-152-3p would increase SETDB1 protein levels, while miR-152-3p mimic treatment would decrease SETDB1 protein levels.

It is also worth exploring not just the effect of miR-152-3p on SETDB1 protein, but also on SETDB1 transcript. If the microRNA affects mRNA stability, significant changes at the transcript level would correspond to changes at the protein level. Additionally, there is at least one instance of a miRNA targeting DNA and thus affecting transcription. MiR-324-3p has been found to target the RelA promoter and induce its expression in an Ago2-mediated way[5]. In order to determine if miR-152-3p has any effect on SETDB1 transcript levels, we conducted a qPCR following anti-miR-152-3p and miR-152-3p mimic treatments.

We also explored the effects of anti-miR-324-3p on cells due to its high expression in HCC as compared to the other four experimentally confirmed miRs that target SETDB1 in humans[24] (Figure 2.1). MiR-324-3p is predicted to target SETDB1 as well. We aimed to compare the effects of miR-152-3p and miR-324-3p on SETDB1.

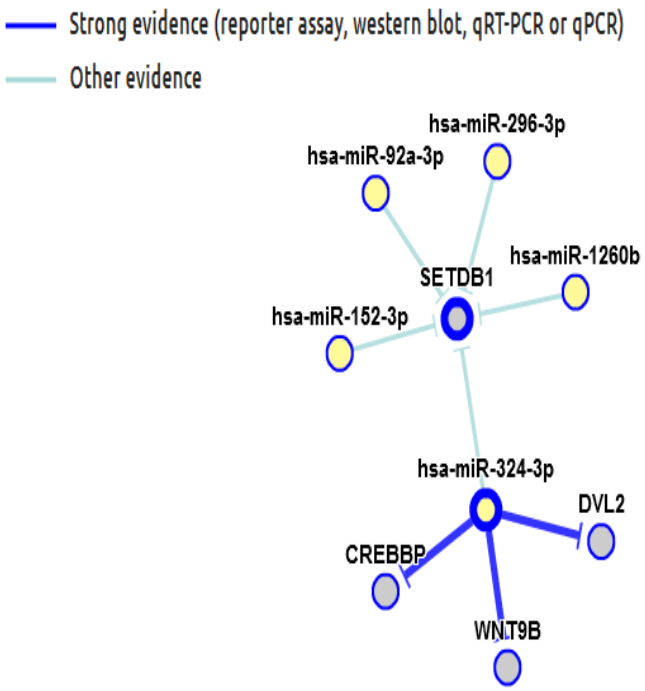


Figure 2.1 MicroRNAs experimentally determined to target SETDB1. According to this tree diagram, sourced from miRTarBase, there are five experimentally confirmed miRNAs that target SETDB1. The three other miR-324-3p targets, aside from SETDB1, are also shown[24].

HepG2, a human hepatocellular carcinoma (HCC) line, and IMR-91, a normal human skin fibroblast line were chosen for studying miR-152-3p and SETDB1 due to the presence of the specific miRNA, mRNA, and protein in the cell lines[54, 55]. Both miR-152 and SETDB1 have

also been heavily implicated in HCC[12, 23, 54, 56], so IMR-91 was chosen as a normal human cell line model. HepG2 and IMR-91 have the advantages of being robust, easy to culture, cell lines.

Methods

miRNA Inhibitor and Mimic Transfections

Cells were transfected in parallel with water in lieu of anti-miR or mimic as a negative control. Multiple time points were repeatedly tested to establish the kinetics of this effect. Cells were initially assayed 24 hours apart for 3 days upon the user guide suggestion to examine cells 6-72 hours after transfection. When a difference was observed at 24 hours, cells were assayed every 6 hours for 24 hours.

HepG2 and IMR-91 cells were treated with 50 nM anti-miR-152-3p, 50 nM anti-miR-324-3p, 5 nM miR-152-3p mimic, or water on the day of passage using the HiPerFect transfection reagent according to the Qiagen manual instructions. Briefly, cells were seeded in cell culture treated plates coated with 0.1% gelatin for IMR-91 cells, or without coating for HepG2 cells, and incubated under normal growth conditions (37°C and 5% CO_2) while preparing transfection complexes. 20 µM stock inhibitor, 20 µM stock mimic, and/or water were diluted in serum free culture media (IMR-91: DMEM with 4.5 g/L glucose and 100X GlutaMax; HepG2: DMEM with 1.0 g/L glucose). HiPerFect transfection reagent was added to a final volume of 3% and tubes were mixed using a vortex. Transfection complexes were formed during a 5-10 minute room temperature incubation before being added drop-wise onto cells.

Western Blots for SETDB1 Protein

Total protein was collected in RIPA buffer 6-72 hours after transfection, quantified using a Lowry assay, and 13.5 µg (Figure 2.2A), 15 µg (Figures 2.2B, 2.3A, and 2.5), or 30 µg (Figures 2.2C and 2.3B) of protein was loaded on 8% polyacrylamide gels. The gels were run at 120V at 4°C and transferred to either nitrocellulose (Figures 2.2C and 2.3B) or PVDF (Figures 2.2A, 2.2B, and 2.3A) at 4°C in 40% methanol transfer buffer at 300 mA for 2 hours. Membranes were blocked in 5% milk/TBST for 1-16 hours and incubated in anti-SETDB1 primary antibody at 1:1000 (Figures 2.2A, 2.2B, and 2.3A) or 1:500 (Figures 2.2C and 2.3B) for 12-16 hours in 5% milk/TBST. Membranes were incubated in anti-mouse secondary antibody for 1 hour at room temperature then visualized under ECL conditions using a Kodak developer.

As a normalization control, membranes were re-probed with anti-β actin primary antibody for 1 hour at room temperature, then with anti-mouse secondary for 1 hour at room temperature (Figures 2.2A, 2.2C, 2.3A, and 2.3B). Alternatively, either 2 µg (Figure 2.2B) or 1.5 µg (Figure 2.5) of total protein was loaded on a 10% gel, probed with anti-β actin primary antibody for 1 hour at room temperature, then with anti-mouse secondary for 1 hour at room temperature.

qPCR for SETDB1 mRNA

Cells were plated in triplicate for each condition. Total RNA was collected from trypsinized cells 6-24 hours after treatment using Qiagen's RNeasy Mini Kit (# 74104) and quantified with a nanodrop. BioRad's iScript kit was used to generate 1 µg cDNA/sample. 10 ng sample and the TaqMan system were used for qPCR. Samples were plated in triplicate and run in an Applied Biosystems 7500 machine. Technical replicate C_T values were averaged for SETDB1

and β-ACTIN (βACT). βACT values were subtracted from SETDB1 to generate ΔC_T. $\Delta\Delta C_T$ was found by subtracting the 0 hour ΔC_T from other sample ΔC_Ts in the time course experiment. In the single-time point experiment, the water ΔC_T was subtracted from SETDB1. Fold change was then calculated based on the $\Delta\Delta C_T$ obtained. 2-way ANOVA with Bonferroni post-test was performed to determine statistical significance.

Results

Effects of miR-152-3p Inhibitor and Mimic on SETDB1 Protein Levels

Preliminary experiments showed that SETDB1 protein expression decreased 24 hours after HepG2 cells were transfected with 50 nM anti-miR-152-3p compared to the control (untreated) sample (Figure 2.2A). In order to identify when the effect was most pronounced, a time-course experiment was done. Samples were collected every 6 hours for 24 hours post-transfection and analyzed for SETDB1 protein expression. At 6 hours, SEDB1 protein expression was below the detection threshold of the western blot (Figure 2.2B). We chose the 12 hour time point for subsequent experiments due to the unambiguous reduction in SETDB1 as compared to the control (dH_2O-treated) samples, which still fell within the detectable range. These results were also seen in IMR-91, a normal human skin fibroblast cells line (Figure 2.3A). HepG2 cells transfected with miR-152-3p mimic showed an increase in SETDB1 protein at 12 hours after treatment (Figure 2.2C). The same results were also seen in IMR-91 cells (Figure 2.3B).

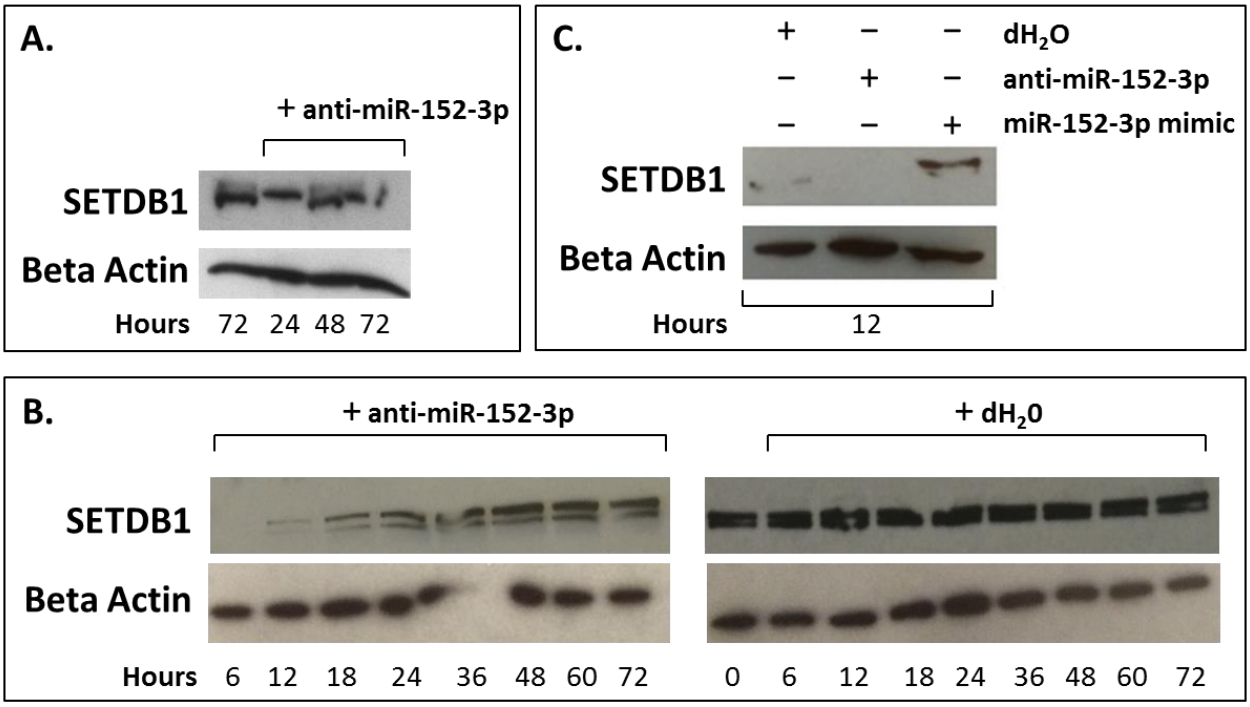


Figure 2.2 SETDB1 protein expression in HepG2 cells transfected with anti-miR-152-3p, miR-152-3p mimic, and dH_2O. HepG2 cells were single-transfected with the described treatment and then collected (**A**) 24 hours apart for 3 days total (**B**) 6 hours apart for 24 hours and then 12 hours apart for a total of 72 hours (**C**) at a single 12 hour time point. Samples were loaded on 8% polyacrylamide gels and transferred to either NC or PVDF and then probed for SETDB1 and beta-actin. The water-treated and untreated samples serve as controls.

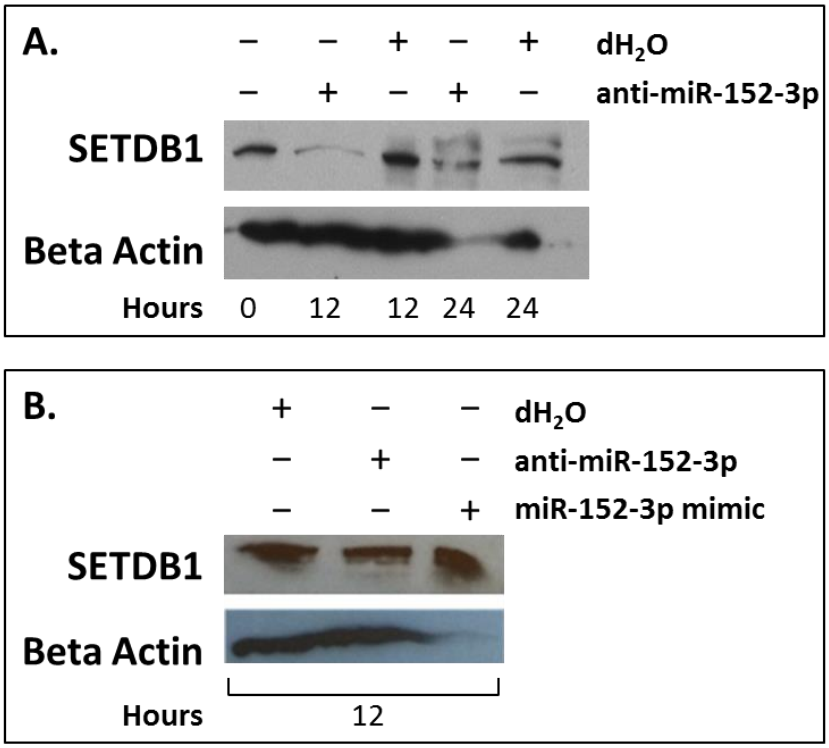


Figure 2.3 SETDB1 protein expression in IMR-91 cells transfected with anti-miR-152-3p, miR-152-3p mimic, and dH_2O. IMR-91 cells were single-transfected and total protein was collected (**A**) every 12 hours after treatment for 24 hours and (**B**) at a single 12 hour time point. Samples were loaded on 8% polyacrylamide gels and transferred to either NC or PVDF and then probed for SETDB1 and beta-actin. The water-treated and untreated samples serve as controls.

Effects of miR-152-3p Inhibitor and Mimic on SETDB1 mRNA Levels

In order to determine if miR-152-3p has any effect on SETDB1 mRNA, we quantified the steady state levels of SETDB1 mRNA following miR-152-3p manipulation. We transfected cells with inhibitor, mimic, or water. We collected RNA every 6 hours for 24 hours after a single

transfection to see if there was any change in SETDB1 mRNA levels in HepG2 cells from the 0 hour, untreated, sample (Figure 2.4, panels A and B). We also collected RNA from both cell lines 12 hours after a single transfection to determine if there was any change in SETDB1 mRNA levels compared to the water-treated control sample (Figure 2.4, panel C). We found no significant difference between SETDB1 mRNA levels in dH_2O-treated, anti-miR-152-3p-treated, or miR-152-3p mimic-treated cells in both cell lines (Figure 2.4).

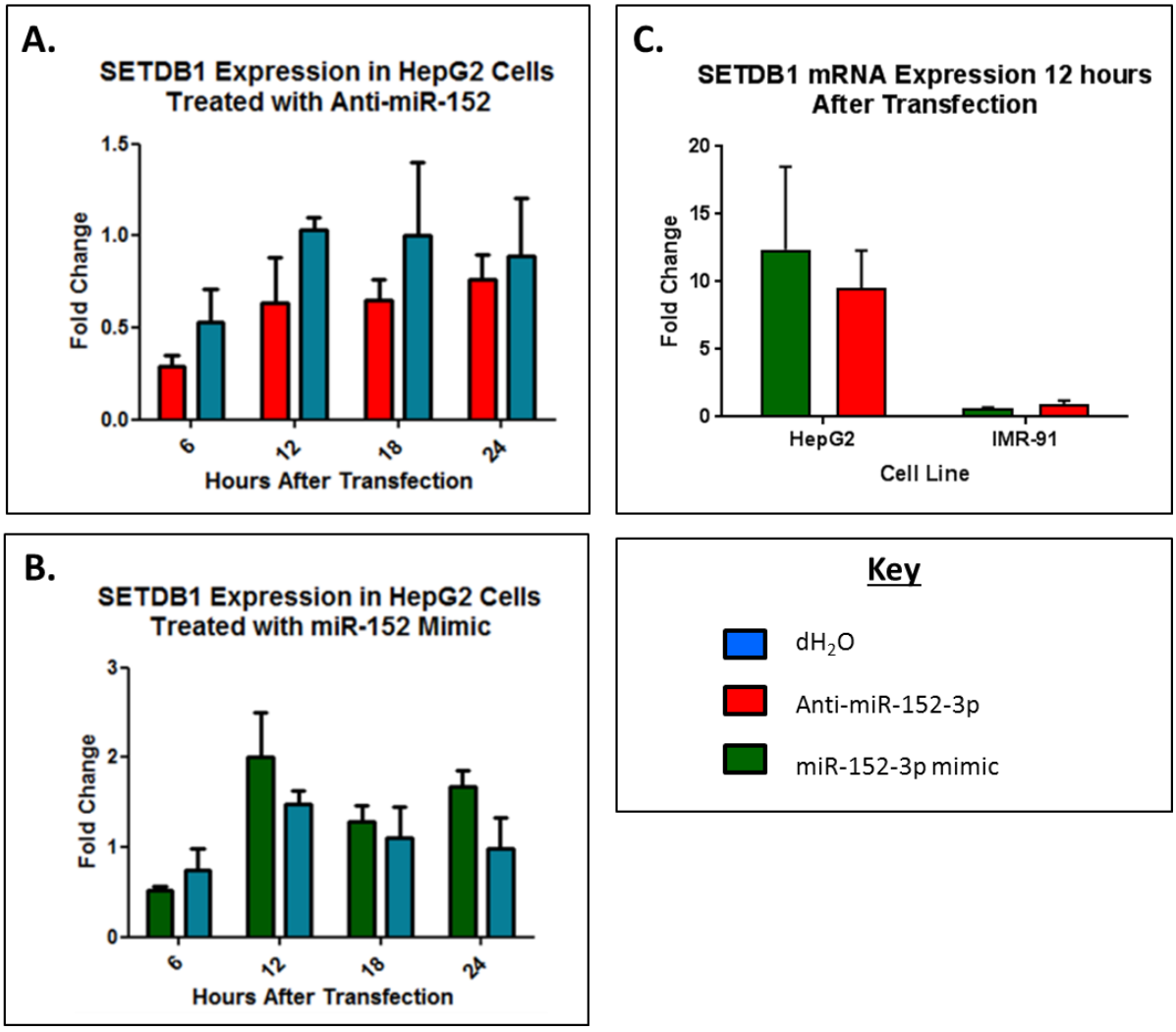


Figure 2.4 SETDB1 mRNA expression following anti-miR-152-3p and miR-152-3p mimic treatment. (**A**) There was no significant difference in SETDB1 mRNA expression between anti-miR-152-3p and dH_2O treated HepG2 cells across time points. (**B**) There was no significant difference in SETDB1 mRNA expression between miR-152-3p mimic and dH_2O treated HepG2 cells across time points. (**C**) There was no significant difference in SETDB1 mRNA expression with anti-miR-152-3p or miR-152-3p mimic treatment as compared to dH_2O treatment in HepG2 and IMR-91 cells.

Effects of miR-324-3p Inhibitor on SETDB1 Protein Levels

We treated cell with miR-324-3p inhibitor to assess the effects of another miRNA thought to target SETDB1. HepG2 and IMR-91 cells transfected with 50 nM miR-324-3p inhibitor were assayed 12 and 24 hours after treatment and no decrease in SETDB1 protein was observed (Figure 2.5).

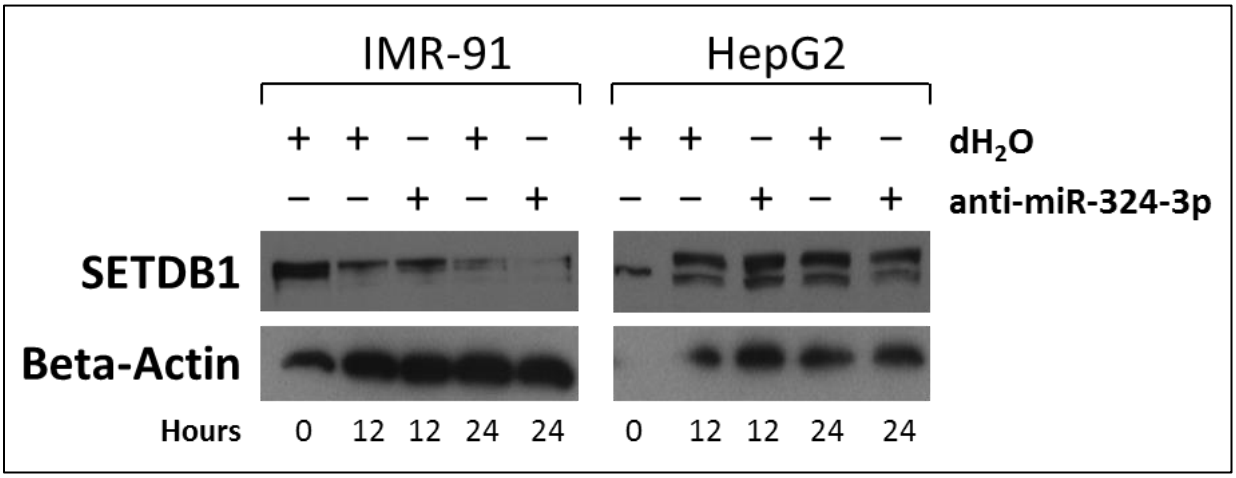


Figure 2.5 SETDB1 protein expression in cells transfected with anti-miR-324-3p and dH_2O. Total protein was collected from cells every 12 hours for 24 hours following a single transfection with either anti-miR-324-3p or water. Samples were loaded on 8% polyacrylamide gels and transferred to PVDF and then probed for SETDB1 and beta-actin. The water-treated and untreated samples serve as controls.

Conclusions and Discussion

Experimental evidences obtained by western blot analysis of SETDB1 protein expression and qPCR of SETDB1 mRNA lead us to conclude miR-152-3p positively regulates SETDB1 protein synthesis. We determine miR-152-3p has no effect on steady-state levels of SETDB1 mRNA since qPCR shows no significant changes in SETDB1 mRNA steady-state levels with and without miR-152-3p mimic and anti-miR-152-3p in both cell lines. Since SETDB1 mRNA levels do not correspond to the pattern of expression seen in SETDB1 protein levels, we believe miR-152-3p stimulates synthesis of SETDB1 protein, rather than of SETDB1 mRNA. Our western blot results for cells transfected with anti-miR-324-3p indicate miR-324-3p may not target SETDB1.

The effects of miR-152-3p mimic and anti-miR-152-3p were most prevalent at 12 hours following treatment, but can be seen up to 24 hours post-transfection. Our western blots show that SETDB1 protein levels return to basal levels over the course of 72 hours after a single transfection with anti-miR-152-3p. The transient nature of these treatments serves as a technical limitation in studying the role of miR-152-3p in cells. Because the inhibition or increase of miR-152-3p is not constitutive, we can only study the effects within the effective window of the treatments. We expect that if the miR-152-3p functions by enhancing SETDB1 transcription or mRNA stability, anti-miR-152-3p treatment would cause a significant decrease in expression of SETDB1 mRNA, and miR-152-3p mimic transfection would cause a significant increase in

steady-state levels of SETDB1 transcript. Inhibition of miR-324-3p has no effect on amounts of SETDB1 protein, as seen in western blots for SETDB1 protein. One of the limitations in this experiment is the time scale at which we assayed for changes in SETDB1 protein following miR-324-3p inhibition.

CHAPTER 3

EXPLORING THE REGULATION OF miR-152-3p WITH THE SETDB1 3'UTR

Introduction

Currently, the only experimental evidence demonstrating SETDB1 is a miR-152 target originates from a cross-linking, ligation, and sequencing of hybrids (CLASH) study which found a binding site for miR-152 in the coding region of SETDB1[2] (Figure 3.1A). While the technique of CLASH boasts many advantages such as a < 2% rate of chimeric reads and the ability to monitor live cell interactions[2], it fails to distinguish biologically meaningful interactions from those which may be transient. This is because CLASH involves crosslinking, which may show binding of a miRNA to targets it does not actually act on. Thus, it was in our interest to further explore the interaction between miR-152 and SETDB1 and determine if it their binding represents a biologically relevant interaction with meaningful consequences in the cell. We chose to examine the 3'UTR because the experimentally determined binding site for miR-152-3p on SETDB1 is in a non-canonical region. Moreover, the miRanda database predicted three binding sites for miR-152-3p on the 3'UTR of SETDB1[14] (Figure 3.1B).

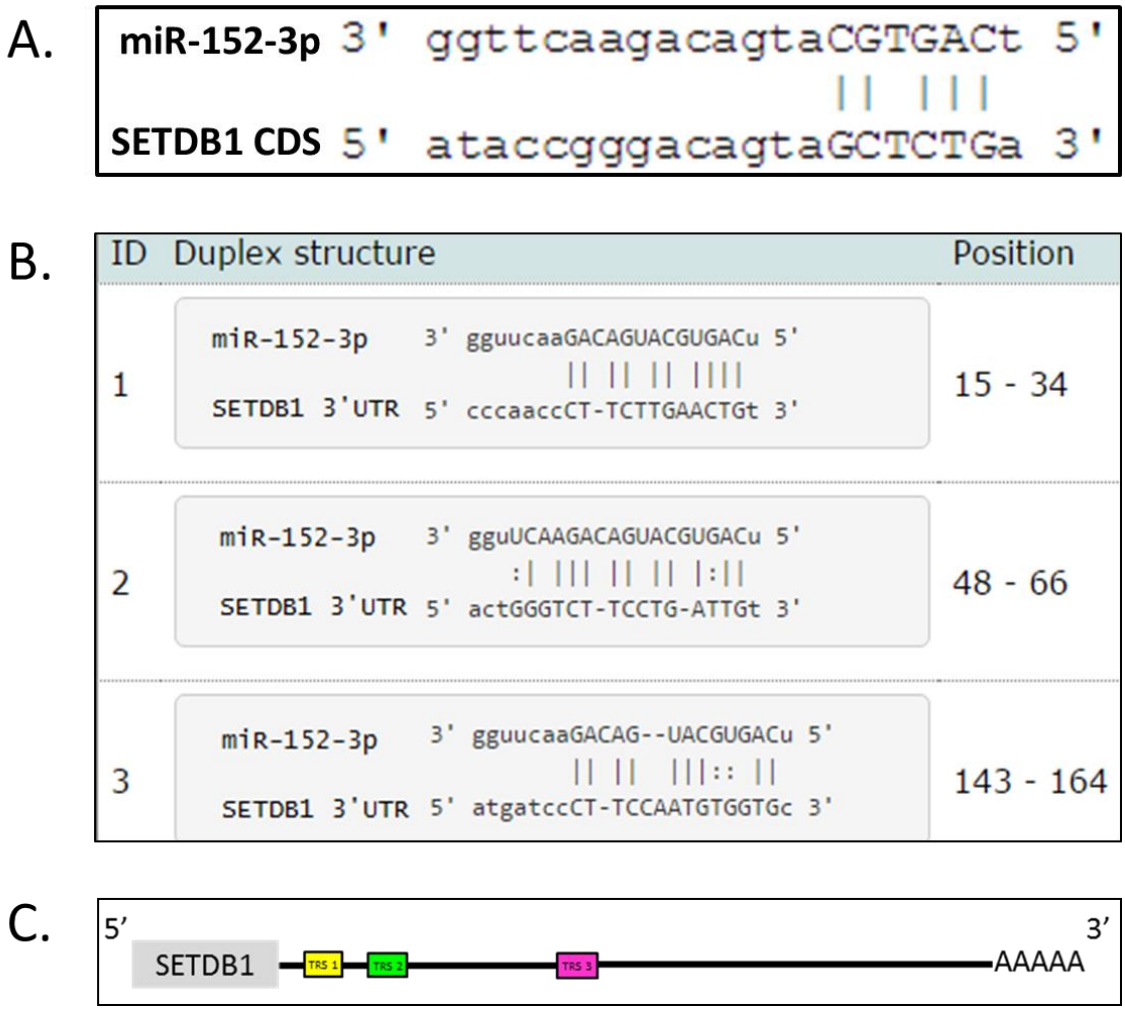


Figure 3.1 Confirmed and predicted miR-152-3p binding sites on SETDB1. (**A**) CLASH-confirmed binding site in the coding region of SETDB1[2]. (**B**) MiRanda Database target recognition sequences (TRS) in the 3'UTR of SETDB1 as shown on miRTarBase[14, 24]. **(C**) SETD1 3'UTR structure with TRS 1 (yellow), TRS 2 (green), and TRS 3 (pink) annotated.

We used a reporter assay to analyze the 3'UTR target sites. Each of the predicted target recognition sequences (TRS) were individually assessed by deleting those sequences in a luciferase reporter construct that contains the SETDB1 3'UTR. Beyond solely confirming the

predicted 3'UTR binding sites, our reporter assay experimental design could also be interpreted for directionality of targeting. For instance, addition of a 3'UTR to the reporter vector should decrease luciferase production due to the introduction of miRNA activity, which destabilizes the luciferase protein. When actual binding sites of a repressive miRNA are mutated, the luciferase levels should increase as compared to the non-mutated. Conversely, if the binding site of an enhancing miRNA is disrupted, luciferase levels should decrease. In the case that a TRS is not real, there would be no change in luciferase upon mutation of the sequence. Two alternative methods of mutating the binding sites are complete deletion of the site and seed sequence disruption, which includes base substitution or base deletion.

Methods

Cloning, Transfection, and Luciferase Assay

The backbone vector pmirGLO, also referred to as the empty vector, was purchased from Promega (Figure 3.2). It contains firefly luciferase (shortened to luciferase, L, or firefly), which is expressed at different levels depending on the presence, absence, or mutation of an insert. Luciferase serves as the experimental readout. The vector also contains *Renilla* luciferase (shortened to *Renilla*, *R*), which is expressed regardless of miRNA activity. *Renilla* serves as a readout for transfection efficiency. Vector pmirGLO is ampicillin resistant and has a multiple cloning site where a miRNA target sequence can be cloned in. The miRNA target sequence chosen was the SETDB1 3'UTR. It was generated from a HepG2 total RNA sample synthesized

into cDNA and PCR amplified using primers specific to the SETDB1 3'UTR and vector. The sequences generated by the InFusion Cloning Primer Design Tool are as follows: forward primer = 5'- CTA GTT GTT TAA ACG AGG ACA GCC TTC TTC CCA - 3' and reverse primer = 5' – GCA GGT CGA CTC TAG CAC AAT CTT ACT TTT ATC ATA TT - 3'. The insert was cloned into pmirGLO using the InFusion Cloning System (Takara).

The SETDB1 3'UTR insert is 367 base pairs and was analyzed for size on a 1% agarose gel. Vector linearity was confirmed through agarose gel electrophoresis as well. Once ligated to the linearized vector, the identity of the insert was verified through DNA sequencing with an Applied Biosystems 3500 Genetic Analyzer (Appendix II: Supplementary Data).

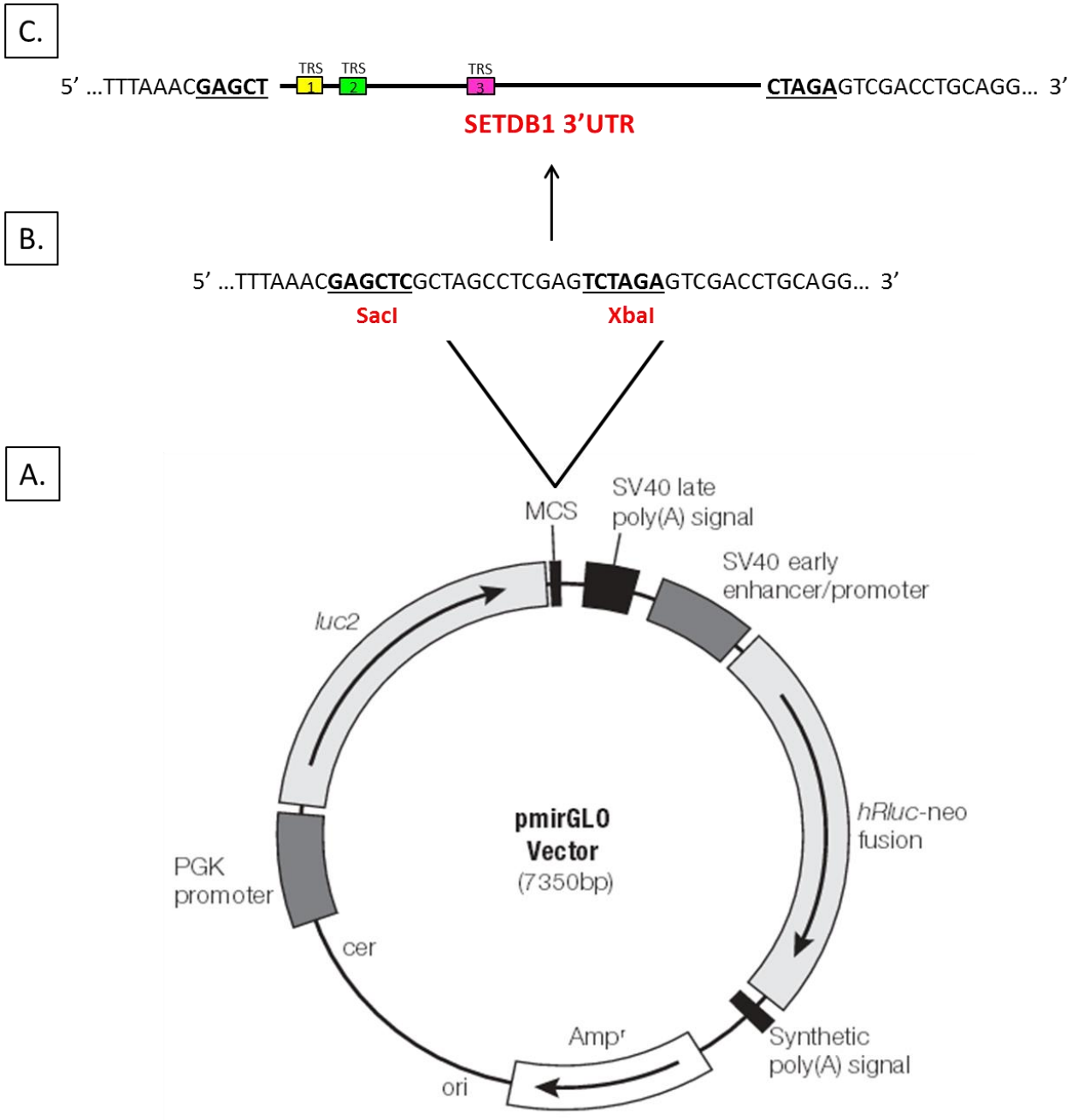


Figure 3.2 Overview of SETDB1 3’UTR reporter vector construction. (**A**) Original vector map for pmirGLO purchased from Promega (#e1330). (**B**) Vector was linearized by digestion with SacI and XbaI restriction enzymes in the multiple cloning site (MCS). (**C**) The final MCS sequence contained the 3’UTR from SETDB1. Cut sites were not preserved. Luciferase expression is driven by

the PGK promoter. Luciferase levels depend on activity occurring at the 3'UTR of SETDB1. The SV40 early enhancer/promoter drives *Renilla*. *Renilla* is expressed in every cell transfected with the vector.

Site-directed mutagenesis was performed by designing primer sets to omit the TRS through PCR amplification. The forward primer is identical to the sequence immediately following the deletion site. The reverse primer is reverse complementary to the sequence directly upstream the deletion site and includes a 15 base pair overlap following the deletion site (Table 3.1). Colonies were screened for successful binding site deletion using a PCR approach. Primers designed to generate a 100 base pair product in the absence of a deletion and approximately 80 base pair product in the presence of a deletion helped us determine which mini preps from colonies were likely to contain the desired mutation. The final vector sequences were confirmed through DNA sequencing with an Applied Biosystems 3500 Genetic Analyzer (Appendix II: Supplementary Data).

Primer Name	Primer Sequence (5' → 3')	SETDB1 3'UTR Sequence Deleted
pE1330S3Upbs1F	CGT TTC CTC AGG AAC TGG GTC TT	CCCAACCCTTCTTGAACTGT
pE1330S3Upbs1R	GTT CCT GAG GAA ACG AAG AAG GCT GTC CTC TAA AGA AGA C	
pE1330S3U2F	TGA ACC CTG ACC CGA AGT CT	ACTGGGTCTTCCTGATTGT
pE1330S3Upbs2R	TCG GGT CAG GGT TCA TCC TGA GGA AAC GAC AGT TCA AGA A	
pE1330S3Upbs3F	TAG CAG GCA GGA TCC CTT CTC	ATGATCCCTTCCAATGTGGTGC
pE1330S3Upbs3R	GGA TCC TGC CTG CTA CTG GTC CAG AAC CCC CAT TTC	

Table 3.1 Primer sets designed to generate target recognition sequence deletions within the 3'UTR of SETDB1. The primer name listed is the name the

primer was order under and is labeled as in the -20°C freezer where it is stored. The primer sequence provided was created based on the rules for generating site-directed mutagenesis using the InFusion Cloning Kit (Takara). Each primer pair generated one of three deletions, as shown in this table.

HepG2 and IMR-91 cells were plated on day 0 at 10,000 cells per well of a white-wall 96-well cell culture treated dish with 0.1% gelatin coating for IMR-91 cells only under aseptic conditions and allowed to adhere overnight in a 37°C humidified incubator with 5% CO_2. On day 1, cells were transfected with DNA using Lipofectamine 2000 according to the product protocol. Transfection parameters were first optimized for timing and reagent concentrations by treating cells with 100 ng DNA diluted in Opti-MEM and the following amounts of Lipofectamine 2000 reagent diluted in 25 µL Opti-MEM: 0.2 µL, 0.3 µL, 0.4 µL, or 0.5 µL. Optimal transfection parameters were determined by comparing the empty vector background subtracted luciferase/*Renilla* ratio to the vector with SETDB1 3'UTR background subtracted luciferase/*Renilla* ratio 24 hours after transfection (Appendix II: Supplementary Data). It was determined that 0.4 µL Lipofectamine was the ideal quantity to use and all other transfections were done with this amount. After optimization, HepG2 cells were evaluated for miRNA-mediated effects on luciferase expression by comparing the empty vector background subtracted luciferase/*Renilla* ratio to the vector with SETDB1 3'UTR background subtracted luciferase/*Renilla* ratio. The ratio of the empty vector represented a maximum value.

On day 2, cells were assayed using the Dual-Glo Luciferase Assay System from Promega. Cells and reagents were equilibrated to room temperature. Dual-Glo reagent was added to each well and incubated at room temperature for at least 10 minutes. During this time, cells

lysed, allowing for firefly luminescence to be measured using an LMAX II Luminometer (Molecular Devices). Next, Stop and Glo Reagent was added to each well. The plate was again incubated at room temperature for at least 10 minutes before measuring *Renilla* luminescence on the luminometer. This protocol is summarized in Figure 3.3.

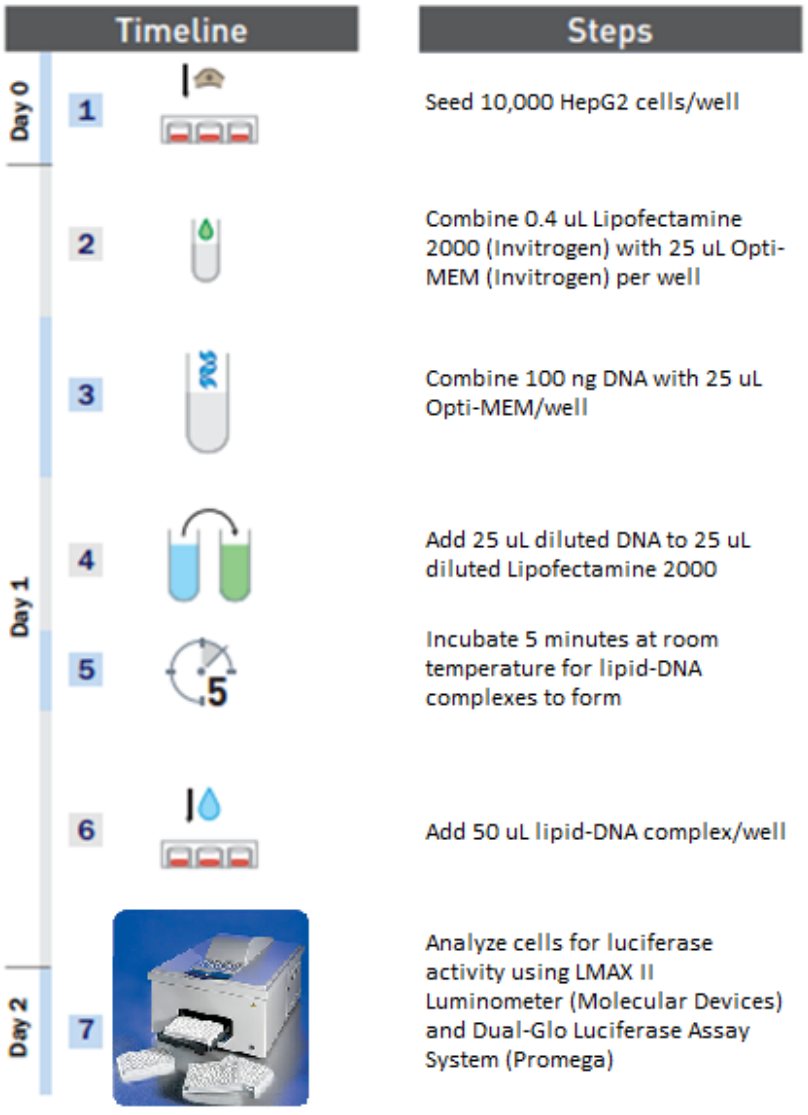


Figure 3.3 Timeline and overview of protocol for transfecting HepG2 cells with reporter vector. This image was modified from the Lipofectamine 2000

protocol provided by Invitrogen. Cell seeding occurred 1 day prior to cell transfection and luminescence was read 1 day following transfection.

In order to account for non-specific fluorescence as a result of auto fluorescence or reagent fluorescence, wells on each plate were mock transfected using Opti-MEM only. These cells were otherwise treated identically to DNA-transfected cells by using the same growth conditions. When Dual-Glo was applied, background luciferase values were obtained. Addition of Stop and Glo created a *Renilla* background value. The luciferase background values were averaged for each trial and subtracted from each well's luciferase value. Next, the background subtracted luciferase value/*Renilla* was taken. This luciferase/*Renilla* (also referred to as firefly/*Renilla*) ratio allows us to compare values across trials. To ensure the reproducibility and reliability of background subtracted luciferase/*Renilla* ratios, a minimum of three independent trials were done with a minimum n=3 per condition in each trial.

Results

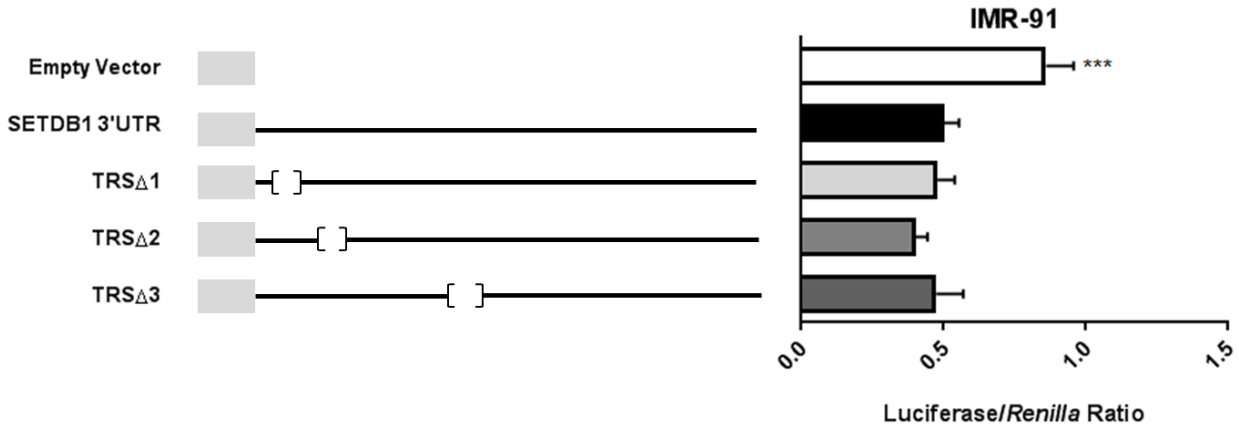
Experimentally Assessing Computationally Predicted miR-152-3p binding sites in the 3'UTR of SETDB1

The empty vector firefly/*Renilla* luminescence ratio represented the absence of all miR activity, as there was no miR target site present. By adding in the 3'UTR from SETDB1, miRNAs could potentially bind to all available target sites within the region. An unpaired t-test

showed the firefly/*Renilla* luminescence ratio was significantly lower in HepG2 ($p < 0.0001$) and IMR-91 ($p < 0.001$) upon addition of the SETDB1 3'UTR, representing miR activity across the entire UTR. Our data shows miR binding has been disrupted in two of the three miR-152-3p sites predicted by miRanda to exist in the 3'UTR of SETDB1 in HepG2 cells only.

In IMR-91 cells, there were no significant changes ($p > 0.01$) in firefly/*Renilla* luminescence with any of the target recognition sequence deletions. In HepG2 cells, when target recognition sequence 1 was removed (TRSΔ1), a significant decrease ($p < 0.001$) in firefly/*Renilla* luminescence was observed. When target recognition sequence 2 was deleted (TRSΔ2), no significant changes occurred. When the third target recognition sequence was deleted (TRSΔ3), the firefly/*Renilla* luminescence ratio was significantly higher ($p < 0.01$) than the vector with un-mutated SETDB1 3'UTR (Figure 3.4).

A.



B.

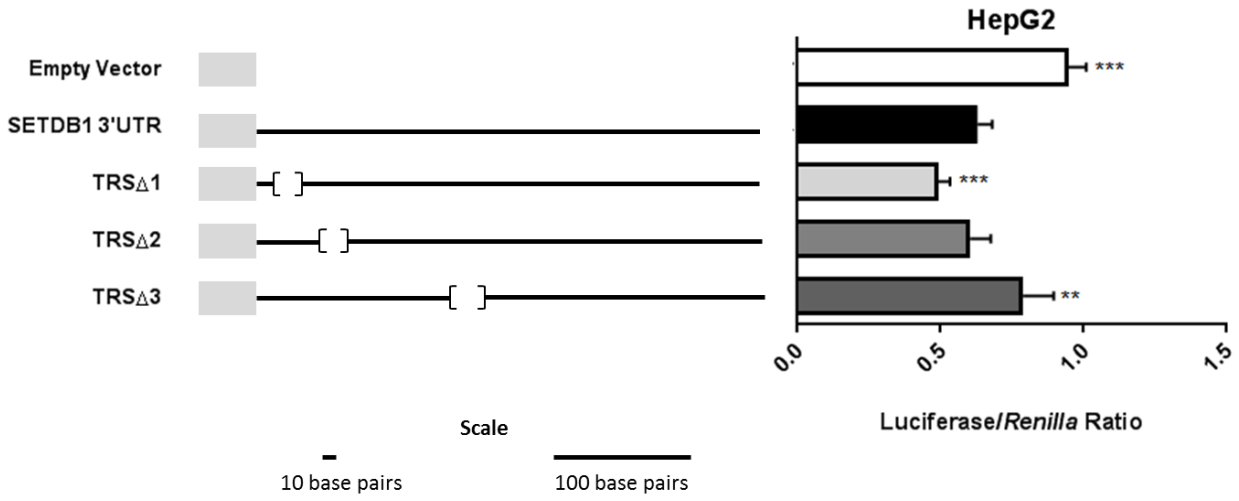


Figure 3.4 Sequence requirements for regulatory activity on the 3'UTR of SETDB1. For all data, significance was determined with unpaired t-test to vector containing SETDB1 3'UTR. (**A**) In IMR-91 cells, there was a significant decrease ($p < 0.001$) in luciferase/*Renilla* luminescence ratio (L/*R*) upon addition of the SETDB1 3'UTR. (**B**) In HepG2 cells, there was also a significant decrease ($p < 0.0001$) in L/*R* upon addition of the SETDB1 3'UTR. L/*R* was significantly lower ($p < 0.001$) with the TRSΔ1 mutation, and significantly higher ($p < 0.01$) with the TRSΔ3 mutation.

Conclusions and Discussion

We provide the first experimental evidence of miRNA regulatory activity on the 3'UTR of SETDB1. Overall, addition of the 3'UTR to the empty vector resulted in a significant ($p < 0.001$) decrease in firefly/*Renilla* as compared to the empty vector alone. This change, which was observed in both cell lines, represents the net effect of all regulatory activity on the entire 3'UTR of SETDB1. We found the firefly/*Renilla* value decreased with target recognition sequence (TRS) 1 deletion and concluded TRS 1 is required for regulation of SETDB1 in HepG2 cells. Deletion of TRS 2 caused no significant changes in luciferase as compared to the vector containing intact 3'UTR. Thus, TRS 2 may not be required for regulation of SETDB1 in HepG2 cells. When TRS 3 was deleted, the firefly/*Renilla* luminescence ratio was significantly higher ($p < 0.01$) than the vector with un-mutated SETDB1 3'UTR in HepG2 cells.

Deletion of the individual target recognition sequences presented different results in each cell line. Since the vectors were the same and the only difference was the cell type, these results may signify a cell type-specific finding. Our data might reflect different miRNA content and expression in different cell lines. In IMR-91 cells, deletion of target recognition sequences 1 and 3 caused no significant change in luciferase/*Renilla* activity. A possible explanation would be that these are not true binding sites for miR-152-3p. However, removal of TRS 1 or TRS 3 resulted in significant changes ($p < 0.001$ and $p < 0.01$, respectively) in luciferase/*Renilla* in HepG2 cells. This paradox leads us to question how these cell line differences can arise. One possibility is that the predominant binding sites for miR-152-3p may be different in different cell lines. Perhaps miR-152-3p largely targets the CDS in some cells and the 3'UTR in others. This is feasible since other miRNAs bind in these regions, but expression of them can vary across tissue

types. Thus, in one cell line, less miR-152-3p may bind in the 3'UTR due to greater activity from other miRNAs binding at that same location.

Another possible interpretation is that there could be different expression levels of different miRNAs in the different cell lines. This possibility seems likely given that there is evidence of tissue-specific expression of miRNAs[1]. Moreover, there could even be the same expression of different miRNAs which target those regions with different functions. For example, if a repressive miRNA is expressed highly in IMR-91 and an enhancer miRNA is expressed highly in HepG2 cells, both of which target the deleted site, we may still obtain significant changes in luciferase with site-directed mutagenesis in HepG2 cells and no changes in luciferase with site-directed mutagenesis in IMR-91 cells. It is worth noting that there are presently no predicted or confirmed repressive miRNAs targeting TRS 1, and searches on RNA22 and microRNA.org yield no miRNAs which could bind to the TRS 1 sequence (5'-CCCAACCCTTCTTGAACTGT-3').

Finally, another potential interpretation is that if there are different expression levels of the same miRNA in HepG2 and IMR-91 cells, we may still see cell line differences. For example, when looking at the predicted miRNAs binding at TRS 3 (Figure 3.5), if it were confirmed that all three miRNAs, miR-152-3p, miR-1260b, and miR-324-3p bind at TRS 3, but the levels were different of these miRNAs in each cell line, site-directed mutagenesis of TRS 3 might cause significant changes in luciferase in HepG2 cells and no changes in luciferase with site-directed mutagenesis in IMR-91 cells. As a hypothetical scenario under this potential interpretation, if miR-1260b is SETDB1 repressive in both cell lines, but is highly expressed in HepG2 cells and expressed at very low levels in IMR-91, and miR-324-3p and miR-152-3p

levels are constant in both cell lines, there may be a significant increase in luciferase with site-directed mutagenesis of TRS 3 in HepG2 cells with no change in IMR-91 cells.

Deletion of TRS 1 or TRS 3 was found to bear significant effects on luciferase/*Renilla* ($p < 0.001$ and $p < 0.01$, respectively) when compared to the empty vector in HepG2 cells. However, their effects were significant in opposing directions. In order to understand these contrasting results, the effects of other miRNAs thought to target at those sites must be considered (Figure 3.5).

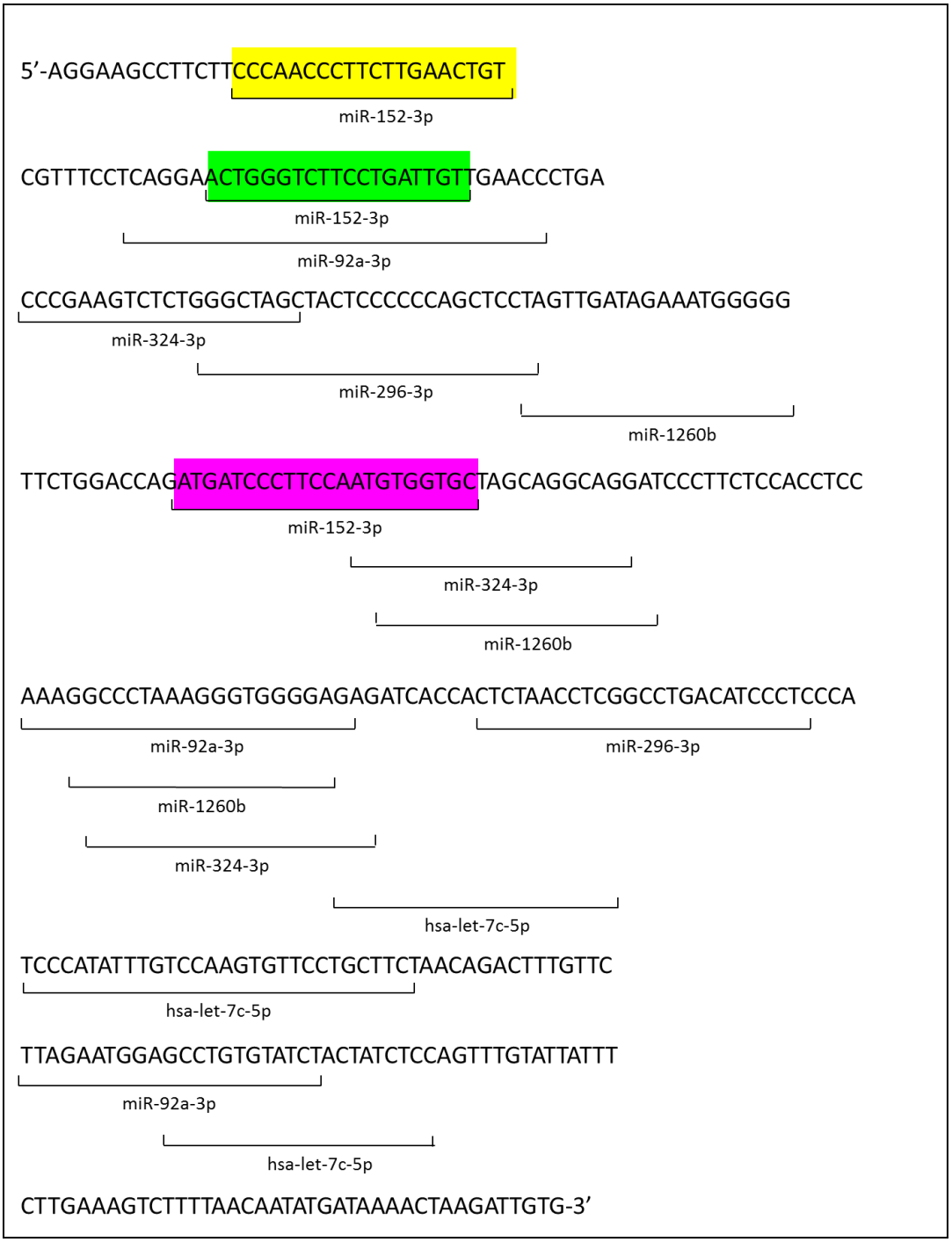


Figure 3.5 MiRNA binding sites in the SETDB1 3’UTR. All miRNAs listed in this figure are determined to target SETDB1 by Helwak, et. al. using CLASH[2]; however, the annotated sequences shown are predicted by miRanda. TRS 1 is

highlighted in yellow, TRS 2 is highlighted in green, and TRS 3 is highlighted in purple. TRS 1, 2, and 3 were deleted using site-directed mutagenesis to form 3 mutant vectors. Deletions of these sites were performed to help determined the necessity of those sequences for SETDB1 regulation.

At TRS 1, there are presently no other miRNAs besides miR-152-3p predicted to target at this site. Since the luciferase/*Renilla* value decreased with TRS 1 deletion, we believe miR-152-3p is positively regulates SETDB1 protein. If the luciferase/*Renilla* value had remained unchanged, we would conclude miR-152-3p does not target at TRS 1. If the luciferase/*Renilla* value had significantly increased, we would conclude that miR-152-3p is repressive of SETDB1.

In contrast to there being no other miRNA binding site overlaps at TRS 1, there are two other miRNAs predicted by miRanda to bind at TRS 3. If we assume that these other miRNAs, miR-1260b and miR-324-3p, are SETDB1 repressive and that at least one of them represents a biologically meaningful interaction with SETDB1, deletion of the site would significantly increase luciferase/*Renilla*. This is precisely what was observed. Interestingly, the luciferase/*Renilla* level doesn't identically match the empty vector value. Possible interpretations are that this may be due to experimental variation, as the difference is not significant, or that it may represent the counter-effects of miR-152-3p on SETDB1. Under the pretense that miR-1260b and/or miR-324-3p are SETDB1 repressive, whether miR-152-3p binds at TRS 3 or not, we might observe the same significant increase in luciferase/*Renilla* that was obtained. For this to be the case, equal activity of all miRs on TRS 3 would have to be assumed. Thus, we can

eliminate the possibility that miR-152-3p activity is greater than that of miR-1260b or miR-324-3p at TRS 3.

Though it was previously found that miR-324-3p targets SETDB1[2], our reporter assay findings in combination with anti-miR-324-3p transfection and subsequent SETDB1 protein analysis do not entirely support a meaningful biological interaction between the two. Since the TRS 3 deletion results would be the same even without miR-324-3p binding, and inhibition of miR-324-3p had no noticeable effect on SETDB1 protein, it may be that the interaction between miR-324-3p and SETDB1 is transient with no downstream effects.

Based on the results obtained upon deletion of TRS 2, we can infer one of two things: either miR-92a and miR-152-3p bind and have opposite regulatory functions, such that miR-92-a is a negative SETDB1 regulator and miR-152-3p is a positive SETDB1 regulator, or that neither miRNA binds to TRS 2. Binding of miRNAs at TRS 2 must be "all or none" simply because these are the only two scenarios, based on the available information, which would result in insignificant findings. Since the results from TRS 1 deletion indicate that miR-152-3p is a positive regulator, a successful interaction at TRS 2 would also show a significant decrease in luciferase/*Renilla*. Conversely, if miR-92a alone binds and does so in a repressive manner, there would be a significant increase in luciferase/*Renilla*. If both miRNAs bind with the described functions and at equal efficiencies, their activities would net no impact on the luciferase/*Renilla* luminescence. Thus, we can infer that either both miRNAs bind at TRS 2 with opposing functions, or neither miRNA binds at TRS 2. In any case, neither of the miRNAs bind alone to TRS 2. These considerations hold true in both cell lines since deletion of TRS 2 was not significant in IMR-91 and HepG2 below $p < 0.01$.

One experiment that we've omitted in this study is to co-transfect cells with miR-152-3p mimic or anti-miR-152-3p and the reporter vector. Doing so may have allowed us to more directly study the effects of miR-152-3p on SETDB1. However, we predicted that if miR-152-3p was inhibited, the effects of other miRNAs that bind to the SETDB1 3'UTR would have dominated. If cells were transfected with miR-152-3p mimic, it may have been possible to get more miR-152-3p binding to the 3'UTR. On the other hand, it is also possible that what we observed was already the maximal binding capacity of miR-152-3p to SETDB1, and adding mimic may not make any difference. Moreover, treatment with anti-miR and mimic needed to be done on the same day as cell seeding, according to the HiPerFect transfection reagent protocol. Then, it took 12 hours post-transfection to see an effect with either anti-miR or mimic, as determined by our own experiments. This timeline was incompatible with the vector transfection and assay timeline, which required cells to be adherent upon transfection, and luminescence was measured on day 2. Thus, given the likelihood that no new knowledge would be obtain through co-transfection and the challenge of weaving together different timelines, this experiment was not pursued.

Another strategy for demonstrating miR-152-3p binding to the SETDB1 3'UTR would have been to do a miRNA binding experiment, such as CLASH. These types of experiments rely on pull downs of miRNA with targets. However, because they are accomplished on fixed cells, the biological relevance of results comes into question. This is particularly true because CLASH does not provide a way to distinguish transient interactions from stable interactions. Therefore, it is essential to employ multiple approaches when confirming a potential miRNA interaction. Furthermore, Helwak, et. al. and others concede that that non-canonical binding sites may not

even be functional[2, 57]. While the technique of CLASH yields only 2% non-specific binding[2], the biological relevance of targets obtained is unclear.

One of the biggest issues with using the SETDB1 3'UTR to study miR-152-3p binding is that there are five other miRNAs which are experimentally confirmed to bind to SETDB1 and predicted to do so at the 3'UTR, and numerous predicted miRNAs, that bind in this 367 base pair region. Several of these binding sites overlap with the miR-152-3p TRSs, confounding the luciferase reporter assay results (Figure 3.5). Since it is not practical to inhibit every miRNA that may bind to the 3'UTR, additional mutation studies could be done. For example, truncating the 3'UTR after the last miR-152-3p target recognition sequence would eliminate some of the other miRNA binding sites. Another option would be to clone in only the target recognition sequences, one per luciferase reporter vector to see the specific effect of miR-152-3p activity on those sites. Alternatively, each binding site could sustain point mutations at the seed sequence to see if there is any impact on luciferase activity. Perhaps through doing multiple mutation studies, additional evidence supporting miR-152-3p binding and negative regulation of SETDB1 translation could be found. However, we primarily relied on miR-152-3p mimic and anti-miR-152-3p treatment followed by western blot analysis of SETDB1 protein for assessing directionality of the miRNA's function, and used the reporter vector results to confirm our western blot results.

CHAPTER 4

EFFECTS OF miR-152-3p REGULATION OF SETDB1 ON HISTONE METHYLATION

Introduction

It is well-established that SETDB1 methylates H3K9me2, thereby producing H3K9me3. Since inhibition of miR-152-3p caused SETDB1 protein levels to decrease compared to water-treated cells, we sought to determine if H3K9me3 levels would also be reduced with anti-miR-152-3p treatment relative to water treated cells. Conversely, when cells were treated with miR-152-3p mimic, we saw an increase in SETDB1 protein compared to water-treated cells. Thus, we predicted that H3K9me3 levels would increase in the presence of miR-152-3p mimic relative to water treated cells. Since H3K9me2 serves as the substrate for SETDB1, levels for H3K9me2 and SETDB1 should be inversely linked under miR-152-3p perturbation. Specifically, we predicted miR-152-3p mimic treatment would decrease H3K9me2 levels compared to water-treated cells and anti-miR-152-3p treatment would cause H3K9me2 levels to increase compared to water treated cells. To study the levels of H3K9me2/3, we performed anti-miR and mimic treatments followed by a histone ELISA for H3K9me2/3. Changes in H3K9me2/3 levels as a result of miR-152-3p inhibitor or mimic would indicate that miR-152-3p has an indirect role in epigenetic gene regulation.

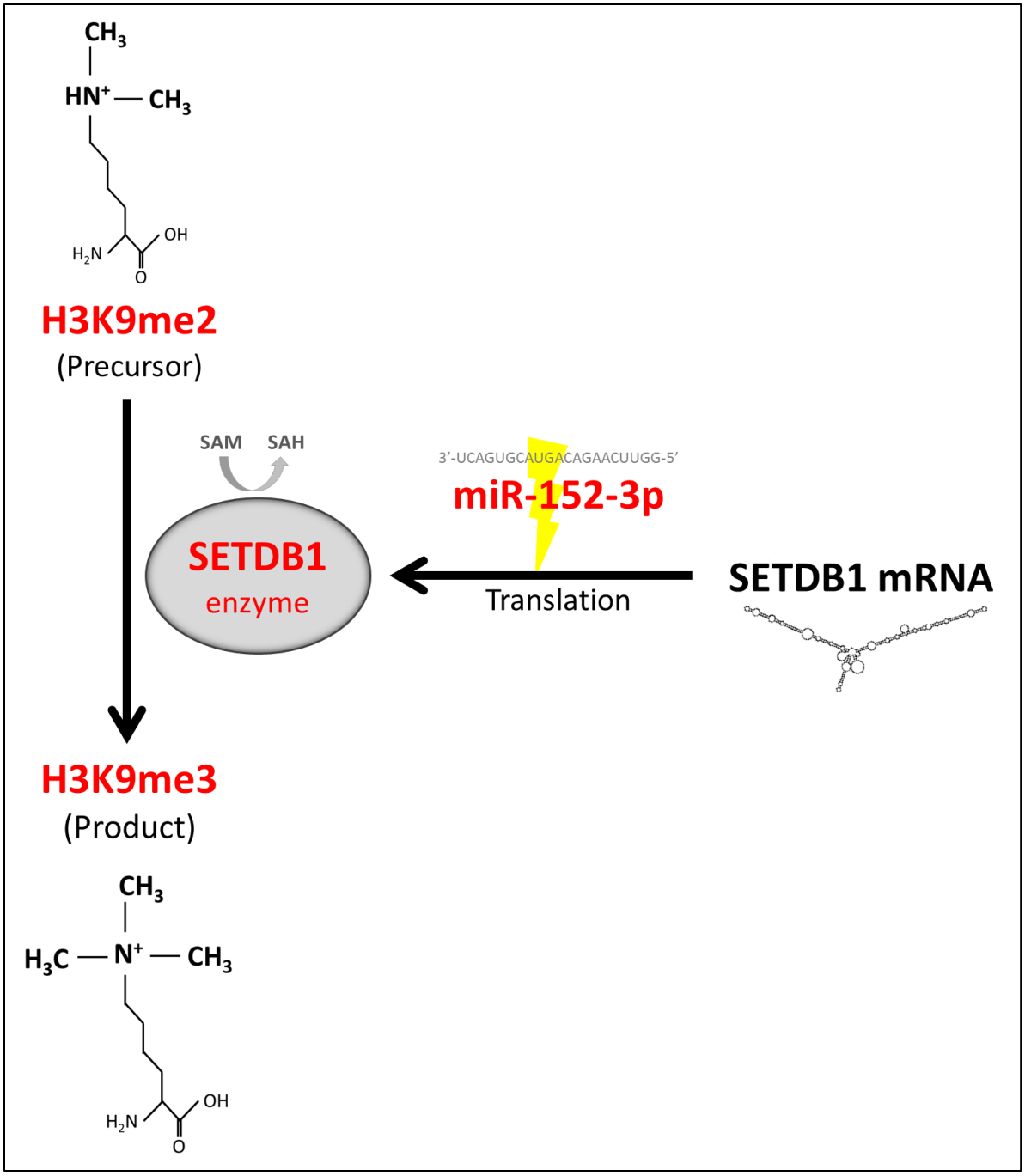


Figure 4.1 Relationship between miR-152-3p, SETDB1, and H3K9me2/3. MiR-152-3p targets SETDB1 mRNA. MiR-152-3p affects the translation of SETDB1, an enzyme which converts H3K9me2 to H3K9me3.

Methods

Preparation of Histones

10^6 cells were transfected according to standard procedures with 50 nM anti-miR-152-3p, 5 nM miR-152-3p mimic, or water. 12 hours after transfection, cells were trypsinized, washed in PBS, and then resuspended in 0.4 M HCl. Cells were incubated on ice for 30 minutes with occasional mixing. Lysate was centrifuged at 11,000 x *g* for 10 minutes at 4°C. The supernatant was collected and 2/5 the total volume of neutralization buffer (1M $NaPO_4$ dibasic, pH to 12.5 using 5 M NaOH, 2.5 mM DTT, and 10 mM PMSF, sometimes substituted with Halt Protease Inhibitor, Thermo Fisher Scientific, #78430) was immediately added. Histones were then aliquoted into single-use portions and stored at -80°C. A Lowry assay was used to quantify histones. Typical histone yields ranged from 0.70 µg/µL to 1.5 µg/µL.

H3K9me2 and H3K9me3 Enzyme-Linked Immunosorbant Assay (ELISA)

Pre-coated H3 capture ELISA plates were obtained from Active Motif for H3K9me2 and H3K9me3 analysis. Crude histone samples were plated and assayed according to kit instructions. In addition to measuring H3K9me2 and H3K9me3 in samples, we also purchased unmethylated H3 isolate from Active Motif to serve as a negative control for the ELISA. Samples were run in triplicate with a range of concentrations. The standard curve was performed in duplicate on each plate.

H3 from crude histone samples was bound to plates with incubation for 1 hour at room temperature. The plate was washed 3 times with wash buffer, then incubated in primary antibody for 1 hour at room temperature with agitation. Afterwards, the plate was washed 3 times with wash buffer. Next, secondary antibody was bound for 1 hour at room temperature without agitation. The plate was washed 3 times with wash buffer prior to beginning the colorimetric reaction.

Developing solution was added to each well and samples were incubated for approximately 30 seconds to 5 minutes under low light until standard curve wells appeared medium to dark blue. At this point, stop solution was added to wells and the absorbance was read at 450 nm (assay wavelength) and 655 nm (reference wavelength).

Data Analysis

Data was processed by first subtracting the reference wavelength absorbance value from the assay wavelength absorbance value. Then, the absorbance value for the blank was subtracted from each value. The final absorbance values for the standard curve wells were plotted against the known concentrations. Linear regression created a line of best fit and an equation which was then used to determine the concentrations of H3K9me2/3 in the samples. T-tests were used to determine statistical significance.

Results

ELISA Controls

No non-specific antibody binding to H3 was observed with H3K9me2 (Figure 4.2) or H3K9me3 (Figure 4.3). We also found significantly higher ($p < 0.0001$) H3K9me2 and H3K9me3 levels in HepG2 cells than in IMR-91 cells (Figure 4.4).

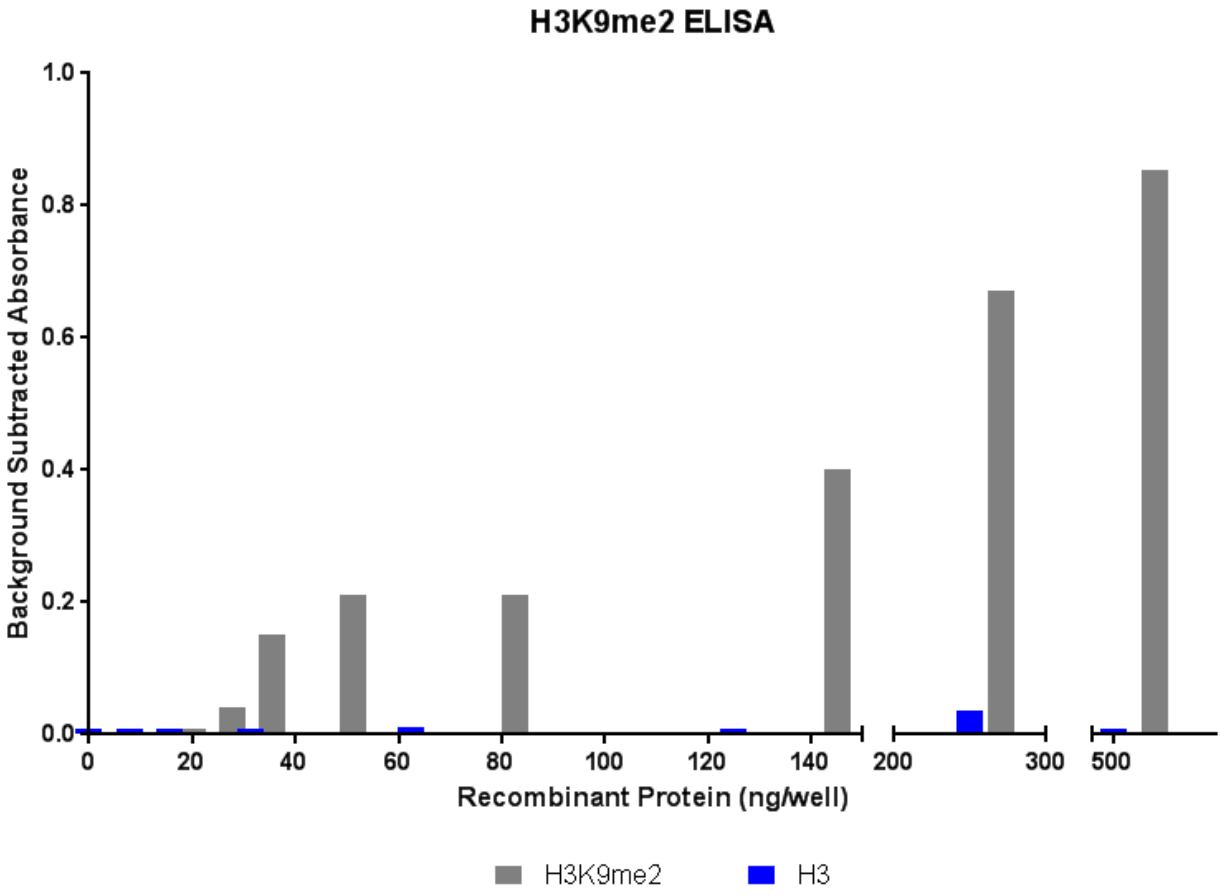


Figure 4.2 Histone H3 and H3K9me2 ELISA results. Unmethylated H3 and H3K9me2 were plated in duplicate at 0 ng/well and 2-fold concentrations ranging from 7.8 ng/well to 500 ng/well for an ELISA measuring H3K9me2. There is

high specificity for H3K9me2, as shown through the background (Opti-MEM only wells) subtracted absorbance values. By contrast, very little to no signal is seen for H3.

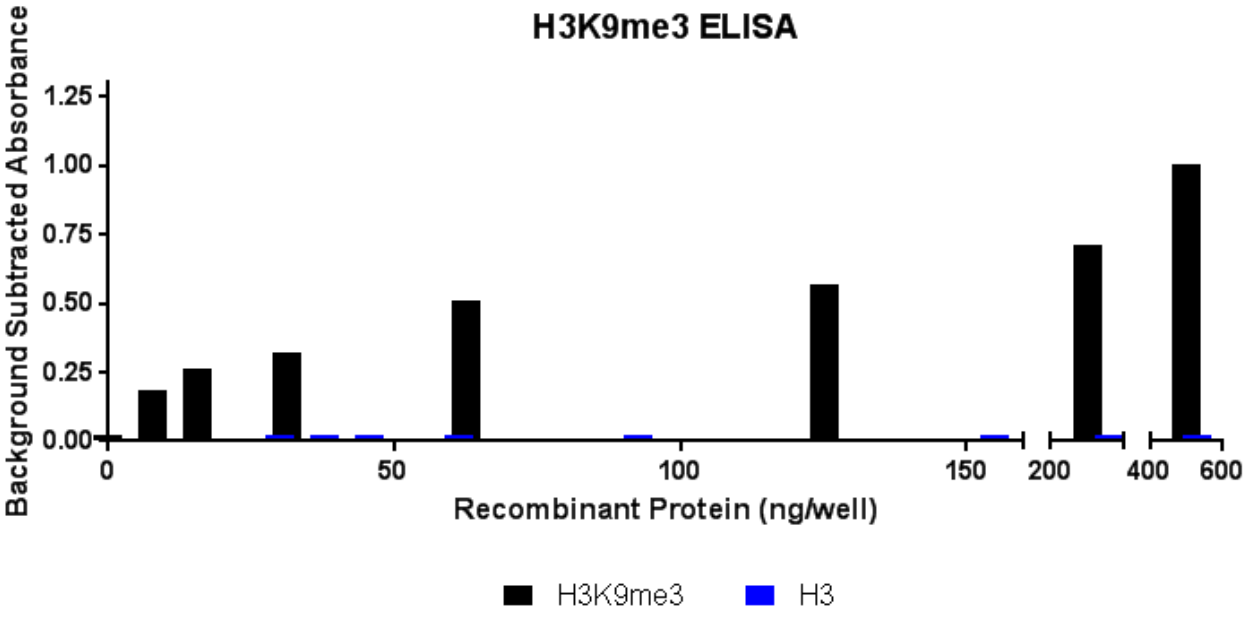


Figure 4.3 Histone H3 and H3K9me3 ELISA results. Unmethylated H3 and H3K9me3 were plated in duplicate at 0 ng/well and 2-fold concentrations ranging from 7.8 ng/well to 500 ng/well for an ELISA measuring H3K9me3. There is high specificity for H3K9me3, as shown through the background (Opti-MEM only wells) subtracted absorbance values. By contrast, very little to no signal is seen for H3.

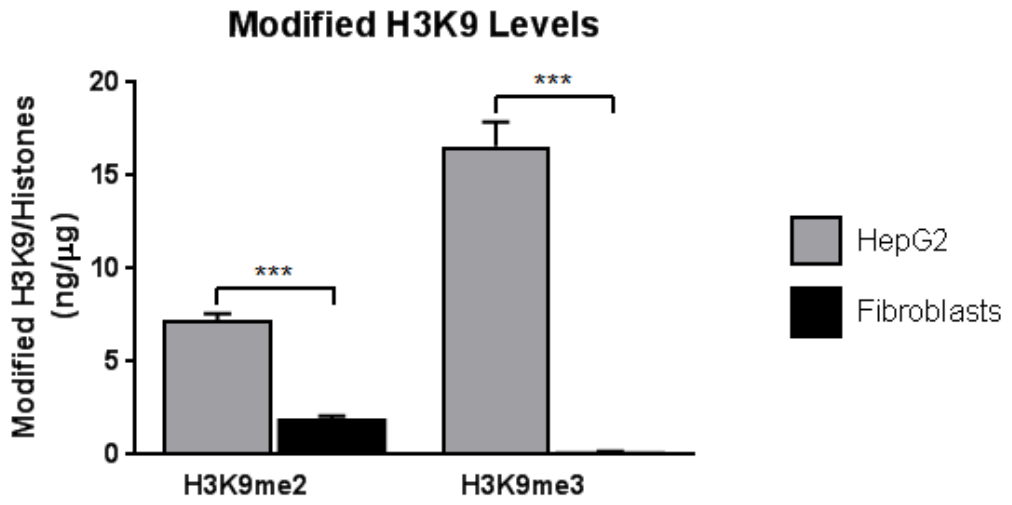


Figure 4.4 Steady state levels of H3K9me2/3 in cell lines. HepG2 cells have significantly more ($p < 0.0001$) H3K9me2 and H3K9me3 than IMR-91 cells.

Effects of miR-152-3p Regulation of SETDB1 on H3K9me2/3

Anti-miR-152 treatment of HepG2 cells caused a significant increase ($p < 0.0001$) in H3K9me2 levels compared to water-treated cells, and no change in H3K9me3 levels compared to water-treated cells (Figure 4.5). Likewise, in IMR-91 cells, there was no significant change in H3K9me3 levels with anti-miR-152-3p treatment compared to water treated cells (Figure 4.6).

MiR-152-3p mimic treatment of HepG2 cells caused a significantly decreased ($p < 0.05$) in H3K9me2 levels compared to water-treated cells, and a significant increase ($p < 0.05$) in H3K9me3 compared to water-treated cells (Figure 4.5). In miR-152-3p mimic treated IMR-91 cells, H3K9me2 levels significantly increased ($p < 0.05$) and H3K9me3 levels significantly decreased ($p < 0.05$) (Figure 4.6).

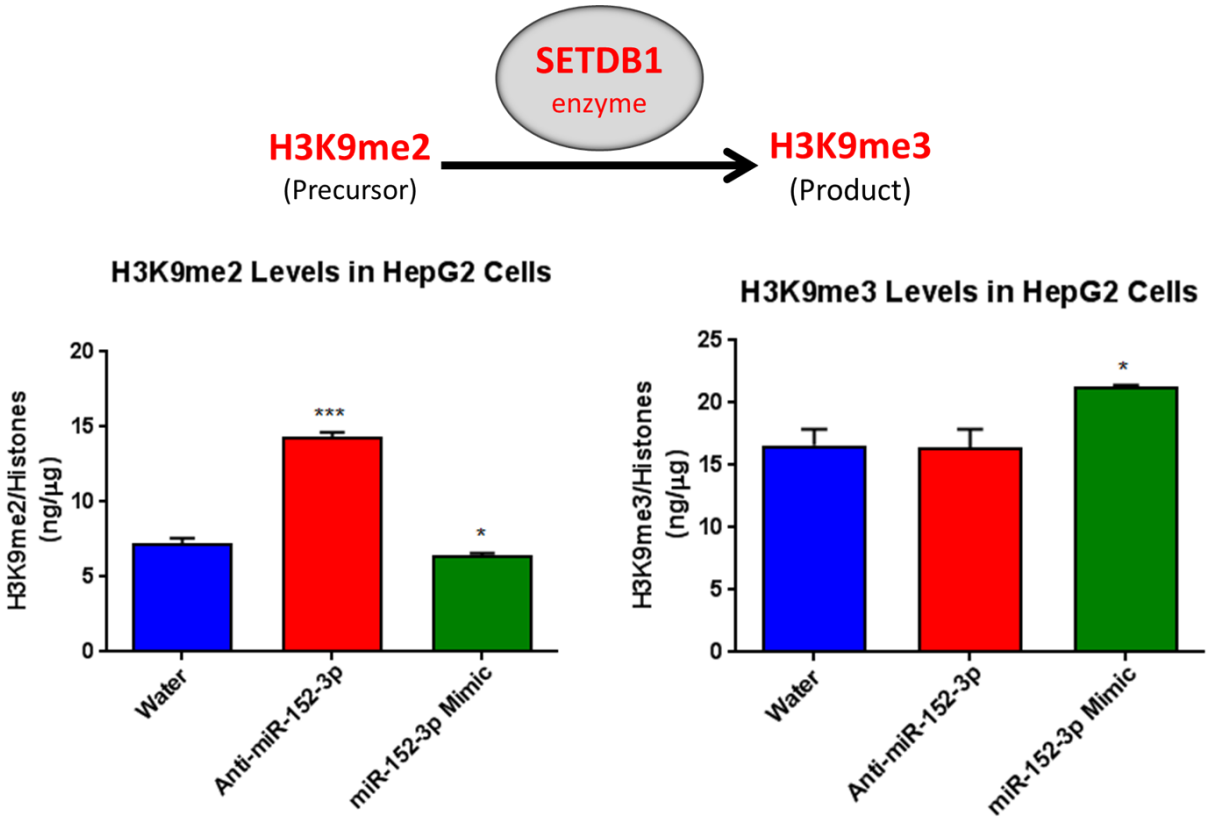


Figure 4.5 ELISA results for H3K9me2/3 levels in HepG2 cells. H3K9me2 levels significantly increase ($p < 0.0001$) with anti-miR-152 treatment and significantly decrease ($p < 0.05$) with miR-152-3p mimic treatment in HepG2 cells as compared to water treated HepG2 cells. H3K9me3 levels significantly increase ($p < 0.05$) in miR-152 mimic treated HepG2 cells relative to water treated HepG2 cells.

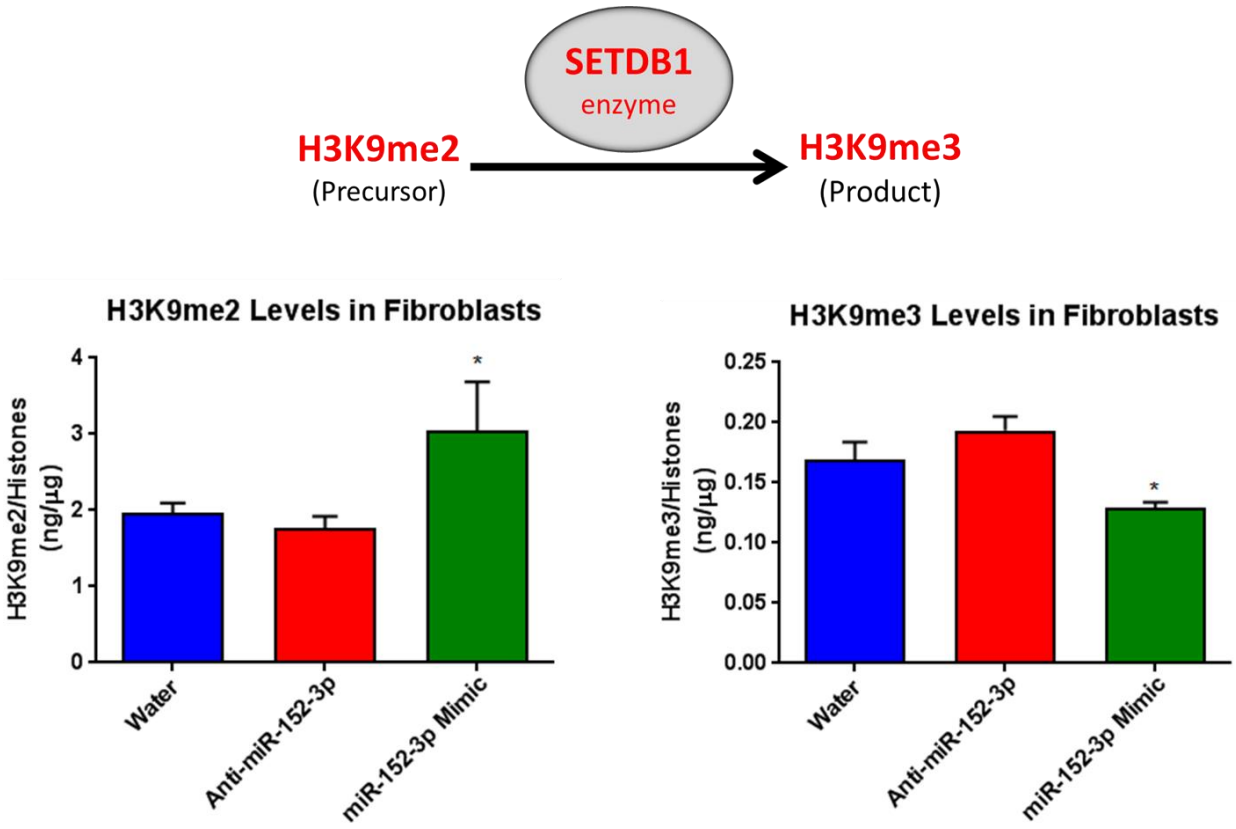


Figure 4.6 ELISA results for H3K9me2 and H3K9me3 levels in fibroblasts. H3K9me2 levels significantly increase ($p < 0.05$) in miR-152 mimic treated IMR-91 cells, but H3K9me3 levels significantly decreased ($p < 0.05$) with miR-152-3p mimic treatment relative to control (water treated) IMR-91 cells. The effects of anti-miR-152-3p treatment on H3K9me2 and H3K9me3 were not significant when compared to control cells in fibroblasts.

Conclusions and Discussion

As predicted, H3K9me2 significantly increased ($p < 0.0001$) with anti-miR-152-3p treatment and H3K9me3 levels significantly increased ($p < 0.05$) with miR-152-3p mimic treatment in HepG2. Under anti-miR-152-3p treatment, miR-152-3p was inhibited and could no longer bind to or positively regulate SETDB1. With less SETDB1 present, there may have been a reduced turnover from H3K9me2 to H3K9me3. This may be why we saw more H3K9me2 with anti-miR-152-3p treatment and no change in H3K9me3, of which there might have been a pre-existing pool. The opposite held true with miR-152-3p mimic treatment; there may have been more miR-152-3p present to upregulate SETDB1, which then may have increased turnover from H3K9me2 to H3K9me3. Therefore, there was a significant decrease in H3K9me2, and a significant increase in H3K9me3 with miR-152-3p mimic treatment.

Ultimately, we found that miR-152-3p regulation of SETDB1 proportionately impacts histone methylation. Decreases in SETDB1 caused by anti-miR-152-3p treatment resulted in increases in H3K9me2. Conversely, increases in SETDB1 caused by miR-152-3p mimic treatment corresponded to increases in H3K9me3. Therefore, miR-152-3p regulation of SETDB1 affects levels of the modified histone proteins involved in epigenetic gene silencing in HepG2 cells. Furthermore, the regulation of SETDB1 by miR-152-3p is biologically relevant in HepG2 cells due to these effects on H3K9me2/3.

Though we predicted anti-miR-152-3p treatment would elicit a decrease in H3K9me3, it did not do so. Instead of undermining our hypothesis, we believe this finding exposes a limitation in the experimental design. Because cells were grown, treated, then assayed for modified histone, there was no way to distinguish H3K9me2 or H3K9me3 resulting directly

from treatment from existing H3K9me2 or H3K9me3. We only knew there was an effect based on statistically significant differences from the control. With this understanding, we speculate that anti-miR-152-3p did not impact H3K9me3 because H3K9me3 was already present. Therefore, the H3K9me3 values for anti-miR-152-3p treated HepG2 cells represented a pre-existing pool of the modified histone.

The best way to address the pre-existing pool of modified histones would be to perform radio-labeling experiments. This way, the newly synthesized H3K9me2 or H3K9me3 would possess an isotope label to allow us to distinguish it from pre-existing H3K9me2 or H3K9me3. In addition to radio-labeling experiments, there are several alternative approaches for the study of histone modifications instead of the histone ELISA we utilized. However, all of these techniques suffer from an inability to distinguish newly modified histones from previously existing histone modifications.

Western blot would have been the least expensive method for observing overall effects of treatment on cells if already optimized. However, it is not directly quantitative in nature and would not provide a statistical analysis of differences. Flow cytometry can also be done to look at histone modifications[59]. With optimization, this would have been a valid method as it is both quantitative and allows for statistical analysis of differences. One final technique that could have been used for studying histone modifications is quantitative mass spectrometry. This method was not chosen as it requires specialized knowledge of mass spectrometry. ELISA was used to study histone modifications as it is quantitative, rapid (3.5 hour procedure), and more cost effective than western blotting since the kit and reagents are already optimized.

Interestingly, the same changes seen in H3K9me2 and H3K9me3 levels in HepG2 cells were not seen in IMR-91 cells. In fact, anti-miR-152-3p treatment had no significant effect on

H3K9me2 or H3K9me3. Yet, miR-152-3p mimic treatment significantly affected both H3K9me2 and H3K9me3 in fibroblasts, but in opposite directions. Notably, HepG2 cells are derived from a 15 year old human with hepatocellular carcinoma, while IMR-91 comes from healthy human fetal skin fibroblasts. The different results from each cell line may indicate the healthy cells respond to increases in SETDB1 levels differently than the cancer-derived HepG2 cells. Since SETDB1 has also been found to have roles in mono- and dimethylation of H3K9[36, 37], it may be that these actions are at play rather than the trimethylating activity of SETDB1. Another strong possibility is that what we observe in IMR-91 cells is akin to background H3K9me2/3, since neither is present at very large quantities in these fibroblasts (Figure 4.4).

CHAPTER 5

EFFECTS OF miR-152-3p POSITIVE REGULATION OF SETDB1 PROTEIN ON HP1α

Introduction

Having shown that miR-152-3p targeted the 3'UTR of SETDB1, upregulated SETDB1 protein, decreased H3K9me2, and increased H3K9me3, we were interested in determining if there was any effect on heterochromatin protein 1 (HP1). HP1 is a structural protein recruited by H3K9me3 and involved in the formation of heterochromatin (Figure 5.1). Of the three isoforms, HP1α was chosen for study due to its localization to constitutive heterochromatin[8].

We predicted that because H3K9me3 levels increased with miR-152-3p mimic treatment, HP1α levels in the nucleus would also increase with miR-152-3p mimic treatment. Likewise, we predicted less HP1α to be recruited to histones with anti-miR-152-3p treatment. In this case, HP1α would appear decentralized (not attached to histones or DNA) under immunofluorescence. The reason we predicted HP1α would be decentralized with anti-miR-152-3p treatment is because there was less H3K9me3 than with miR-152-3p mimic treatment. However, as we saw in our ELISA results, there may be a pre-existing pool of HP1α, thereby confounding our results.

We visualized HP1α under inverted microscopy and replicated the imaging in HepG2 cells using confocal microscopy and fluorescent intensity analysis. Since localization of HP1α is cell cycle dependent[7, 53], cell cycle analysis with flow cytometry was done. This allowed us to distinguish fluctuations in HP1α due to cell cycle progression from fluctuations resulting from anti-miR-152-3p or miR-152-3p mimic transfection.

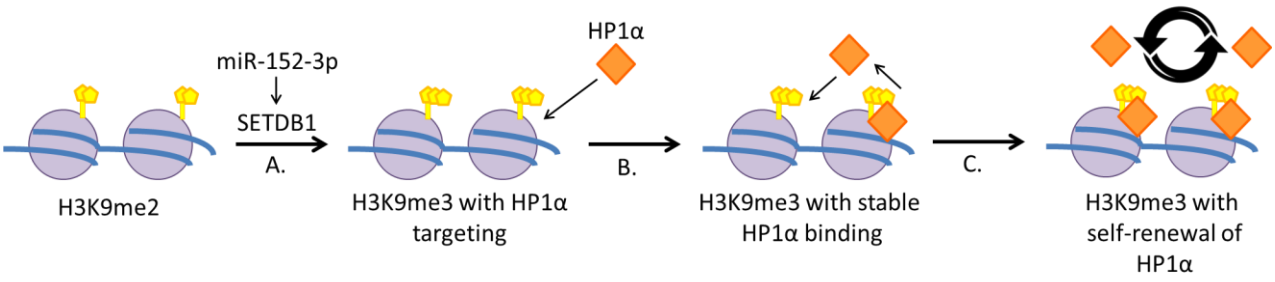


Figure 5.1 Recruitment of HP1α. (**A**) miR-152-3p upregulates SETDB1, which trimethylates H3K9me2 and forms H3K9me3. The N-terminus chromodomain of HP1α binds to H3K9me3, while the hinge domain of HP1α binds DNA and linker histone. (**B**) H3K9me3 and HP1α binding at H3K9me3 stimulate subsequent HP1α recruitment. (**C**) As HP1α binds, heterochromatin is formed. Heterochromatin formation recruits additional HP1α.

Methods

HP1α Immunofluorescence

For inverted microscopy, cells were grown in a 12-well cell culture treated plate for 12 hours after transfection with anti-miR-152-3p, water, or miR-152-3p mimic. Cells were washed in PBS, fixed in 4% PFA/PBS for 15 minutes, then washed 3 times with 1% BSA/TBS.

Permeabilization was done with 0.25% Triton X-100/TBS for 10 minutes. Cells were then blocked in 5% goat serum/TBS for 1-3 hours at room temperature. They were incubated in 1:400 anti-HP1α antibody (ab9057) for 16 hours at 4°C. Next, cells were washed 3 times in 1% BSA/TBS and incubated in Alexa 488 secondary at 1:800 for 1 hour at room temperature. Cells were then washed, DAPI stained, and washed again prior to visualization. For confocal microscopy, the protocol was identical with the exception that cells were grown on cell culture treated chamber slides (Ibidi).

For all immunofluorescence, cells were plated in triplicate. Images were taken from across each well beginning with the DAPI setting, so as not to allow the HP1α signal to influence image selection. For fluorescence intensity analysis using confocal microscopy, each field contained a minimum of 1 dozen cells from which data could be collected. Cells where discrete nuclei could not be distinguished were omitted from statistical analysis. Region of interest (ROI) boxes of uniform area were drawn at the center of nuclei under FITC and DAPI. Fluorescent values were then recorded. This was done for over 78 water treated cells, 75 anti-miR-152-3p treated cells, and 94 miR-152 mimic treated cells. One-way ANOVA was done to determine mean differences between groups. Cells stained with the primary antibody omitted were also visualized to provide a background control.

HP1α Flow Cytometry and Cell Cycle Analysis

10^6 cells per condition were grown in 6-well cell culture treated plates. Cells were trypsinized and washed then fixed with 70% ethanol at -20°C. Cells were permeabilized with 0.1% Tween-20 in PBS for 20 minutes at room temperature. They were then incubated in

primary antibody (anti-HP1α, ab9057) for 20 minutes at a dilution of 1:50. Secondary staining with Alexa 647 occurred for 20 minutes at a dilution of 1:400. Cells were RNase treated prior to propidium iodide (PI) staining and analyzed using a BD LSRFortessa X-20.

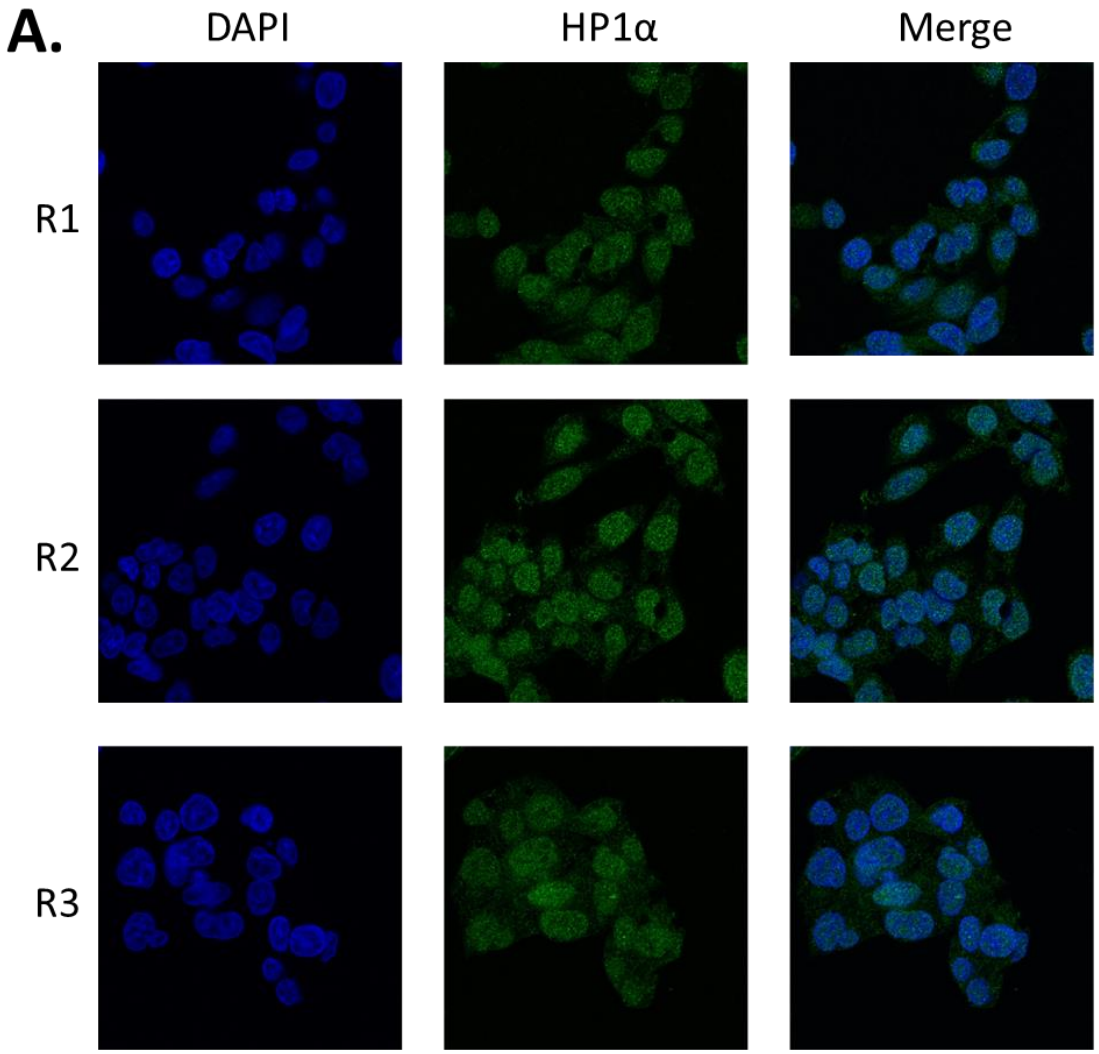
Unstained water treated HepG2 cells, antibody (HP1α + Alexa 647) stained water treated HepG2 cells, and propidium iodide (PI) stained water treated HepG2 cells were chosen as compensation controls for flow cytometry. Cells treated with anti-miR-152-3p and miR-152-3p mimic were plated in duplicate to allow one set of cells for each condition to be stained with secondary antibody (Alexa 647) plus PI and the other set to be stained with HP1α, Alexa 647, and PI. These two staining conditions were also done for water treated cells. Using this combination of parameters allows us to look at the expression of HP1α at G0-G1, S, and G2-M stages of the cell cycle. PI stains all nucleic acids, thus, cells were RNase treated prior to PI staining to ensure DNA-specific staining.

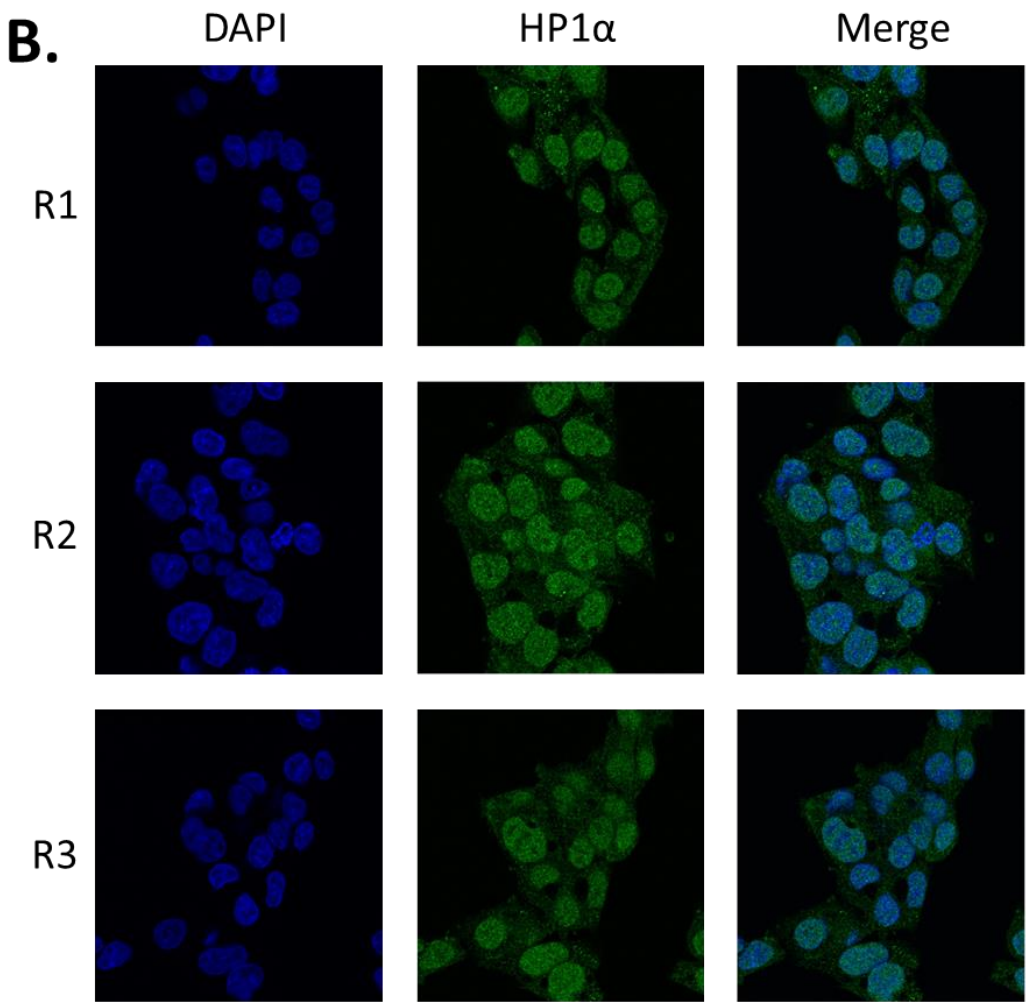
Results

Effects of miR-152-3p Regulation of SETDB1 on HP1α

HP1α images from inverted microscopy can be found in Appendix II: Supplementary Data. These images were selected and organized by Madeline Dorso (2016). Images from confocal microscopy are presented in Figure 5.2. A single representative image is shown per

identically-treated replicate well (R1-R4). When quantified, anti-miR-152-3p treated HepG2 cells had a significantly higher ($p < 0.0001$) HP1α intensity/area than water treated HepG2 cells, as shown by FITC staining (Figure 5.3). Raw ROI values were obtained by Madeline Dorso (2016).




B.
DAPI
HP1α
Merge
R1
R2
R3

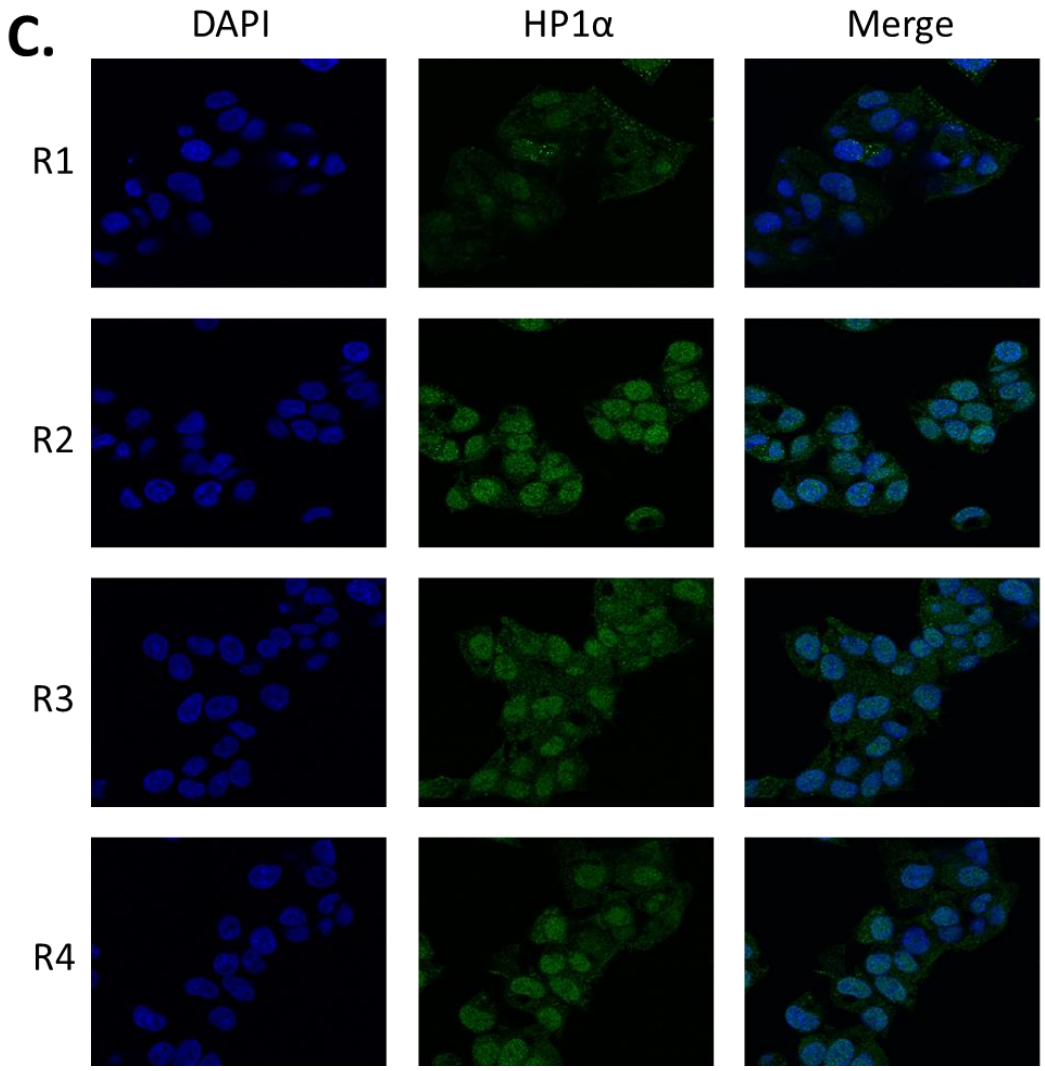


Figure 5.2 HP1α localization in HepG2 cells fixed 12 hours post-transfection. All wells were visualized under a Nikon A1R Spectral Confocal Microscope. Representative cells from replicate wells are shown in horizontal panels as R1, R2, R3, and R4. HP1α was visualized under the FITC channel. (**A**) HepG2 cells mock transfected with water. (**B**) HepG2 cells transfected with anti-miR-152-3p. (**C**) HepG2 cells transfected with miR-152-3p mimic.

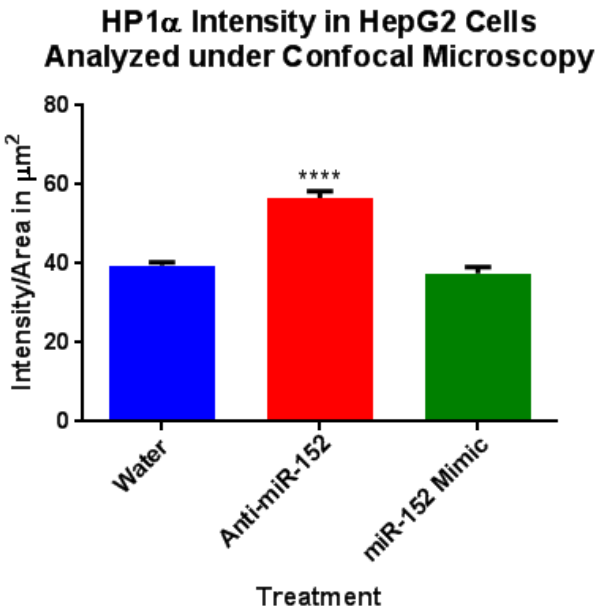


Figure 5.3. Nuclear HP1α staining intensity in HepG2 cells. HP1α was quantified by fluorescent intensity in the FITC channel using NIS Elements. Fluorescence from the FITC channel was calculated by first subtracting the background fluorescence from the raw FITC values and then by dividing the background subtracted value by the area of the ROI. Statistical analysis by one-way ANOVA found a significance of $p < 0.0001$ for anti-miR-152-3p treated cells compared to water treated cells.

Flow cytometry data showed that amounts of HP1α increased throughout the cell cycle stages G0-G1, S, and G2-M in control, water-treated, cells. We also found that HP1α levels increased across both experimental treatment conditions, anti-miR-152-3p and miR-152-3p mimic, for each of the cell cycle stages. Because there was only one sample per condition, statistical significance could not be determined (Figure 4.9).

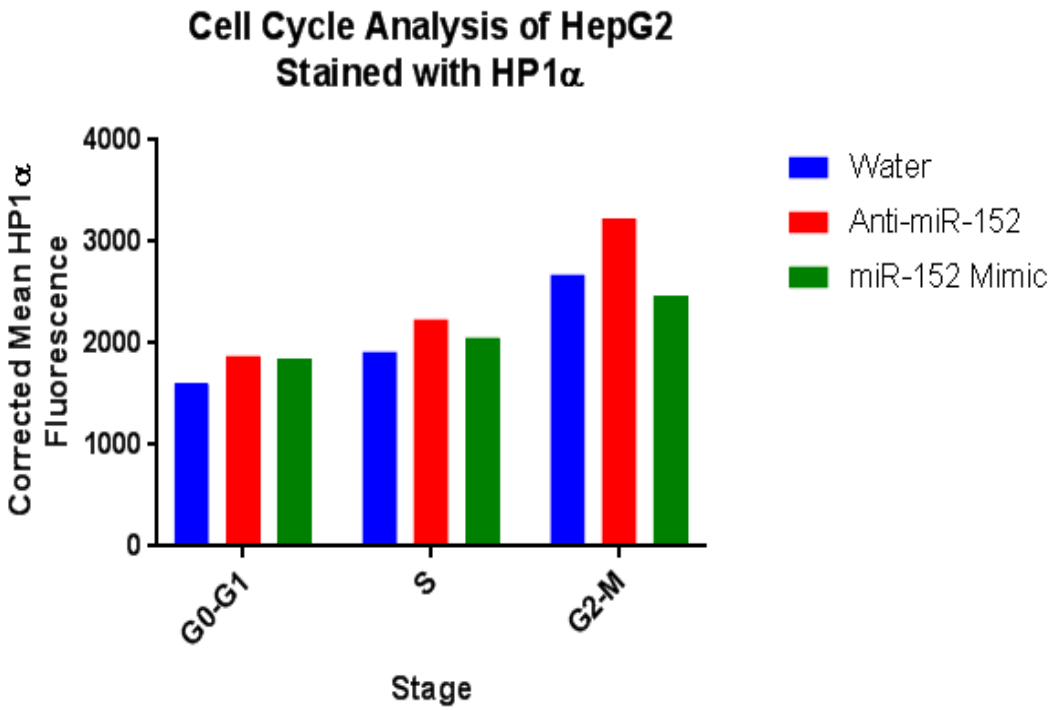


Figure 5.4. HP1α levels in treated HepG2 cells across stages of the cell cycle. HepG2 cells were treated with either water, anti-miR-152-3p, or miR-152-3p mimic for 12 hours before being fixed for flow cytometry with a BD LSRFortessa X-20 instrument. One sample was collected for each treatment condition. The corrected mean HP1α fluorescence was calculated by subtracting the APC-A mean for samples stained with anti-HP1α primary antibody, Alexa 647 secondary antibody, and propidium iodide from the APC-A mean for samples stained only with Alexa 647 secondary antibody and propidium iodide.

Conclusions and Discussion

Overall, there was no discernable difference in HP1α localization among treatment conditions as visualized under confocal microscopy. With confocal microscopy, we concluded HepG2 cells mock transfected with water showed an even dispersion of HP1α in the nucleus and cytoplasm (Figure 5.2A). We also concluded that HepG2 cells transfected with anti-miR-152-3p appeared to have an increased amount of HP1α in both the nucleus and cytoplasm compared to water treated cells, with even staining in both regions, rather than changes in localization (Figure 5.2B). Finally, we concluded that HepG2 cells transfected with miR-152-3p mimic did not seem to have a clear pattern of HP1α location or abundance relative to water or anti-miR-152-3p treated cells (Figure 5.2C).

We noticed a difference in staining intensity of HP1α across samples such that cells treated with anti-miR-152-3p had significantly higher FITC staining ($p < 0.0001$) under immunofluorescent microscopy. We hypothesized that this may have been due to an increase in the quantity of HP1α, better detection of HP1α in those wells each time, or that miR-152-3p affected cell cycle progression. Since there is an established paradigm between HP1α levels and stage in the cell cycle, we looked at the effects of miR-152-3p on HP1α with flow cytometry. We found that HP1α levels increased across all treatments. As neither our previous work nor the literature can provide a suitable explanation for the observations that both anti-miR-152-3p and miR-152-3p mimic treatments increase HP1α throughout the cell cycle, we concluded that miR-152-3p may not actually bear an effect on HP1α. Instead, changes in HP1α localization or quantity observed across treatments could be due to the inherent fluctuations in HP1α described by Sugimoto, et al.[53]. It is worth noting that there is bleed over from each phase when doing cell

cycle analysis. Thus, if a true S-phase is to be seen, an additional stain such as PCNA should be used. For our purposes, a clear S-phase was not necessary.

While the effects on H3K9me2 and H3K9me3 did not appear to translate into profound changes in HP1α, there still exists the possibility that the other two isoforms of HP1 may be affected. Heterochromatin protein 1 beta (HP1β) would be an interesting protein to study as part of a future project. HP1β is also localized to constitutive heterochromatin[8] and thus, may be affected by changes in H3K9me3. Although miR-152-3p regulates SETDB1 and affects histone methylation, there is currently no known effect on HP1α recruitment or heterochromatin formation.

CHAPTER 6

FUTURE DIRECTIONS AND PRELIMINARY STUDIES

There are two major areas of focus for which we have performed some preliminary studies. This section is divided into two parts, I and II, to reflect each of these areas of research. Part I unites existing research on miR-152-3p targeting of DNMT1 with the body of work presented thus far. In it, a potential feedback loop between regulation for DNMT1 and SETDB1 is described. In Part II, we propose a potential application for miR-152-3p targeting of SETDB1. We set forth a case for manipulation of miR-152-3p as a means of facilitating hepatocyte differentiation. Though we cannot presently draw any conclusions from the preliminary data presented, we anticipate pursuing both of these areas in the future.

I. The Potential Role of miR-152-3p as a Molecular Switch between Histone and DNA Methylation

Background

DNA Methylation and DNA Methyltransferase 1 (DNMT1)

DNA methylation is important for proper epigenetic gene regulation. Dysregulation of methylation has been linked to diseases such as cancer, Rett syndrome, ATR-X syndrome, and diabetes[60]. DNA methyltransferases are a category of enzymes which methylate cytosine nucleotides in DNA. Like SETDB1, they receive their methyl group from the methyl donor, SAM[61]. There are two types of DNA methyltransferases, *de novo* and maintenance. DNA methyltransferases 3A and 3B are both *de novo* methyltransferases, meaning they can add new methyl groups to cytosine. *De novo* methyltransferases are active largely during early development. DNA methyltransferase 1 (DNMT1) is a methylation maintenance enzyme and as such, functions to preserve methylation patterns after DNA replication.

MiR-152 has a well-established down-regulatory effect on DNMT1. According to Huang, et. al., transfection of HepG2 cells with miR-152 mimic reduces DNMT1 protein and anti-miR-152 treatment increases DNMT1 protein[56]. These results are supported by Ji, et. al. who also found that 16HBE (normal human bronchial epithelial) and HEK293T (human embryonic kidney) cells treated with miR-152 mimic reduce DNMT1 protein levels and that miR-152 targets the 3'UTR of DNMT1[62]. Additionally, Wang, et. al. found treatment of dairy cow

mammary epithelial cells (DCMEC) with anti-miR-152 results in increased DNMT1 protein relative to miR-152 treated DCMEC, although protein levels appear to be less than the control cells. This group concludes that DNMT1 is downregulated by miR-152 via 3'UTR targeting, which they studied using a luciferase reporter assay[63]. Thus, miR-152, the miRNA we show up-regulates a histone methyltransferase, also down-regulates a DNA methyltransferase.

The idea that DNA methylation and histone methylation are connected is not novel. Researchers have found that DIM-5, an H3K9me3 KMT, is required for DNA methylation in *N. crassa*, a type of filamentous fungi[64]. With the H3K9me3 mark in place, HP1 is recruited[50]. The CSD of HP1 interacts with DIM-2, a DNA cytosine methyltransferase[65]. Interestingly, DNA methylation is abolished when the gene encoding HP1, *hpo*, is knocked out[66]. This molecular pathway provides just one example of the related mechanisms between DNA methylation and histone methylation.

Introduction

The proper maintenance of histone methylation and DNA methylation is crucial for normal cell function. Dysregulation of methylation can affect development and cause disease. Cancer is one such example where the epigenetic status of cells is disrupted[60, 67, 68]. In fact, DNA promoter hypermethylation resulting in the silencing of miRNA genes is increasingly being

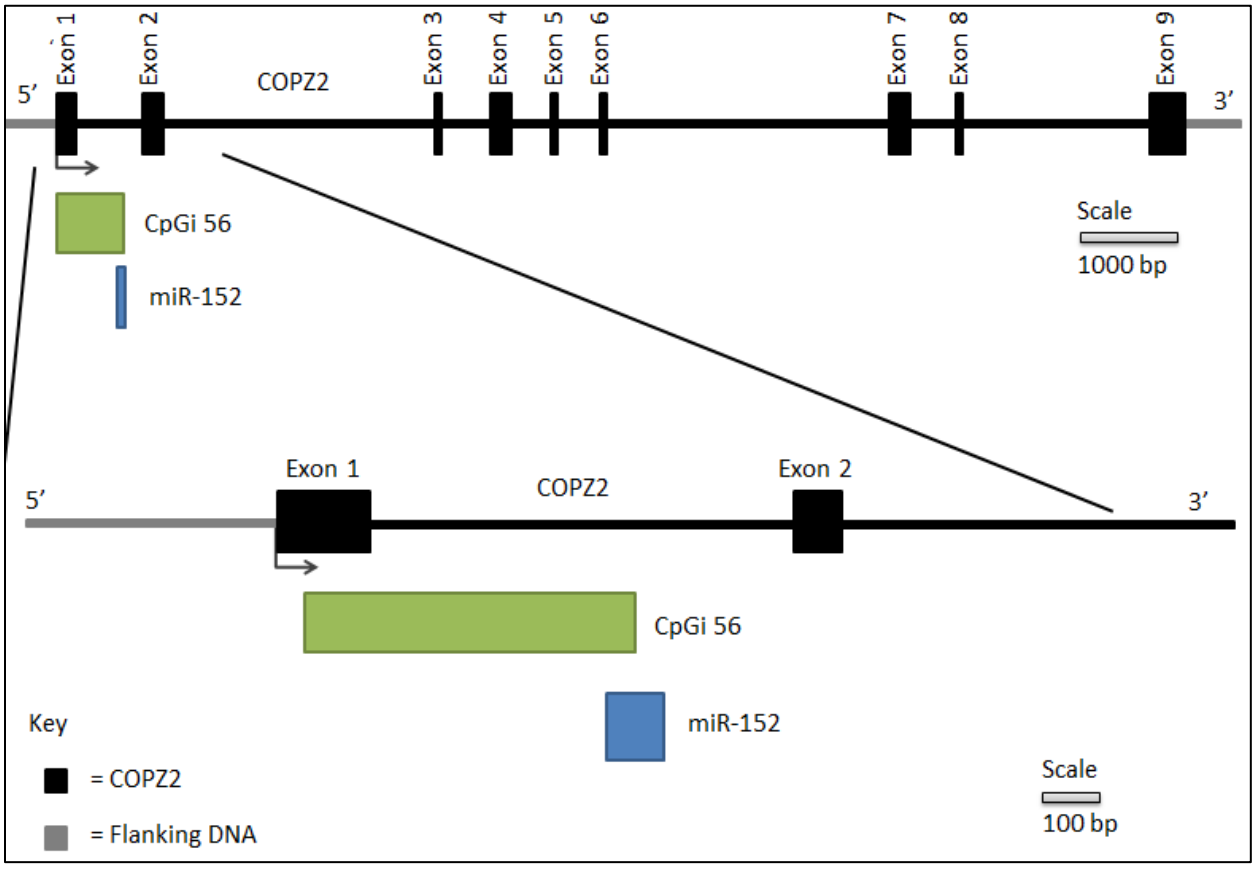
understood as a hallmark of cancer[69, 70]. MiRNA profiling of tumors and cancer cell lines suggests that miRNA levels are also dysregulated in cancer[23]. Yet, the literature remains scant on how, precisely, hypermethylation of a miRNA gene promoter impacts the role of that miRNA and the functioning of its targets.

MiRNA-152 has a well-established role in binding to and repressing DNMT1, a maintenance DNA methyltransferase[56, 62, 63]. Recent work published by Sengupta, et. al. identifies a feedback loop between miR-152 and DNMT1, wherein DNMT1 downregulates miR-152 and miR-152 inhibits DNMT1[71]. When considering that the promoter for COPZ2, the precursor gene of which miR-152 is processed from, is hypermethylated[11], a possible regulatory mechanism involving miR-152, DNMT1, SETDB1, and H3K9me3 begins to emerge. By applying the DNMT1/miR-152 feedback loop to the targeting and regulation of SETDB1, a new mechanism emerges relating DNA methylation to histone methylation.

According to the UCSC Genome Browser, miR-152 is processed from the first intron in the COPZ2 gene. 40 base pairs of the miR-152 transcript and CpG island (CpGi) 56 overlap[72] (Figure 6I.1). There are three sites which can be methylated in this overlapping 40 base pair region. Tsuruta, et. al. noted that methylation at this CpG correlates to decreased levels of miR-152 and COPZ2 transcript[69]. Additionally, within the miR-152 gene lies a CpG island which has been found to be heavily methylated in ovarian cancer[69]. We propose the same would be true in HepG2 cells. We predict that DNMT1 methylates at this particular CpG site, as alluded to in the work by Sengupta, et. al.[71]. The literature fails to connect methylation at the CpG island within miR-152 and global DNA methylation (GDM) to DNMT1 and SETDB1 regulation.

We hypothesize miR-152 downregulation by DNMT1 increases GDM and promoter methylation for COPZ2, while decreasing SETDB1 protein, which thereby decreases H3K9me3 and causes gene silencing. Thus, we predict that gene silencing is achieved by two means resulting from miR-152 suppression by DNMT1: increased GDM and increased histone methylation. This work would demonstrate for the first time that a miRNA can act as a molecular switch between histone and DNA methylation.

A.



B.

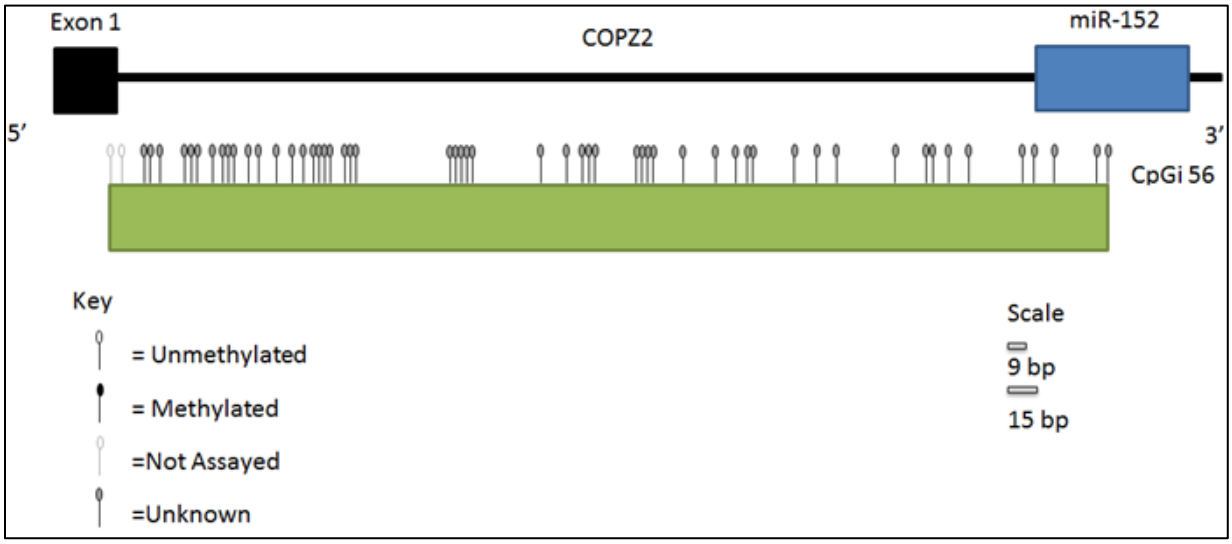


Figure 6I.1 Locations of COPZ2, miR-152, and CpGi 56. (**A**) The COPZ2 gene body ± 500 base pairs is drawn relative to the locations for CpGi 56 and the miR-152 gene. (**B**) Methylation sites on CpGi 56 are shown in relation to miR-152 and COPZ2. Diagrams prepared by Heather White, 2017.

Methods

Cell Culture Conditions and Transfection

HepG2 and IMR-91 cells were transfected twice in order to allow sufficient time for cells to respond to treatment with changes in methylation status. On day 1, cells were passaged and

plated at 5 x 10^5 cells per well of a 6-well dish then transfected as described in Chapter 2 with 5 μL inhibitor, 5 μL water, or 0.5 μL mimic in 500 μL serum-free media with 10 μL HiPerFect transfection reagent. Cells were grown for 24 hours and then 1/4 of cells were passaged into a new well. These cells were treated again on day 2 with the same quantities of reagents. Cells were grown for 48 hours, and then collected on day 4. Treatments were done in triplicate. Cells were washed in DPBS before samples were split in half to have two sets of samples, one for preparing RNA from and the other for preparing gDNA from.

Sample Preparation for Bisulfite Sequencing and LC/MS

Phenol/chloroform extraction was used to collect genomic DNA (gDNA). For LC/MS, 1 μg DNA was hydrolyzed with DNA Degradase Plus (Zymo, E2020) according to the manufacturer's protocol. Briefly, sample was hydrolyzed at 37°C for at least 2 hours and then inactivated at 70°C for 20 minutes. Samples were equilibrated to room temperature, quickly centrifuged, and added to 175 μL of 0.1% formic acid to generate a final concentration of 5 ng/μL.

Bisulfite Sequencing

Samples designated for promoter methylation analysis were split in half in order to bisulfite convert one set, while keeping the other set unconverted. Both sets underwent the same basic work flow. First, gDNA samples were PCR amplified for the region of interest. Next, the insert (either bisulfite converted or unconverted) was cloned into a lacZ reporter vector and

DH5α cells were transformed. White colonies were picked, cultured over night for mini preps, and then DNA was isolated, sequenced, and analyzed. For bisulfite conversion to take place, we used the EZ DNA Methylation Kit from Zymo (catalog #D5001). Unconverted samples were PCR amplified using Q5 High Fidelity Master Mix (NEB, #M0492). The samples then underwent Taq addition of 3'A (NEB #M0273) in order to make them compatible with TOPO TA Cloning (Invitrogen, #K4520-01).

The primers used to generate the region of interest are shown in Table 6I.1 and maps are provided in Figures 6I.2 and 6I.3 graphically showing where the primers bind. Primers were designed and maps were prepared by Heather White (2016).

Primer Abbreviation	Primer Description	Primer Sequence (5'→ 3')	Product Size (base pairs)
FWURG6	Unconverted DNA forward region 6	CATGCACTGACTGCTCCAGA	576
RVURG6	Unconverted DNA reverse region 6	CTGGCCACGTCCGCA	
FWBSRG1	Bisulfite Converted DNA forward region 1	GTTAAGGAGGAGAGAGGGAGAGGAGAAGTGGG	400
RVBSRG1	Bisulfite Converted DNA reverse region 1	TCTAAATTAACTACTCCCCCTAACCCTACTC	
FWBSRG2	Bisulfite Converted DNA forward region 2	GGGTAGGGATAGAGTAGGGTTAGGGGGAGTAG	466
RVBSRG2	Bisulfite Converted DNA reverse region 2	AAAATCCTTCCRAACCCAAATTCTATCATACAC	

Table 6I.1 Primers used to generate PCR products for bisulfite converted and unconverted DNA. The names of primers, a brief description of the areas they

span, the sequences are provided. The size of the final product resulting when primer pairs are used together is also shown.

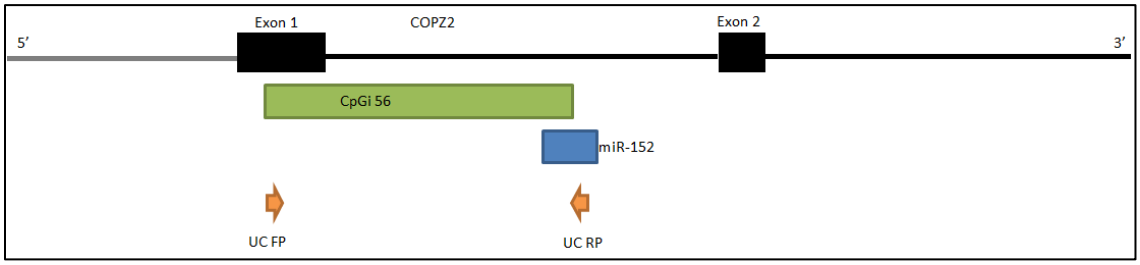


Figure 6I.2 Graphical representation of the primer recognition sites (unconverted DNA). The primer recognition sites within the COPZ2 gene to PCR amplify unconverted DNA are shown. UC FP denotes the unconverted forward primer, and UC RP denotes the unconverted reverse primer.

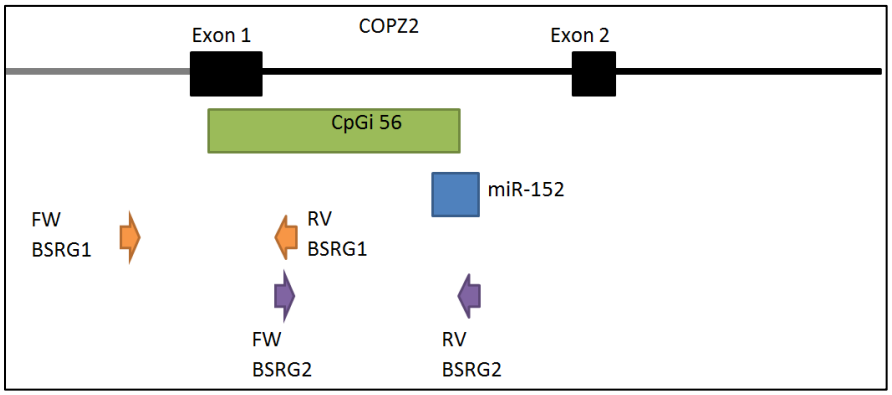


Figure 6I.3 Graphical representation of the primer recognition sites (bisulfite converted DNA). The primer recognition sites within the COPZ2 gene to PCR amplify bisulfite converted DNA are shown. FW and RV stand for forward and

reverse, respectively. BSRG is an acronym for bisulfite region, to indicate the region which will be bisulfite converted.

qPCR for COPZ2

Total RNA was collected using TRIzol Reagent according to the protocol provided. Following resuspension, RNA was quantified and analyzed for purity with a nanodrop. BioRad's iScript kit was used to generate 1 µg cDNA/sample. 10 ng sample and the TaqMan system were used for qPCR, which was run in triplicate in an Applied Biosystems 7500 machine. Technical replicate C_T values were averaged for COPZ2 and β-ACTIN (βACT). βACT values were subtracted from COPZ2 to generate ΔC_T. $\Delta\Delta C_T$ was found by subtracting the water-treated ΔC_T from other sample ΔC_T values. This was done for each water-treated ΔC_T replicate (i.e. anti-miR-152-3p treated replicate #1 ΔC_T – water-treated replicate #1 ΔC_T, anti-miR-152-3p treated replicate #1 ΔC_T – water-treated replicate #2 ΔC_T, anti-miR-152-3p treated replicate #1 ΔC_T – water-treated replicate #3 ΔC_T, etc.). Fold change was then calculated as $2^{-(avg\Delta CT)}$ for each sample. Significance was determined using a 1-way ANOVA with raw C_T values to asses if there were significant differences among biological replicates.

Preliminary Data

qPCR for COPZ2

COPZ2 expression increased as compared to water treated HepG2 cells 12 hours following anti-miR-152-3p or miR-152-3p mimic treatment (Figure 6I.4). These samples were the same ones used to obtain SETDB1 expression data in Chapter 2. However, even when cells were transfected twice, there was no significant difference between the treatment conditions in HepG2 cells.

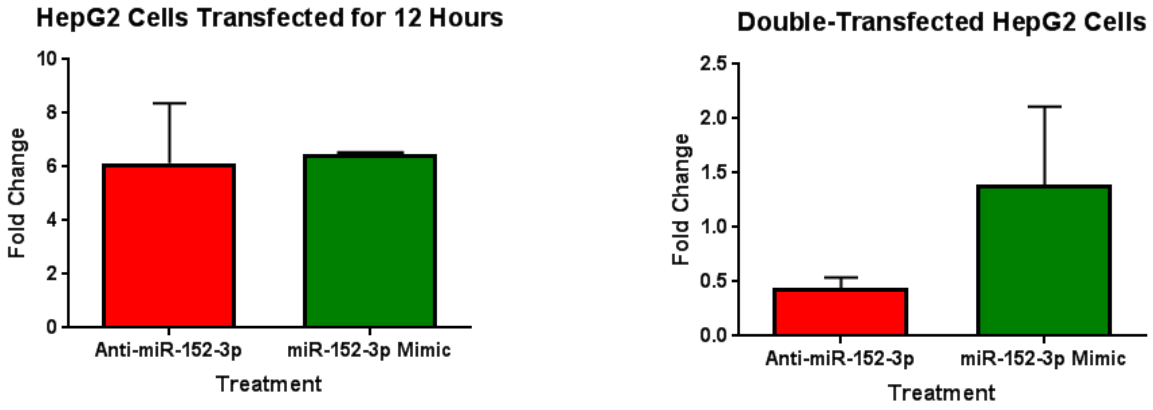


Figure 6I.4 COPZ2 expression in HepG2 cells. Cells were treated with anti-miR-152-3p and miR-152-3p mimic. Fold change was normalized to both β actin expression and COPZ2 expression in water-treated cells. There was no significant difference ($p > 0.05$) between COPZ2 fold change in anti-miR-152-3p and miR-152-3p mimic treated HepG2 cells.

As with HepG2 cells, COPZ2 expression increased regardless of treatment as compared to control conditions in IMR-91 cells transfected for 12 hours (Figure 6I.5). Oddly, results could not be analyzed for IMR-91 samples transfected twice as there appears to be an unanticipated effect on our chosen normalization control, β-actin (Table 6I.2).

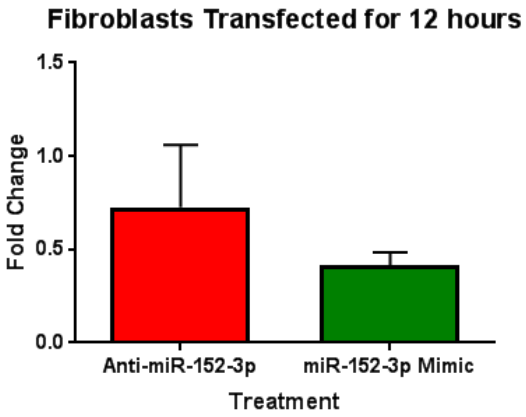


Figure 6I.5 COPZ2 expression in IMR-91 cells. Cells were treated with anti-miR-152-3p and miR-152-3p mimic for 12 hours. Fold change was normalized to both β actin expression and COPZ2 expression in water-treated cells. There was no significant difference ($p > 0.05$) between COPZ2 fold change in anti-miR-152-3p and miR-152-3p mimic treated IMR-91 cells.

		Beta Actin			COPZ2		
Treatment Condition	Bio. Replicate	C_T	Avg C_T	C_T Std Dev	C_T	Avg C_T	C_T Std Dev
dH20	1	32.86	33.13	0.65	32.50	32.28	0.27
		32.52			31.90		
		34.03			32.43		
	2	30.32	30.31	0.23	30.09	29.95	0.34
		30.04			30.28		
		30.59			29.48		
	3	33.13	34.03	0.88	31.71	31.56	0.21
		35.22			31.26		
		33.75			31.72		
anti-miR	1	36.79	36.38	0.95	32.88	32.78	0.13
		35.06			32.87		
		37.28			32.60		
	2	U	U	U	35.19	34.24	0.69
		U			33.91		
		U			33.60		
	3	U	U	U	35.53	35.87	0.33
		U			36.20		
		U			U		
mimic	1	U	U	U	34.28	34.25	0.29
		U			34.60		
		U			33.88		
	2	37.10	36.10	1.04	33.39	33.20	0.29
		34.66			33.42		
		36.53			32.80		
	3	U	35.58	0.00	37.13	36.76	0.26
		35.58			36.58		
		U			36.58		

Table 6I.2 Cycle threshold (C_T) values for IMR-91 cells transfected twice. C_T values were undetermined (U) for the majority of IMR-91 cells transfected with anti-miR-152-3p or miR-152-3p mimic.

Bisulfite Sequencing

Sequence alignments for preliminary data can be found in Appendix II: Supplementary Data.

II. miR-152-3p in Hepatocyte Differentiation

Introduction

MiRNAs can be studied in the contexts of development and differentiation. By determining which miRNAs are expressed during the different stages of these processes, we can begin to modulate them and drive pluripotency, lineage specification, and cell fate for therapeutic purposes. For example, miR-145 targets the 3'UTRs and down-regulates key pluripotency genes, OCT4, SOX2, and KLF4. MiR-145 thereby suppresses pluripotency in human embryonic stem cells (hESCs). One might then expect inhibition of miR-145 would induce pluripotency. Yet, Xu, et. al. found differentiation was actually impaired, despite elevated levels of OCT4, SOX2, and KLF4[73]. Thus, it is important to go beyond simply determining which miRNAs are expressed during stages of development or differentiation and to actually study the effects of individual miRNAs.

MiR-152 may have a role in hepatocyte differentiation. According to work by Kim, et.al., miR-152 was found to be among the top ten miRNAs which increased in expression during the transition from definitive endoderm to hepatocyte cell during the differentiation of human embryonic stem cells (Figure 6II.1)[19]. This work was performed using hESCs and a 23 day differentiation protocol. However, the group did not look at expression of these miRNAs during the intermediate stages of hepatic lineage induction or hepatoblast formation. Nor did they conduct studies on the necessity or functions of any of these miRNAs. However, other research groups have examined some of the miRNAs identified to change in expression during these

stages. For example, overexpression of miR-122 has been found to promote hepatic differentiation[74]. Likewise, miR-152 may also be involved in hepatocyte differentiation.

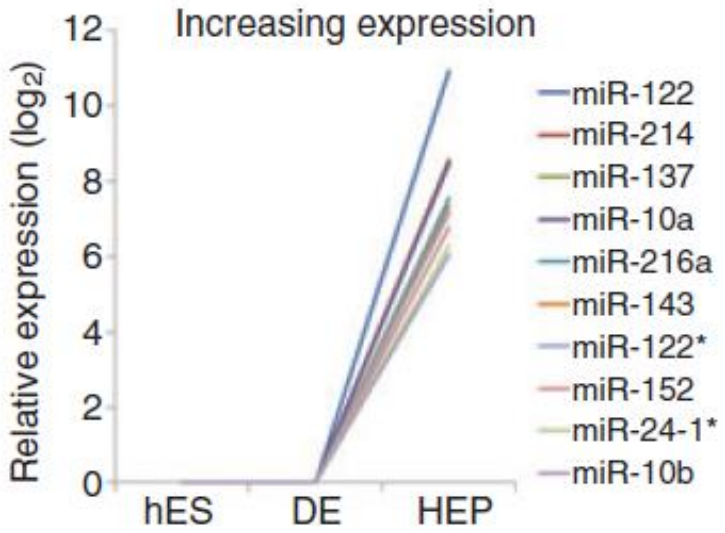


Figure 6II.1 MicroRNAs increasing in expression. This graph was adapted from Kim, et.al.[19] and shows the top 10 miRNAs which increase in expression during differentiation from definitive endoderm (DE) to hepatocytes (HEP).

MiR-152-3p has three relevant categories of targets (Table S.1) which may implicate it in hepatocyte differentiation. The first category is methylation and includes SETDB1 and DNMT1. As established by the work presented in Chapters 2 through 5, miR-152-3p enhances SETDB1 activity. Others have shown that miR-152-3p also down-regulates DNMT1[56, 62, 63]. This feedback loop may facilitate the expression of hepatic genes, while silencing stem cell genes. One way to study this would be to complete a qPCR analysis of LDLr, CYP genes, and AFP/ALB on anti-

miR-152-3p, miR-152-3p mimic, and control (water-treated) cells throughout the stages of differentiation. If there is a significant difference among the treatment and control groups for any of those genes, further studies can be done to confirm that it is due to miR-152-3p effects on methylation.

The second relevant category of miR-152-3p targets is generally involved in differentiation. This includes FGF2, KLF4, and WNT1. The final category of miR-152-3p targets is insulin-signaling related. Insulin-signaling may be required for *in vitro* hepatocyte differentiation (Winfried Kruger, 2013). Thus, miR-152-3p targeting of IGF1R and IRS1 may be relevant. Finally, the possibility exists that multiple targets may be relevant, should miR-152-3p be found necessary for hepatocyte differentiation. A recent study published by Fráguas, et. al. showed miR-29 targeting of DNA demethylation enzymes and wnt signaling molecules impairs early reprogramming[75].

With so many possible avenues for miR-152-3p involvement in hepatocyte differentiation, the central questions are:

1. Does miR-152-3p affect the progress of cells from definitive endoderm to hepatocyte?
2. If miR-152-3p affects the progress of cells from definitive endoderm to hepatocyte, via which target(s) does this occur?
3. Can modulation of miR-152-3p positively contribute to hepatocyte differentiation?

By addressing these questions, one may be able to improve the quality of HLCs, shorten differentiation time, and determine the necessity and function of miR-152 in stem cell differentiation to hepatocytes.

There are several steps that must be taken prior to embarking on any experiments that address the specific questions presented here. First, if an iPSC model is to be used, the lines must be validated for pluripotency. Next, the differentiation protocol should be qualified. The hepatocyte differentiation protocol employed in our laboratory is summarized in Figure 6II.2. Finally, a protocol by which adherent cells growing in multiple layers can be efficiently transfected with miRNA inhibitors and mimics should be developed. Once these challenges are overcome, a simple first experiment is to profile miR-152 expression across all 4 stages of differentiation. Next, mRNA and protein expression analysis should be done for the 7 relevant targets described. Finally, miR-152-3p inhibitor and mimic can be introduced and the effects can be studied. Appropriate follow-up experiments can be designed to address any new questions that emerge.

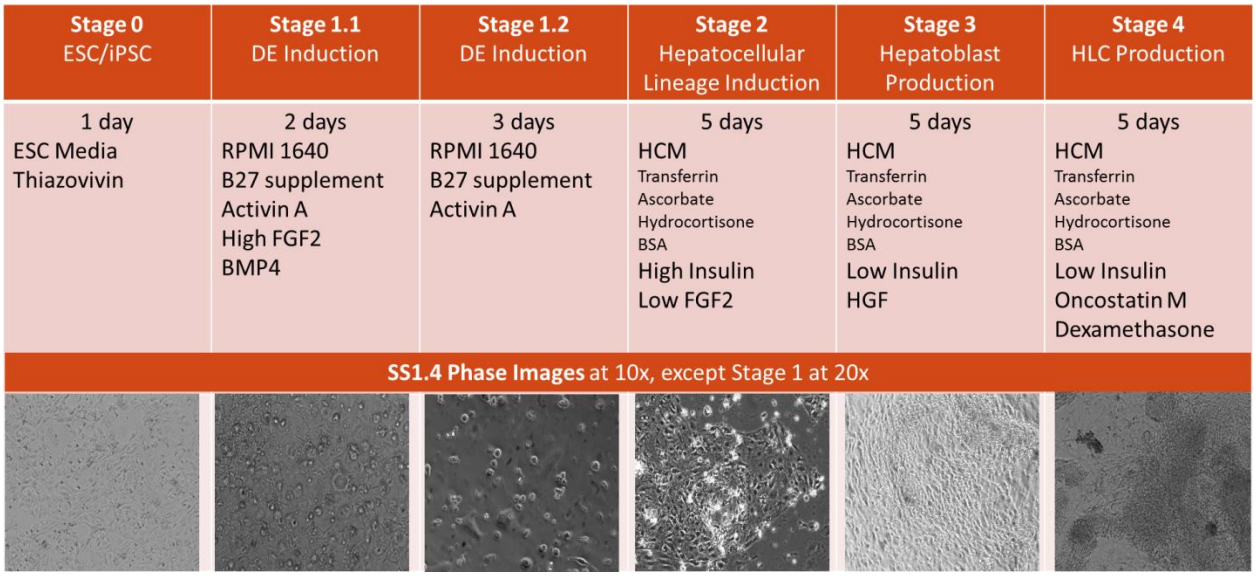


Figure 6II.2 Hepatocyte differentiation protocol. Our research group's hepatocyte differentiation protocol is summarized and phase/contrast images of

cells at each stage are shown. There are 4 stages of hepatocyte differentiation. Stage progression depends on the composition of growth media. Cell morphology begins to resemble hepatocytes at Stage 2.

The data presented represents preliminary findings and must be followed up before any conclusions regarding the role of miR-152 in hepatocyte differentiation can be made.

Methods

iPSC Culture Conditions

Induced pluripotent stem cell line SS1.4 was generated from human fibroblast cell line IMR-91 (Coriell, #I91S-04). Cells were grown at 37°C in a 5% CO_2 incubator under sterile conditions on hESC-qualified Matrigel (Corning #354277) in mTeSR1 (Stem Cell technologies #05857). Media was changed daily and cells were mechanically passaged every 4-6 days prior to hepatocyte differentiation.

Hepatocyte Differentiation

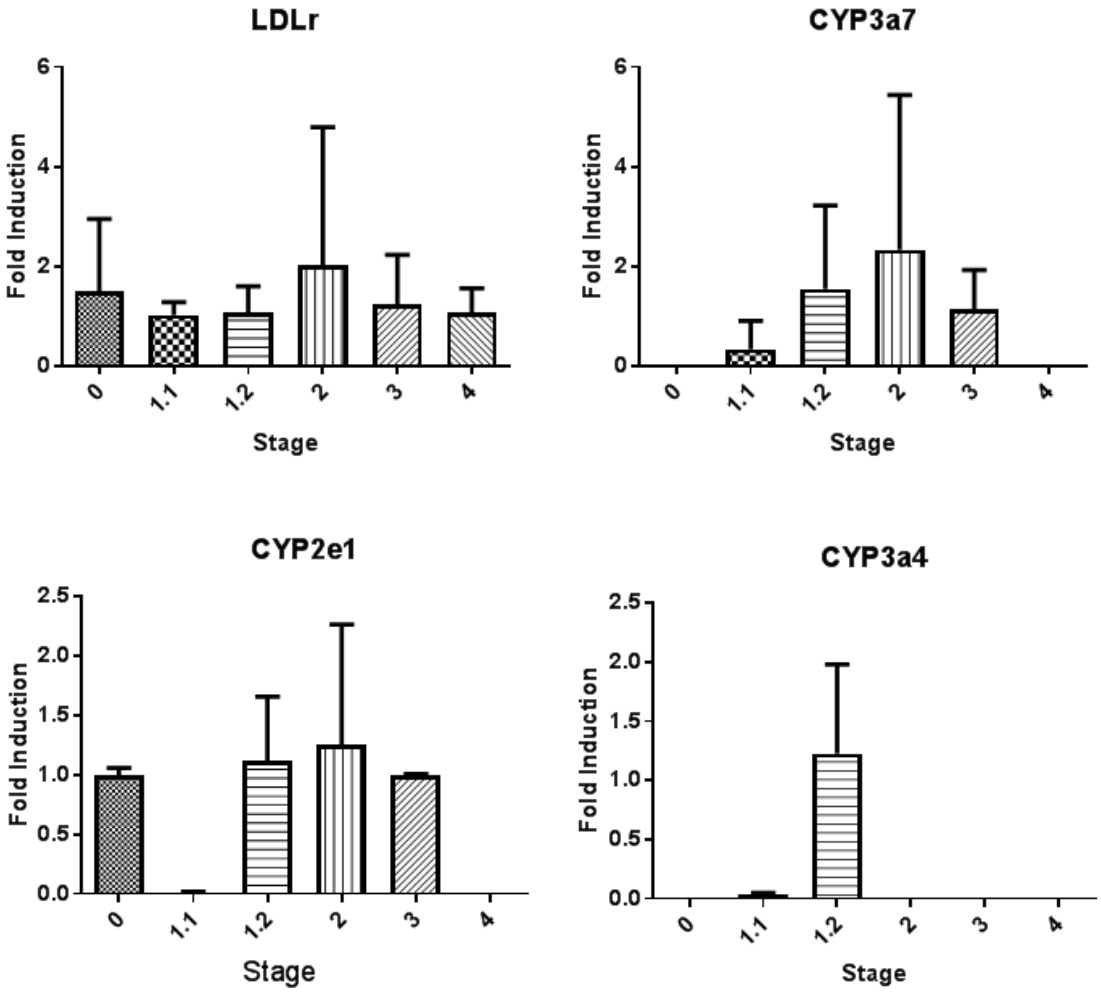
On day 0, SS1.4 cells were dissociated using Accutase and seeded at 100,000 cells per cm^2 in mTESR1 supplemented with 1x Thioazovivin. On day 1, media is changed to Stage 1.1 (S1.1) media consisting of RPMI 1640 with 1 x B27, 1 x NEAA, 100 ng/mL Activin A (Prospec Inc.), 20 ng/mL FGF2 (R&D Systems), and 10 ng/mL BMP4 (Prospec Inc.). S1.1 media is replaced with fresh S1.1 media on day 2. On days 3-5, media is changed daily with S1.2 media. The S1.2 media composition is RPMI 1640 with 1 x B27, 1 x NEAA, and 100 ng/mL Activin A. On day 6, S1.2 media is replaced with HCM.S2. The base media, HCM (Promocell), is supplemented with Transferrin, Ascorbate, Hydrocrtisone, and BSA. 5 μg/mL Insulin, 20 ng/mL BMP4, and 10 ng/mL FGF2 are added to create HCM.S2, which is changed daily for 5 days. On day 11, HCM.S2 is replaced with HCM.S3 (supplemented HCM with 20 ng/mL Insulin, 20 ng/mL HGF). Media is again changed daily for 5 days. On day 16, media is changed to HCM.S4, which consists of supplemented HCM with 20 ng/mL Insulin, 10 ng/mL OSM (dissolved in 0.3% BSA/PBS) and 0.1 μM Dexamethsone. Media is changed daily for 5 days. This protocol was developed by Winfried Kruger (2013).

qPCR for miR-152

RNA was isolated from cells at each stage of differentiation using miRNeasy Mini Kit (Qiagen, # 217004). cDNA was generated using the miScript PCR Starter Kit (Qiagen, #218193). miScript Primer Assays for RNU6 and miR-152 were used in SYBR green-based qPCR according to kit instructions on a Bio-Rad CFX-96 Real-Time System.

Preliminary Data

HLC Validation



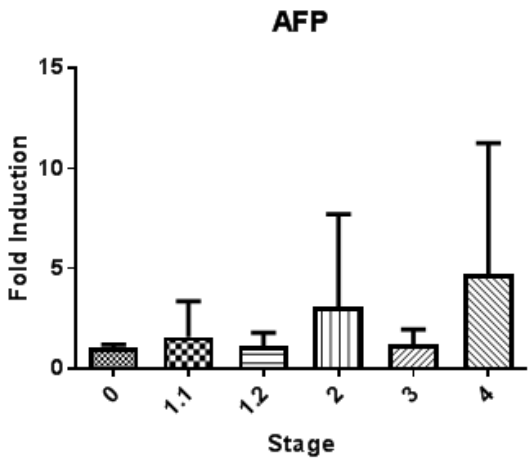


Figure 6II.3 Fold induction of HLC genes. Fold induction was calculated according to the data analysis method published by Livak, et. al.[76]. We observed an incomplete progression to HLC. Gene expressions were normalized to β-actin. Differentiation to HLC was performed according to a protocol written by Winfried Kruger (2013).

qPCR for miR-152

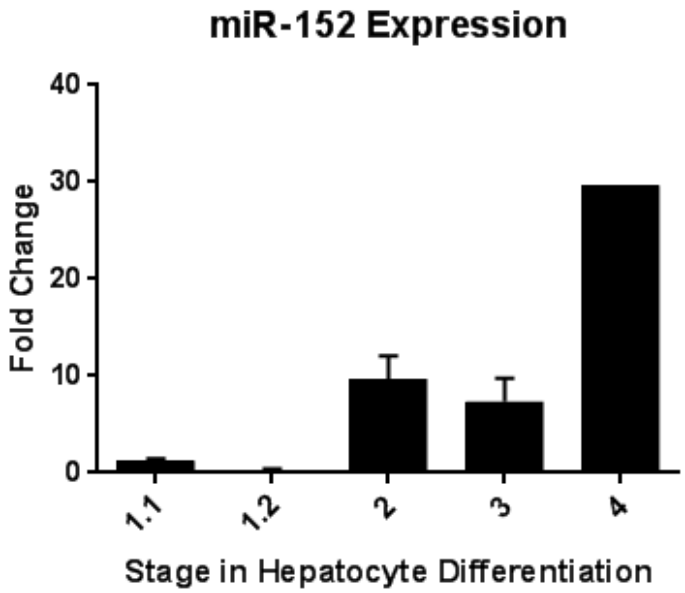


Figure 6II.4 miR-152 expression during HLC induction from iPSC. Expression was calculated as fold change from Day 0 iPSC line SS1.4 derived from IMR-91. MiR-152 levels showed a positive trend during Stage 4, though significance could not be determined at Stage 4 due to sample size. Two values were obtained for fold change at Stage 4, they were 29.7 and 131.2. The fold change of 29.7 was chosen as the representative value for depiction, though both values illustrated the increase in miR-152 expression seen at Stage 4.

REFERENCES

1. Bartel DP. MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 2004 Jan 23;116(2):281-97.

2. Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell 2013 Apr 25;153(3):654-65.

3. Orom UA, Nielsen FC, Lund AH. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 2008 May 23;30(4):460-71.

4. Masaki T, Arend KC, Li Y, Yamane D, McGivern DR, Kato T, Wakita T, Moorman NJ, Lemon SM. miR-122 stimulates hepatitis C virus RNA synthesis by altering the balance of viral RNAs engaged in replication versus translation. Cell Host Microbe 2015 Feb 11;17(2):217-28.

5. Dharap A, Pokrzywa C, Murali S, Pandi G, Vemuganti R. MicroRNA miR-324-3p induces promoter-mediated expression of RelA gene. PLoS ONE 2013;8(11):e79467.

6. Wang H, An W, Cao R, Xia L, Erdjument-Bromage H, Chatton B, Tempst P, Roeder RG, Zhang Y. mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. Mol Cell 2003;12(2):475 487.

7. Minc E, Allory Y, Worman HJ, Courvalin J, Buendia B. Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. Chromosoma 1999;108(4):220 234.

8. Zeng W, Ball AR,Jr, Yokomori K. HP1: Heterochromatin binding proteins working the genome. Epigenetics 2010 May 16;5(4):287-92.

9. Eissenberg JC, Elgin SC. HP1a: A structural chromosomal protein regulating transcription. Trends Genet 2014 Mar;30(3):103-10.

10. Ryu H, Lee J, Hagerty SW, Soh BY, McAlpin SE, Cormier KA, Smith KM, Ferrante RJ. ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in huntington's disease. Proc Natl Acad Sci U S A 2006 Dec 12;103(50):19176-81.

11. Dang Y, Zeng J, He R, Rong M, Luo D, Chen G. Effects of miR-152 on cell growth inhibition, motility, suppression and apoptosis induction in hepatocellular carcinoma cells. Asian Pac J Cancer Prev 2014;15:4969.

12. Wong CM, Wei L, Law CT, Ho DW, Tsang FH, Au SL, Sze KM, Lee JM, Wong CC, Ng IO. Up-regulation of histone methyltransferase SETDB1 by multiple mechanisms in hepatocellular carcinoma promotes cancer metastasis. Hepatology 2016 Feb;63(2):474-87.

13. Henke JI, Goergen D, Zheng J, Song Y, Schüttler CG, Fehr C, Jünemann C, Niepmann M. microRNA-122 stimulates translation of hepatitis C virus RNA. Embo J 2008 European Molecular Biology Organization;27(24):3300-10.

14. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: Targets and expression. Nucleic Acids Res 2008 Jan;36(Database issue):D149-53.

15. Kim YK, Kim VN. Processing of intronic microRNAs. Embo J 2007;26:775.

16. Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Fitzgerald S, Gil L, et al. Ensembl 2016. Nucleic Acids Research 2015 December 19;44(D1):D710-6.

17. Mourelatos Z, Dostie J, Paushkin S, Sharma A, Charroux B, Abel L, Rappsilber J, Mann M, Dreyfuss G. miRNPs: A novel class of ribonucleoproteins containing numerous microRNAs. Genes & Dev 2002;16:720.

18. Diederichs, S. & Haber, D. A. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. Cell 2007;131:1097.

19. Kim N, Kim H, Jung I, Kim Y, Kim D, Han Y. Expression profiles of miRNAs in human embryonic stem cells during hepatocyte differentiation. Hepatology Research 2011;41(2):170-83.

20. Crist CG, Montarras D, Pallafacchina G, Rocancourt D, Cumano A, Conway SJ, Buckingham M. Muscle stem cell behavior is modified by microRNA-27 regulation of Pax3 expression. Proc Natl Acad Sci U S A 2009 Aug 11;106(32):13383-7.

21. Nampoothiri SS, Rajanikant GK. Decoding the ubiquitous role of microRNAs in neurogenesis. Mol Neurobiol 2016 Feb 24.

22. Wang S, Wang L, Dou L, Guo J, Fang W, Li M, Meng X, Man Y, Shen T, Huang X, et al. MicroRNA 152 regulates hepatic glycogenesis by targeting PTEN. Febs J 2016 May;283(10):1935-46.

23. Chen H, Liu H, Zou H, Chen R, Dou Y, Sheng S, Dai S, Ai J, Melson J, Kittles RA, et al. Evaluation of plasma miR-21 and miR-152 as diagnostic biomarkers for common types of human cancers. J Cancer 2016 Feb 5;7(5):490-9.

24. Chou CH, Chang NW, Shrestha S, Hsu SD, Lin YL, Lee WH, Yang CD, Hong HC, Wei TY, Tu SJ, et al. miRTarBase 2016: Updates to the experimentally validated miRNA-target interactions database. Nucleic Acids Res 2016 Jan 4;44(D1):D239-47.

25. Veenendaal MV, Painter RC, de Rooij SR, Bossuyt PM, van der Post JA, Gluckman PD, Hanson MA, Roseboom TJ. Transgenerational effects of prenatal exposure to the 1944-45 dutch famine. Bjog 2013 Apr;120(5):548-53.

26. Painter RC, Osmond C, Gluckman P, Hanson M, Phillips DI, Roseboom TJ. Transgenerational effects of prenatal exposure to the dutch famine on neonatal adiposity and health in later life. Bjog 2008 Sep;115(10):1243-9.

27. Turner B. Reading signals on the nucleosome with a new nomenclature for modified histones. Nat Struct Mol Biol 2005;12:110.

28. Jenuwein T. Re-SET-ting heterochromatin by histone methyltransferases. Trends Cell Biol 2001;11(6):266 273.

29. Bedford M. Arginine methylation at a glance. J Cell Sci 2007;120(24):4243.

30. Zhang X, Wen H, Shi X. Lysine methylation: Beyond histones. Acta Biochim Biophys Sin 2012;44(1):14-27.

31. Qian C, Zhou MM. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. Cell Mol Life Sci 2006 Dec;63(23):2755-63.

32. Chiang P, Gordon R, Tal J, Zeng G, Doctor B, Pardhasaradhi K, McCann P. **S-adenosylmethionine and methylation**. Faseb J 1996;10(4):471.

33. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 2004;119(7):941 953.

34. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y. Histone demethylation by a family of JmjC domain-containing proteins. Nature 2005;439(7078):811 816.

35. Borgel J, Tyl M, Schiller K, Pusztai Z, Dooley CM, Deng W, Wooding C, White RJ, Warnecke T, Leonhardt H, et al. KDM2A integrates DNA and histone modification signals through a CXXC/PHD module and direct interaction with HP1. Nucleic Acids Res 2016:gkw979.

36. Rivera C, Saavedra F, Alvarez F, Diaz-Celis C, Ugalde V, Li J, Forne I, Gurard-Levin ZA, Almouzni G, Imhof A, et al. Methylation of histone H3 lysine 9 occurs during translation. Nucleic Acids Res 2015 Oct 30;43(19):9097-106.

37. Loyola A, Tagami H, Bonaldi T, Roche D, Quivy JP, Imhof A, Nakatani Y, Dent SY, Almouzni G. The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. EMBO Rep 2009 Jul;10(7):769-75.

38. Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ,3rd. SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev 2002 Apr 15;16(8):919-32.

39. Hashimoto H, Vertino PM, Cheng X. Molecular coupling of DNA methylation and histone methylation. Epigenomics 2010(5):657.

40. Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J, Billis K, García Girón C, Hourlier T, et al. The ensembl gene annotation system. Database 2016 January 1;2016:baw093-.

41. Lee JK, Kim KC. DZNep, inhibitor of S-adenosylhomocysteine hydrolase, down-regulates expression of SETDB1 H3K9me3 HMTase in human lung cancer cells. Biochem Biophys Res Commun 2013 Sep 6;438(4):647-52.

42. Ishimoto K, Kawamata N, Uchihara Y, Okubo M, Fujimoto R, Gotoh E, Kakinouchi K, Mizohata E, Hino N, Okada Y, et al. Ubiquitination of lysine 867 of the human SETDB1 protein upregulates its histone H3 lysine 9 (H3K9) methyltransferase activity. Plos One 2016;11(10):e0165766.

43. Tachibana M, Sugimoto K, Fukushima T, Shinkai Y. Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. J Biol Chem 2001 July 6;276(27):25309.

44. Barski A, Cuddapah S, Cui K, Roh T, Schones D, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. Cell 2007 May 18;129(4):823.

45. Peters AHFM, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AAHA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, et al. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell 2003 12;12(6):1577-89.

46. Sims J, Houston S, Magazinnik T, Rice J. A trans-tail histone code defined by monomethylated H4 lys-20 and H3 lys-9 demarcates distinct regions of silent chromatin. J Biol Chem 2006;281(18):12760.

47. Rougeulle C, Chaumeil J, Sarma K, Allis C, Reinberg D, Avner P, Heard E. Differential histone H3 lys-9 and lys-27 methylation profiles on the X chromosome. Mol Cell Biol 2004;24(12):5475.

48. Escamilla-Del-Arenal M, da Rocha S, Spruijt C, Masui O, Renaud O, Smits A, Margueron R, Vermeulen M, Heard E. Cdyl, a new partner of the inactive X chromosome and potential reader of H3K27me3 and H3K9me2. Mol Cell Biol 2013;33(24):5005.

49. Bannister A, Zegerman P, Partridge J, Miska E, Thomas J, Allshire R, Kouzarides T. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 2001 March 1;410(6824):120.

50. Verschure P, van der Kraan I, de Leeuw W, van der Vlag J, Carpenter A, Belmont A, van Driel R. In vivo HP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation. Mol Cell Biol 2005;25(11):4552.

51. Bártová E, Večeřa J, Krejčí J, Legartová S, Pacherník J, Kozubek S. The level and distribution pattern of HP1β in the embryonic brain correspond to those of H3K9me1/me2 but not of H3K9me3. Histochem Cell Biol 2016;145(4):447.

52. Aucott R, Bullwinkel J, Yu Y, Shi W, Billur M, Brown JP, Menzel U, Kioussis D, Wang G, Reisert I, et al. HP1-β is required for development of the cerebral neocortex and neuromuscular junctions. J Cell Biol 2008;183(4):597.

53. Sugimoto K, Tasaka H, Dotsu M. Molecular behavior in living mitotic cells of human centromere heterochromatin protein HPLalpha ectopically expressed as a fusion to red fluorescent protein. Cell Struct Funct 2001 Dec;26(6):705-18.

54. Huang S, Xie Y, Yang P, Chen P, Zhang L. HCV core protein-induced down-regulation of microRNA-152 promoted aberrant proliferation by regulating Wnt1 in HepG2 cells. PLoS One 2014 Jan 9;9(1):e81730.

55. Mancini M, Saintigny G, Mahe C, Annicchiarico-Petruzzelli M, Melino G, Candi E. MicroRNA-152 and -181a participate in human dermal fibroblasts senescence acting on cell adhesion and remodeling of the extra-cellular matrix. Aging (Albany NY) 2012 Nov;4(11):843-53.

56. Huang J, Wang Y, Guo Y, Sun S. Down-regulated microRNA-152 induces aberrant DNA methylation in hepatitis B virus-related hepatocellular carcinoma by targeting DNA methyltransferase 1. Hepatology 2010 Jul;52(1):60-70.

57. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife 2015 Aug 12;4:10.7554/eLife.05005.

58. Katoh H, Shibata T, Kokubu A, Ojima H, Fukayama M, Kanai Y, Hirohashi S. Epigenetic instability and chromosomal instability in hepatocellular carcinoma. Am J Pathol 2006 Apr;168(4):1375-84.

59. Watson M, Chow S, Barsyte D, Arrowsmith C, Shankey TV, Minden M, Hedley D. The study of epigenetic mechanisms based on the analysis of histone modification patterns by flow cytoametry. Cytometry A 2014;85:78.

60. Kelly TK, De Carvalho DD, Jones PA. Epigenetic modifications as therapeutic targets. Nat Biotechnol 2010 Oct;28(10):1069-78.

61. Labahn J, Granzin J, Schluckebier G, Robinson DP, Jack WE, Schildkraut I, Saenger W. Three-dimensional structure of the adenine-specific DNA methyltransferase M.taq I in complex with the cofactor S-adenosylmethionine. Proc Natl Acad Sci 1994;91(23):10957.

62. Ji W, Yang L, Yuan J, Yang L, Zhang M, Qi D, Duan X, Xuan A, Zhang W, Lu J, et al. MicroRNA-152 targets DNA methyltransferase 1 in NiS-transformed cells via a feedback mechanism. Carcinogenesis 2012;34(2):446 453.

63. Wang J, Bian Y, Wang Z, Li D, Wang C, Li Q, Gao X. MicroRNA-152 regulates DNA methyltransferase 1 and is involved in the development and lactation of mammary glands in dairy cows. PLoS ONE 2014;9(7):e101358.

64. Tamaru H, Selker EU. A histone H3 methyltransferase controls DNA methylation in *neurospora crassa*. Nature 2001;414:277.

65. Honda S, Selker EU. Direct interaction between DNA methyltransferase DIM-2 and HP1 is required for DNA methylation in *neurospora crassa*. Mol Cell Biol 2008;28(6044).

66. Freitag M, Hickey PC, Khlafallah TK, Read ND, Selker EU. HP1 is essential for DNA methylation in *neurospora*. Mol Cell 2004;13:427.

67. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007;128:683.

68. Ballestar E, Esteller M. Epigenetic gene regulation in cancer. Adv Genet 2008;61:247-67.

69. Tsuruta T, Kozaki K, Uesugi A, Furuta M, Hirasawa A, Imoto I, Susumu N, Aoki D, Inazawa J. miR-152 is a tumor suppressor microRNA that is silenced by DNA hypermethylation in endometrial cancer. Cancer Res 2011;71(20):6450.

70. Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setién F, Casado S, Suarez-Gauthier A, Sanchez-Cespedes M, Gitt A, et al. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. Cancer Res 2007 American Association for Cancer Research;67(4):1424-9.

71. Sengupta D, Deb M, Rath SK, Kar S, Parbin S, Pradhan N, Patra SK. DNA methylation and not H3K4 trimethylation dictates the expression status of miR-152 gene which inhibits migration of breast cancer cells via DNMT1/CDH1 loop. Exp Cell Res 2016;346:176.

72. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res 2002;12(6):996.

73. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell 2009;137:647.

74. Davoodian N, Lotfi AS, Soleimani M, Mowla SJ. MicroRNA-122 overexpression promotes hepatic differentiation of human adipose tissue-derived stem cells. J Cell Biochem 2014;11(9):1582.

75. Fráguas MS, Eggenschwiler R, Hoepfner J, Schiavinato JLdS, Haddad R, Oliveira LHB, Araújo AG, Zago MA, Panepucci RA, Cantz T. MicroRNA-29 impairs the early phase of reprogramming process by targeting active DNA demethylation enzymes and wnt signaling. Stem Cell Research 2017 3;19:21-30.

76. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. Methods 2001 Dec;25(4):402-8.

APPENDIX I

PROTOCOLS

Transfection of HepG2 or IMR-91 Cells with miRNA Mimic, miRNA Inhibitor, or Water in a 10 cm Dish

(Protocol modified from Qiagen, last updated 09/2013)

Cell seeding and transfection are performed on the same day. Allow 1-2 hours from start-to-finish for preparing the hood, counting cells, seeding cells, and transfecting cells, depending on the number of plates being used. For IMR-91 cells, coat plates in 0.1% gelatin for 30 minutes at 37°C prior to cell seeding. During the 30 minute coating step, or 30 minutes prior to initiating protocol, set up the cell culture hood.

Begin by warming media (IMR-91: DMEM with 4.5 g/L glucose, 100X GlutaMax, and 10% FBS; HepG2: DMEM with 1.0 g/L glucose and 10% FBS) at 37°C in bead or water bath, equilibrating 0.05% Trypsin-EDTA (Life Technologies, #25300054) and Dulbecco's PBS (Life Technologies, #14190) to room temperature, and obtaining fresh ice. Prepare serum-free media (IMR-91: DMEM with 4.5 g/L glucose and 100X GlutaMax; HepG2: DMEM with 1.0 g/L glucose). Label 15 mL tubes for preparing transfection complexes in. Label treatment conditions on the 10 cm dishes for HepG2 cells at this time and for IMR-91 cells when the 30 minute coating step is complete. Label 40 mL Falcon tubes for passaging cells. Sanitize and place

necessary tube stands, 5 mL and 10 mL serological pipettes, pipette aids, micro pipette tips, micro pipettors, etc. in the hood.

Prior to initiating protocol, passage and count cells.

1. Seed 2.4 x 10^6 HepG2 cells or 1.26 x 10^6 IMR-91 cells (under passage #20) in 8.452 mL media in transfecting with mimic or 8.43 mL media if transfecting with water or anti-miR per well of a 10 cm dish.
2. Incubate cells until transfection complexes are formed (after step 4 is complete) at 37°C with 5% CO_2.
3. In the labeled 15 mL tubes, dilute 2.5 µL mimic, 25 µL anti-miR, or 25 µL water in 1.5 mL of the appropriate serum-free media for each cell line. Add 45 µL HiPerFect Transfection Reagent (Qiagen, #301705) per tube. These quantities are fixed per single 10 cm dish. For replicates, multiple these values (ex: For 2 x HepG2 + mimic plates, use: 5 µL mimic, 3 mL DMEM + 1.0 g/L glucose, and 90 µL HiPerFect transfection reagent).
4. Vortex 15 mL tubes containing treatment, media, and transfection reagent briefly to mix. Quick spin centrifuge if need. Incubate at room temperature for 5-10 minutes to form transfection complexes.
5. Add the complexes drop-wise onto cells and swirl the plate to ensure even distribution of reagents and cells.
6. Incubate plates at 37°C with 5% CO_2 until cells are ready to collect.

InFusion Cloning Protocol for Preparing Reporter Vector + SETDB1 3'UTR

1. Prepare linearized vector using restriction enzyme digest (RED) and check background.

    a. **RED pE1330:** 0.5 µL pE1330 DNA (1000ng, Promega, #e1330) + 1 µL SacI-HF (NEB, R3156S) + 1 µL XbaI (NEB, R0145T) + 5 µL NEBuffer + 42.5 µL dH2O → 37°C overnight → 65°C for 20 minutes → add 10 µL sample buffer and gel purify on 1% agarose gel run at 90V with Macherey Nagel gel purification kit (#740609) provided with InFusion Cloning kit. Nanodrop = 948ng/20 µL.

    b. **Negative control**: transformed Stellar Competent Cells with 2.37 µL linearized e1330 (5 ng) according to the Stellar Competent Cells protocol #PT5055-2 provided through Clontech: http://www.clontech.com/US/Products/Cloning_and_Competent_Cells/Competent_Cells/ibcGetAttachment.jsp?cItemId=17313&fileId=5877577&sitex=10020:22372:US Plated 100 µL on LB+amp (A). Pelleted remaining cells at 6000 rpm for 5 min then resuspended in 100 µL and plated on LB+amp (B). After incubation at 37°C overnight, saw 3 colonies on A and 2 colonies on B.

2. Design PCR primers to amplify the SETDB1 3'UTR using the InFusion Primer Design program or by following the guidelines provided in the manual.

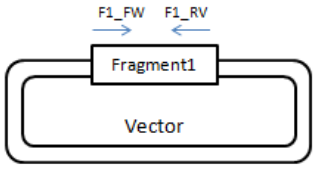
    a. Entered pE1330 vector sequence, selected XbaI and SacI RE, did not preserve restriction site (doing so creates primers that are > 5°C apart), get fragmen.t

    b. Resulted in 14 bases excised and 7336bp linearized vector, inserting fragment based off Genome Browser's SETDB1 3'UTR DNA sequence. Gave:

The result of the In-Fusion Primer design

Both In-FusionR primers are in 5' -> 3' orientation.
Check cloning diagram below to make sure the designed primers will produce correct clone.

Fragment 1 FW Primer (F1_FW)	CTAGTTGTTTAAACGAGGACAGCCTTCTTCCCA	Tm : 96	GC : 45.5 %
Fragment 1 RV Primer (F1_RV)	GCAGGTCGACTCTAGCACAATCTTAGTTTTATCATATT	Tm : 104	GC : 36.8 %



Cloning Diagram

1. PCR Product.
Fragment1(+) :
CTAGTTGTTTAAACGAGGACAGCCTTCTTCCCA.....AATATGATAAAACTAAGATTGTGCTAGAGTCGACCTGC
Fragment1(-) :
GATCAACAAATTTGCTCCTGTCGGAAGAAGGGT.....TTATACTATTTTGATTCTAACACGATCTCAGCTGGACG

2. Linearized vector *Sac* I *Xba* I
Vector(+) : ....GCCGTGTAATTCTAGTTGTTTAAACGAGCT CTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGG....
Vector(-) : ....CGGCACATTAAGATCAACAAATTTGC TCAGCTGGACGTCCGTACGTTCGACTAGGCC....

3. Annealing
Fragment1(+) : 5'- CTAGTTGTTTAAACGAGGACAGCCTTCTTCCCA....AATATGATAAAACTAAGATTGTGCTAGAGTCGACCTGC -3'
Fragment1(-) : 3'- GATCAACAAATTTGCTCCTGTCGGAAGAAGGGT....TTATACTATTTTGATTCTAACACGATCTCAGCTGGACG -5'
Vector(+) : ....GCCGTGTAATTCTAGTTGTTTAAACGAGCT -3' 5'- CTAGAGTCGACCTGCAGGCAT
Vector(-) : ....CGGCACATTAAGATCAACAAATTTGC -5' 3'- TCAGCTGGACGTCCGTA

Output

You can obtain the two types of export file. See below. (If you would like to the detail, See the tutorial.)

All Sequence of Cloning construct	Export File of Vector & Insert Fragment sequence (Text Format)
Cloning Diagram	Export File of Cloning Diagram

*The Tm is incorrect due to glitch in programing. Actual melting temperatures are 55.8°C for FWP and 44.6°C for RVP, according to Integrated DNA Technologies Tm Calculator.

3. PCR amplify the SETDB1 3'UTR using CloneAmp HiFi PCR Premix and HepG2 cDNA

   a. Convert HepG2 RNA to cDNA using the iScript cDNA Synthesis Kit (BioRad, #1708841) by combining the following in a PCR tube and running the Thermocycler on the iScript setting:

      4 µL 5x iScript Reaction Mix

      2.69 µL dH_20

2.31 µL RNA template (1 µg RNA)

1 µL iScript Reverse Transcriptase (add last)

b. In order to generate insert, follow the protocol provided with CloneAmp HiFi PCR Premix according to Option A (98°C for 10 seconds, 55°C for 5 or 15 seconds, 72°C for 5 seconds/kb = 5 seconds total). Prepare the following in a PCR tube and run on CloneAmp setting in the Thermocycler:

12.5 µL PCR Pre-mix

1 µL each of 5 pmol FWP/RVP

1.8 µL of template (90 ng)

8.7 µL dH_20

c. Confirm PCR product size on 1% agarose gel at 90V

4. Either purify PCR product or treat with Cloning Enhancer and follow with appropriate Enzyme Premix treatment protocol (see p 9-10 of User Manual). Conduct positive control in parallel using provided 2 kb control insert.
5. Transform Stellar Competent Cells with Reaction Mixture/Control Insert according to protocol provided (see link under 1b).
6. Pick colonies, create overnights, make mini preps and confirm presence of insert through diagnostic digests (PmeI/DraI + NotI, for example). Sequence sample that looks promising.
7. Maxi prep pE1330+SETDB1 3' UTR (here on called pE1330S3U)/pE1330 (to be used as positive control since no insert is present)
8. Transfect HepG2 with pE1330S3U/pE1330 using Lipofectamine 2000 (optimization will be required – follow protocol and analyze cells at 24, 48, and 72 hours post-transfection) to determine experimental parameters for all transfections. Conduct in 6 replicates.

Lipofectamine 2000 Transfection Protocol for HepG2 and IMR-91 Cells

1. **Day 0.** Plate 10,000 cells/well of a 96-well white wall plate (Fisher Scientific, #08-771-26) in 75 µL growth media. Plate 6 replicates per sample, remembering to include extra for positive and negative controls as well as a blank for when the plate is read.
2. **Day 1.** Prepare reagents for transfection as described:

   A. Combine 1200 ng DNA + Opti-MEM to 60 µL.

Name	µL DNA	µL Opti-MEM
pE1330	0.60	59.4
pE1330S3U	14.88	45.1
pE1330S3Upbs1$^-$	8.00	52.0
pE1330S3Upbs2$^-$	2.24	57.8
pE1330S3Upbs3$^-$	2.14	57.9

   These DNA samples are stored at -20°C in the Cell Culture room.

   B. Combine 300 µL Opti-MEM + 24 µL Lipofectamine 2000
3. For sample wells, combine equal parts (60 µL A + 60 µL B) of diluted DNA with diluted Lipofectamine 2000 in a fresh microcentrifuge tube. For control wells, combine 120 µL Opti-MEM + 4.8 µL Lipofectamine 2000 in a fresh microcentrifuge tube. Incubate tubes at room temperature for 5 minutes. Transfect cells with 10 µL solution per well (sample: 100 ng DNA/well). There will be enough solution to transfect 6 x HepG2 sample wells, 6 x IMR-91 sample wells, 6 x HepG2 control wells, and 6 x IMR-91 control wells.
4. **Day 2.** Change media to a volume of 75 µL/well and begin the Dual Glo Luciferase Assay System Protocol.

Dual Glo Luciferase Assay System Protocol and Data Analysis

(Modified from Promega, #E2920, last updated 09/2015)

1. Prepare reagents according to manufacturer's protocol
2. Equilibrate reagents and cells to room temperature
3. Add 75 µL/well of Dual-Glo Reagent then wait 10 minutes to allow cells to lyse
4. Measure firefly luminescence on the LMAX II luminometer
5. Add 75 µL/well of Dual-Glo Stop & Glo Reagent and wait 10 minutes before measuring *Renilla* luminescence
6. Analyze data accordingly:
    A. Subtract Blank well luminescence from raw luciferase luminescence values

$$LUC_{raw} - BLANK = LUC$$

    B. Subtract Blank well luminescence from raw *Renilla* luminescence values

$$REN_{raw} - BLANK = REN$$

    C. Create a luciferase luminescence to *Renilla* luminescence ratio by dividing the background subtracted values

LUC/REN = luciferase/*Renilla* ratio

APPENDIX II

SUPPLEMENTARY DATA

Confirmed hsa-miR-152-3p Targets

Target	Description of Gene/Protein
ADAM17	Disintegrin involved in autoimmune diseases
B4GALT5	Glycoprotein; Beta-1,4-galactosyltransferase 5
C18orf25	Protein uncharacterized
CCKBR	Gastrin and cholecystokinin receptor
CCND1	Cyclin D 1
CD151	Cell surface glycoprotein, member of transmembrane 4 superfamily proteins
CLOCK	Circadian Locomotor Output Cycles Kaput
CSF1	Macrophage colony stimulating factor; cytokine
DNMT1	Maintenance DNA methyltransferase
EMP1	Epithelial Membrane Protein 1; localized to tight junctions
FGF2	Fibroblast growth factor 2
FGFR3	Fibroblast growth factor receptor 3; promotes HCC metastasis
HLA-G	Histocompatibility antigen
IGF1R	Insulin growth factor
IRS1	Insulin receptor substrate
ITGA5	Fibronectin and fibrinogen receptor; integrin alpha chain family
KIAA0232	Nuclear phosphoserine protein
KLF4	Transcription factor; regulates cell proliferation, differentiation, apoptosis and can be used for reprogramming somatic cells
KRAS	GTPase
MAFB	Transcriptional activator or repressor in lineage-specific hematopoiesis regulation
MOSPD1	Implicated in mesenchymal and epithelial cell development
MXD1	MAX dimerization protein 1; forms a DNA-binding protein complex, MAD-MAX
PHACTR2	Phosphatase and Actin Regulator 2; involved in platelet activation, signaling and aggregation
RMND5A	Required For Meiotic Nuclear Division 5 Homolog A
SETDB1	Histone methyltransferase
STRADB	Pseudokinase involved in the activation of serine/threonine kinase 11; involved in cell polarity and metabolism
TACC3	Motor spindle protein
TGFA	Transforming growth factor; epidermal growth factor
WNT1	Wnt-signaling, oncogenesis, development and differentiation

Table S.1 List of miR-152-3p targets. The targets listed here all have strong evidence supporting their interactions with miR-152-3p. Strong evidence is defined as originating from a reporter assay, western blot, or qPCR[24].

Lipofectamine 2000 Transfection Optimization Data

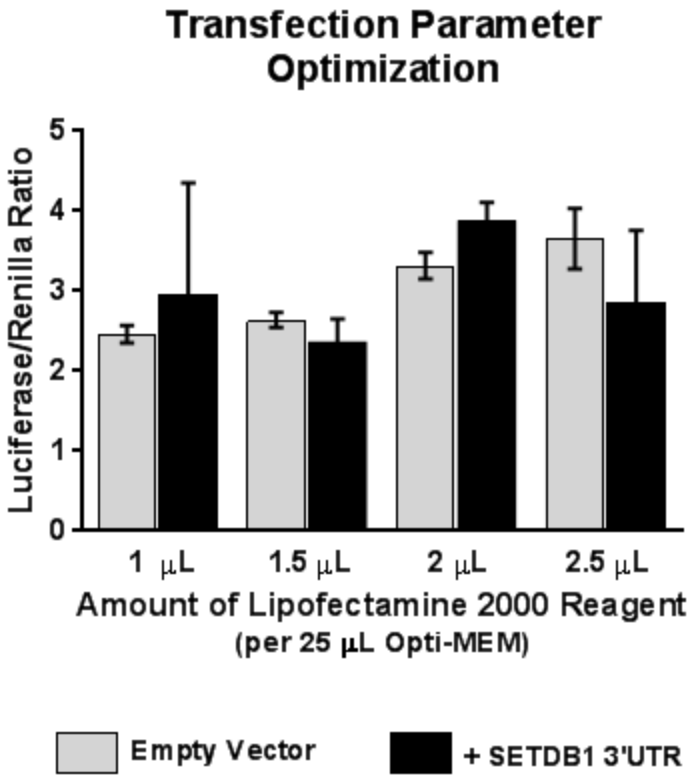
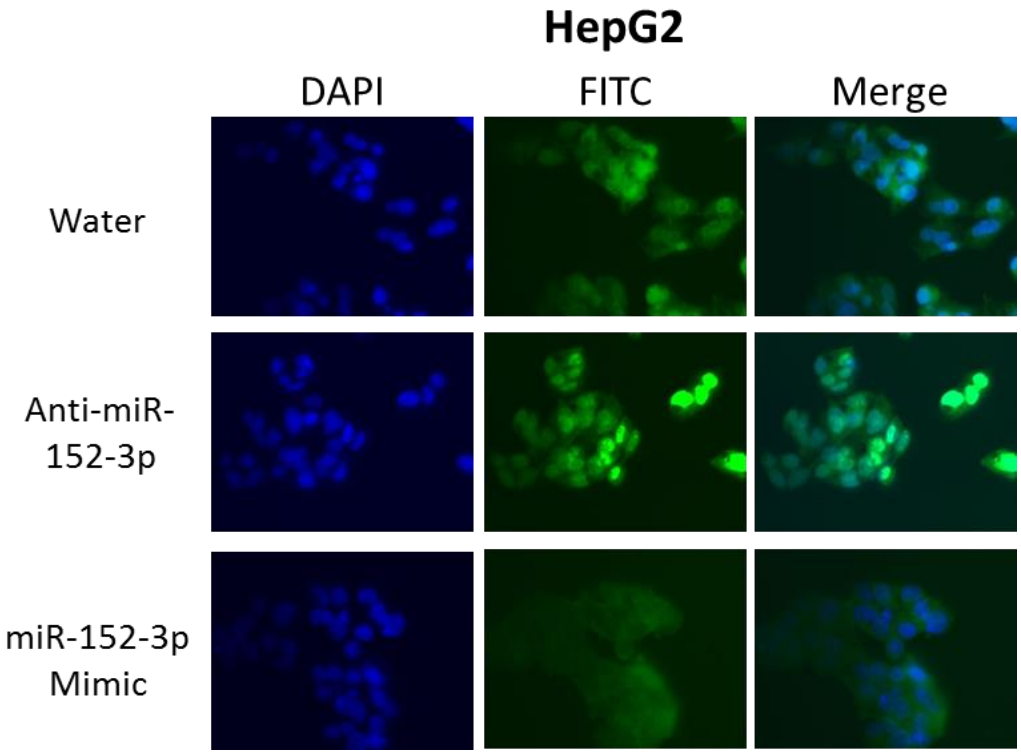


Figure S.1 Transfection parameter optimization results. In order to determine what quantity of transfection reagent to use, HepG2 cells were transfected with the specified quantities of Lipofectamine 2000 using either empty vector (Promega, #e1330) or vector + SETDB1 3'UTR and luminescence was calculated. The 2 μL quantity was chosen for subsequent experiments due to the consistent, robust observed luciferase/*Renilla* ratio.

Inverted Immunofluorescence Images



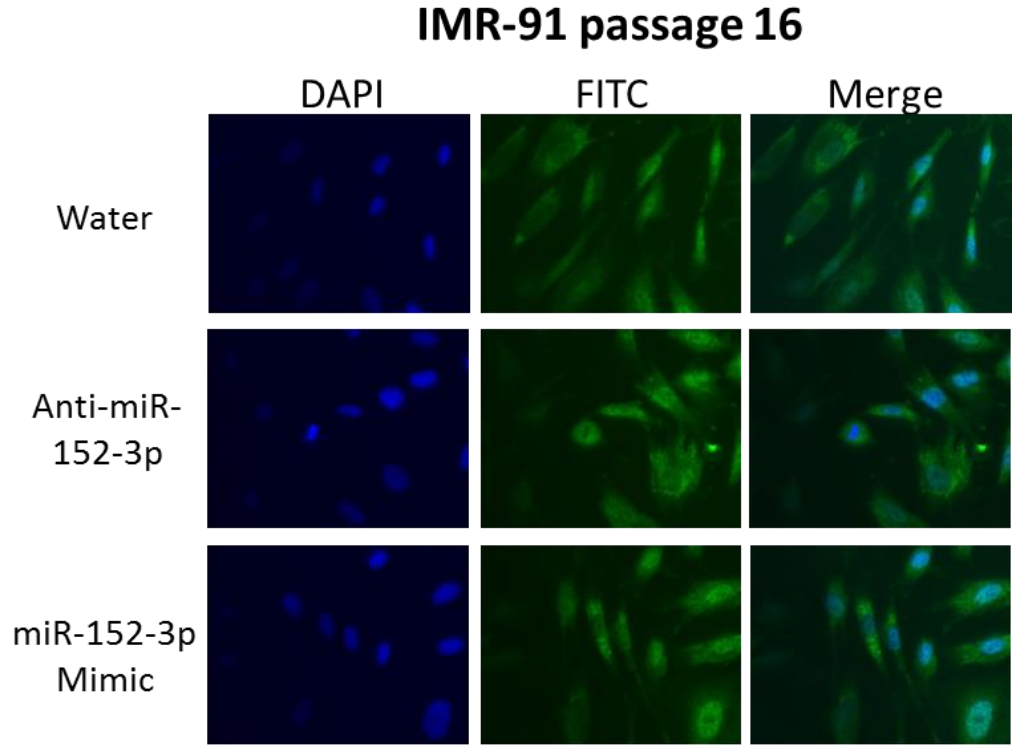


Figure S.4 HP1α visualization under inverted microscopy. Representative 40x magnification inverted microscopy images of HP1α staining in HepG2 and IMR-91 cells treated with water, anti-miR-152-3p, and miR-152-3p with 1.8 second exposure for FITC. Cell staining and image collection performed by Madeline Dorso (2016).

Key		
TRS 1	**TRS 2**	**TRS 3**
Plasmid Sequence	SETDB1 3'UTR Sequence	
Base Substitution	- - - - - - - - - - - -	Deletion

Sequencing Results for Vector + SETDB1 3'UTR (pE1330S3U)

Forward Sequence for pE1330S3U

5'-CGTGATTCTAGTTGTTTAAACGAGGAAGCCTTCTT**CCCAACCCTTCTTGAACTGT**

CGTTTCCTCAGGA**ACTGGGTCTTCCTGATTGT**TGAACCCTGACCCGAAGTCTCTGG

GCTAGCTACTCCCCCCAGCTCCTAGTTGATAGAAATGGGGGTTCTGGACCAG**ATGAT**

CCCTTCCAATGTGGTGCTAGCAGGCAGGATCCCTTCTCCACCTCCAAAGGCCCTAA

AGGGTGGGGAGAGATCACCACTCTAACCTCGGCCTGACATCCCTCCCATCCCATATT

TGTCCAAGTGTTCCTGCTTCTAACAGACTTTGTTCTTAGAATGGAGCCTGTGTATCTA

CTATCTCCAGTTTGTATTATTTCTTGAAAGTCTTTTAACAATATGATAAAACTAAGAT

TGTGCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGA

AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGG

GGCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACA

ACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTA

TTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTT

ATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTAAAACCTCTAC

AAATGTGGTAAAATCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGG

TATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTGCGCAGC

ACCATGGCCTGAAATAACCTCTGAAAGAGGGAACTTGGTTAGGTACCTTTCTGAGGC

GGAAGAACCAGCTGTGGGATGTGTGTCAGTTTAGGGTGTGGGAAAGTCCCCAGGCT

CCCCAGCAGCAGAAGTATGCCAAAGCATGCATCTCATAGTCAGCACCCAAGGTGTG

CAAGTCCCCCAGGCTC-3’

Reverse Sequence for pE1330S3U

5’-TTCCCTGAGTCGATCTAGCACAATCTTAGTTTTATCTATTGTTAAAAGACTTTCA

AGAAATAATACAAACTGGAGATAGTAGATACACAGGCTCCATTCTAAGAACAAAGT

CTGTTAGAAGCAGGAACACTTGGACAAATATGGGATGGGAGGGATGTCAGGCCGAG

GTTAGAGTGGTGATCTCTCCCCACCCTTTAGGGCCTTTGGAGGTGGAGAAGGGATCC

TGCCTGCTA**GCACCACATTGGAAGGGATCAT**CTGGTCCAGAACCCCCATTTCTATC

AACTAGGAGCTGGGGGGAGTAGCTAGCCCAGAGACTTCGGGTCAGGGTTCA**ACAAT**

CAGGAAGACCCAGTTCCTGAGGAAACG**ACAGTTCAAGAAGGGTTGGG**AAGAAG

GCTGTCCTCGTTTAAACAACTAGAATTACACGGCGATCTTGCCGCCCTTCTTGGCCTT

AATGAGAATCTCGCGGATCTTGCGGGCGTCCAACTTGCCGGTCAGTCCTTTAGGCAC

CTCGTCCACGAACACAACACCACCGCGCAGCTTCTTGGCGGTTGTAACCTGGCTGGC

CACATAGTCCACGATCTCCTTCTCGGTCATGGTTTTACCGTGTTCCAGCACGACGACT

GCGGCGGGCAGCTCGCCGGCATCGTCGTCGGGCAGGCCGGCGACCCCGGCGTCGAA

GATGTTGGGGTGTTGCAGCAGGATGCTCTCCAGTTCGGCTGGCGCTACCTGGTAGCC

CTTGTATTTGATCAGGCTCTTCAGCCGGTCCACGATGAAGAAGTGCTCGTCCTCGTC

CCAGTAGGCGATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTTTGT

AGCCTCGGGGTTGTTAACGTAGCCGCTCATGATCATGGGGCCACGGACGCACAGCT

CGCCGCGCTGGTTCACACCCAGTGTCTTACCCGGTGTCCCAAGTCCACCCACCTTAG
CCTCGAGACGCACCCACCCTTGCCCTACTGCGCCAGGCTTTGTCGTCCCCTTTCGGG
GGGGTGAT-3’

Sequencing Results for Vector + SETDB1 3'UTR with Target Recognition Sequence 1 Deletion (pE1330S3UTRSΔ1)

Forward Sequence for pE1330S3UTRSΔ1

5'-CCGTGTATTCAGTTGTTTAAACGAGGCAGCCTTCTT- - - - - - - - - - - - - - - - - - -CGT

TTCCTCAGGA**ACTGGGTCTTCCTGATTGT**TGAACCCTGACCCGAAGTCTCTGGGCT

AGCTACTCCCCCCAGCTCCTAGTTGATAGAAATGGGGGTTCTGGACCAG**ATGATCC**

CTTCCAATGTGGTGCTAGCAGGCAGGATCCCTTCTCCACCTCCAAAGGCCCTAAAG

GGTGGGGAGAGATCACCACTCTAACCTCGGCCTGACATCCCTCCCATCCCATATTTG

TCCAAGTGTTCCTGCTTCTAACAGACTTTGTTCTTAGAATGGAGCCTGTGTATCTACT

ATCTCCAGTTTGTATTATTTCTTGAAAGTCTTTTAACAATATGATAAAACTAAGATTG

TGCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAA

GGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGG

CGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC

TAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTT

GTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATG

TTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTAAAACCTCTACAAA

TGTGGTAAAATCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTAT

TTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTGCGCAGCAC

CATGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGG

AAAGAACCAGCTGTGGAATGTGTGTCAGTTTAGGGTGTGGAAAAGATCCCCCACGC

TTCCCCAGCAGCAGAAGGTAATTGCAAAAGCATGGCATTCTTCCAATTTAGTTCAGG

CCAACCCAGGTTGTTGGTAAAGATCCCCCAGGGCTTCCCCCCAGGCATGCAGGAAG

TTATTGGCAAAAGCATGCCATCTTCAATTAGTTCCAGCAACCATTAGTTCCCGGCCC

CCTTTAAACCTTTCCCGGC-3’

Reverse Sequence for pE1330S3UTRSΔ1

5’-CATGCCTGCAGTCGCTCTGCACAATCTTAGTTTTATCTATTGTTAAAAGACTTTCA

GAAATAATACAAACTGGAGATAGTAGATACACAGGCTCCATTCTAAGAACAAAGTC

TGTTAGAAGCAGGAACACTTGGACAAATATGGGATGGGAGGGATGTCAGGCCGAGG

TTAGAGTGGTGATCTCTCCCCACCCTTTAGGGCCTTTGGAGGTGGAGAAGGGATCCT

GCCTGCTA**GCACCACATTGGAAGGGATCAT**CTGGTCCAGAACCCCCATTTCTATCA

ACTAGGAGCTGGGGGGAGTAGCTAGCCCAGAGACTTCGGGTCAGGGTTCA**ACAATC**

AGGAAGACCCAGTTCCTGAGGAAACG- - - - - - - - - - - - - - - - - - -AAGAAGGCTGTCC

TCGTTTAAACAACTAGAATTACACGGCGATCTTGCCGCCCTTCTTGGCCTTAATGAG

AATCTCGCGGATCTTGCGGGCGTCCAACTTGCCGGTCAGTCCTTTAGGCACCTCGTC

CACGAACACAACACCACCGCGCAGCTTCTTGGCGGTTGTAACCTGGCTGGCCACATA

GTCCACGATCTCCTTCTCGGTCATGGTTTTACCGTGTTCCAGCACGACGACTGCGGC

GGGCAGCTCGCCGGCATCGTCGTCGGGCAGGCCGGCGACCCCGGCGTCGAAGATGT

TGGGGTGTTGCAGCAGGATGCTCTCCAGTTCGGCTGGGGCTACCTGGTAGCCCTTGT

ATTTGATCAGGCTCTTCAGCCGGTCCACGATGAAGAAGTGCTCGTCCTCGTCCCAGT

AGGCGATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTTTGTAGCCT

CGGGGTTGTTAACGTAGCCGCTCATGATCATGGGGCCACGGACGCACAGCTCGCCG

CGCTGGTTCACACCCCAGTGTCTTACCCGGTGTCCAAGTCCACCACCTTAGCCTCGA

AAGAATGGGCACCACCTTTGCCCTACTGCGCCAGGCTTGTCGTCCCCCTTTCGGGGG

GTGATCAGAAATGACGCTTGGTTGTTTCTGTCATGCCGTAGCCTTGGCGATGCCTGG

TAGTGGAAGCGGTTGGCACGGCTCACTACCTCCTGCTGAGCGGCG-3’

Sequencing Results for Vector + SETDB1 3'UTR with Target Recognition Sequence 2 Deletion (pE1330S3UTRSΔ2)

Forward Sequence for pE1330S3UTRSΔ2

5'-CGTGTATCAGTTGTTTAACGAGGAAGCCTTCTT**CCCAACCCTTCTTGAACTGT**CC

GTTTCCTCAGGA- - - - - - - - - - - - - - - - - - TGAACCCTGACCCGAAGTCTCTGGGCTA

GCTACTCCCCCCAGCTCCTAGTTGATAGAAATGGGGGTTCTGGACCAG**ATGATCCC**

TTCCAATGTGGTGCTAGCAGGCAGGATCCCTTCTCCACCTCCAAAGGCCCTAAAGG

GTGGGGAGAGATCACCACTCTAACCTCGGCCTGACATCCCTCCCATCCCATATTTGT

CCAAGTGTTCCTGCTTCTAACAGACTTTGTTCTTAGAATGGAGCCTGTGTATCTACTA

TCTCCAGTTTGTATTATTTCTTGAAAGTCTTTTAACAATATGATAAAACTAAGATTGT

GCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAG

GAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGC

GGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACT

AGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG

TAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGT

TTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTAAAACCTCTACAAAT

GTGGTAAAATCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTATT

TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTGCGCAGCACC

ATGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAA

AGAACCAGCTGTGGAATGTGTGTCAGTTAGGGGTGTGGAAAGTCCCCAGGCTCCCC

AGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAAACCCAGGTGTGG

AAAGTCCCCCAGGGCTCCCCCAGCAGGGCAGAAGTATGCAAAGCATGCATCTCAAA

TTTAAGTCAGCACCATAGTCCCCGCCCTAACCTCGCCATCCCGCCCCTACTACGACC

AGTACGTCCAATTCTTCCGG-3’

Reverse Sequence for pE1330S3UTRSΔ2

5’-AAGCCTGCAGTCGATCTAGCACAATCTTAGTTTTATCATATTGTTAAAAGACTTTC

CAAGAAATAATACAAACTGGAGATAGTAGATACACAGGCTCCATTCTAAGAACAAA

GTCTGTTAGAAGCAGGAACACTTGGACAAATATGGGATGGGAGGGATGTCAGGCCG

AGGTTAGAGTGGTGATCTCTCCCCACCCTTTAGGGCCTTTGGAGGTGGAGAAGGGAT

CCTGCCTGCTA**GCACCACATTGGAAGGGATCAT**CTGGTCCAGAACCCCCATTTCTA

TCAACTAGGAGCTGGGGGGAGTAGCTAGCCCAGAGACTTCGGGTCAGGGTTCA- - - -

- - - - - - - - - - - - - - -TCCTGAGGAAACG**ACAGTTCAAGAAGGGTTGGG**AAGAAGGCT

GTCCTCGTTTAAACAACTAGAATTACACGGCGATCTTGCCGCCCTTCTTGGCCTTAAT

GAGAATCTCGCGGATCTTGCGGGCGTCCAACTTGCCGGTCAGTCCTTTAGGCACCTC

GTCCACGAACACAACACCACCGCGCAGCTTCTTGGCGGTTGTAACCTGGCTGGCCAC

ATAGTCCACGATCTCCTTCTCGGTCATGGTTTTACCGTGTTCCAGCACGACGACTGCG

GCGGGCAGCTCGCCGGCATCGTCGTCGGGCAGGCCGGCGACCCCGGCGTCGAAGAT

GTTGGGGTGTTGCAGCAGGATGCTCTCCAGTTCGGCTGGGGCTACCTGGTAGCCCTT

GTATTTGATCAGGCTCTTCAGCCGGTCCACGATGAAGAAGTGCTCGTCCTCGTCCCA

GTAGGCGATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTTTGTAGC

CTCGGGGTTGTTAACGTAGCCGCTCATGATCATGGGGCCACGGACGCACAGCTCGCC

GCGCTGGTTCACACCCAGTGTCTTACCGGTGTCCAAGTCCACCACCTTAGCCTCGAA

GAAGGCACCACCTTGCCTACTGCGCCAGGCTTGTCGTCCCCTTCGGGGGTGATCAGA

ATGGCGCTGGGTTGTTTCTGTCAGGCCGTAGCCCTGGCGGATGCCCTGGGTAGGTGG

GAAGCGTTTGACCACGGCCTCACCCTACCTACTGCTGGAGCGCCGGCCCCCCGTCGC

TTGCGAATCTCGTGCCAGTGCTAGTCGTACCTGGTCGAATGCAAGAAGACTC-3’

Sequencing Results for Vector + SETDB1 3'UTR with Target Recognition Sequence 3 Deletion (pE1330S3UTRSΔ3)

Forward Sequence for pE1330S3UTRSΔ3

5'-CCGTGTATCAGTTGTTTAAACGAGAAAGCCTTCTT-**CCAACCCTTCTTGAACTGT**

CGTTTCCTCAGGA**ACTGGGTCTTCCTGATTGT**TGAACCCTGACCCGAAGTCTCTGG

GCTAGCTACTCCCCCCAGCTCCTAGTTGATAGAAATGGGGGTTCTGGACCAG------

----------------AGCAGGCAGGATCCCTTCTCCACCTCCAAAGGCCCTAAAGGGTG

GGGAGAGATCACCACTCTAACCTCGGCCTGACATCCCTCCCATCCCATATTTGTCCA

AGTGTTCCTGCTTCTAACAGACTTTGTTCTTAGAATGGAGCCTGTGTATCTACTATCT

CCAGTTTGTATTATTTCTTGAAAGTCTTTTAACAATATGATAAAACTAAGATTGTGCT

AGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAA

GCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCGGCC

GCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAA

TGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAC

CATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCA

GGTTCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTAAAACCTCTACAAATGTG

GTAAAATCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTATTTTCT

CCTTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTGCGCAGCACCATGG

CCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAA

CCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCCAGCAG

GCAGAAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAAACCCAGGTGTGGGAA

AGTCCCCAGGCTCCCCAAGCAAGCCAGAAGTATGCAAAGCCATGGCATTCCTTCCAT

TTAGTTCAGGCACCAATAGTCCCCGGCCCCCCCT-3’

Reverse Sequence for pE1330S3UTRSΔ3

5’- ATGGCCTGCAGTCGCTCTGCACAATCTTAGTTTTATCTATTGTTAAAAGACTTTTA

AGAAATAATACAAACTGGAGATAGTAGATACACAGGCTCCATTCTAAGAACAAAGT

CTGTTAGAAGCAGGAACACTTGGACAAATATGGGATGGGAGGGATGTCAGGCCGAG

GTTAGAGTGGTGATCTCTCCCCACCCTTTAGGGCCTTTGGAGGTGGAGAAGGGATCC

TGCCTGCTA- - - - - - - - - - - - - - - - - - - -CTGGTCCAGAACCCCCATTTCTATCAACTAG

GAGCTGGGGGGAGTAGCTAGCCCAGAGACTTCGGGTCAGGGTTCA**ACAATCAGGA**

AGACCCAGTTCCTGAGGAAACG**ACAGTTCAAGAAGGGTTGGG**AAGAAGGCTGTC

CTCGTTTAAACAACTAGAATTACACGGCGATCTTGCCGCCCTTCTTGGCCTTAATGA

GAATCTCGCGGATCTTGCGGGCGTCCAACTTGCCGGTCAGTCCTTTAGGCACCTCGT

CCACGAACACAACACCACCGCGCAGCTTCTTGGCGGTTGTAACCTGGCTGGCCACAT

AGTCCACGATCTCCTTCTCGGTCATGGTTTTACCGTGTTCCAGCACGACGACTGCGG

CGGGCAGCTCGCCGGCATCGTCGTCGGGCAGGCCGGCGACCCCGGCGTCGAAGATG

TTGGGGTGTTGCAGCAGGATGCTCTCCAGTTCGGCTGGGGCTACCTGGTAGCCCTTG

TATTTGATCAGGCTCTTCAGCCGGTCCACGATGAAGAAGTGCTCGTCCTCGTCCCAG

TAGGCGATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTTTGTAGCC

TCGGGGTTGTTAACGTAGCCGCTCATGATCATGGGGCCACGGGACGCACAGCTCGC

CGCGCTGGTTCACACCCAGTGTCTTACCGGTGTCCAAGTCCACCACCTTAGCCTCGA

AAGAAGGCACCACCCTTGCCTACTGCGCCAGGGCTTTGTCGTCCCCTTTCGGGGGGT

GATCAGTTGGCGCTGATTGTTTTCTGTCAAGAGGTAGCCTGCCGATGCCTGGTAGGT

TGGAGCGTTGCCACCGGTCTCACCCTTAACCTC-3’

Sequence Alignments for Bisulfite Sequencing of CpG Island 56

Experimental methods, data processing, and NCBI BLAST alignments were performed by Heather White (2016).

Bisulfite-Converted Control (Zymo, #D5024) Forward Sequence Alignment

NW Score	Identities	Gaps	Strand
298	178/202(88%)	20/202(9%)	Plus/Plus

```
Query  1    TGAAGGAGGTTACGGGTAAGTCGTTTTGACGTAGACGTTTTATTAGGGTCGCGCGT  56
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    GGAGTGAAGGAGGTTACGGGTAAGTCGTTTTGACGTAGACGTTTTATTAGGGTCGCGCGT  60

Query  57   TCGTCGTTCGTTATATATCGTTCGTAGTATTCGTGTTTAGTTTCGTAGTGGCGTTTGACG  116
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   TCGTCGTTCGTTATATATCGTTCGTAGTATTCGTGTTTAGTTTCGTAGTGGCGTTTGACG  120

Query  117  TCGCGTTCGCGGGTAGTTACGATGAGGCGGCGATAGATTAGGTATAGGGTTTTATCGTTT  176
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  TCGCGTTCGCGGGTAGTTACGATGAGGCGGCGATAGATTAGGTATAGGGTTTTATCGTTT  180

Query  177  TTAAGGGCGAATTCCAGCACAC  198
            ||
Sbjct  181  TT                     182
```

Bisulfite-Converted Control (Zymo, #D5024) Reverse Sequence Alignment

NW Score	Identities	Gaps	Strand
322	180/193(93%)	11/193(5%)	Plus/Plus

```
Query  1    CCCTTGGAGTGAAGGAGGTTACGGGTAAGTCGTTTTGAGGTAGACGTTTTATTASGGTCG  60
            |||||||||||||||||||||||||||||||||| |||||||||||||||| |||||
Sbjct  1          GGAGTGAAGGAGGTTACGGGTAAGTCGTTTTGACGTAGACGTTTTATTAGGGTCG  55

Query  61   CGCGTTCGTCGTTCGTTATATATCGTTCGTAGTATTCGTGTTTAGTTTCGTAGTGGCGTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56   CGCGTTCGTCGTTCGTTATATATCGTTCGTAGTATTCGTGTTTAGTTTCGTAGTGGCGTT  115

Query  121  TGACGTCGCGTTCGCGGGTAGTTACGATGAGGCGGCGATAGATTAGGTATAGGGTTTTAT  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  116  TGACGTCGCGTTCGCGGGTAGTTACGATGAGGCGGCGATAGATTAGGTATAGGGTTTTAT  175

Query  181  CGTTTTTAAGGGC  193
            |||||||
Sbjct  176  CGTTTTT         182
```

Unconverted HepG2 Mock-Treated Region 6 Forward Sequence Alignment

Homo sapiens chromosome 17, alternate assembly CHM1_1.1

Sequence ID: NC_018928.2 **Length:** 81289577 **Number of Matches:** 1

Range 1: 46179574 to 46180109 GenBank Graphics ▼ Next Match ▲ Previous M

Score	Expect	Identities	Gaps	Strand
985 bits(533)	0.0	535/536(99%)	0/536(0%)	Plus/Plus

Features: coatomer subunit zeta-2

```
Query  1         GTATCACAGAACCTGGGCCgggggggACAGCGGGCCGAGCCTCCTTCTTCCAGCTGATCC  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  46179574  GTATCACAGAACCTGGGCCGGGGGGGACAGCGGGCCGAGCCTCCTTCTTCCAGCTGATCC  46179633

Query  61        CTGGCCGGGCTGGACCTGCGCTATCAGCGCGCCCCCAGGCCCTGGCCGGGGTAGACTCCA  120
               |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  46179634  CTGGCCGGGCTGGACCTGCGCTATCAGCGCGCCCCCAGGCCCTGGCCCGGGTAGACTCCA  46179693

Query  121       GAAGCATGCCCATCACCCACCGGTGATGGGAGCCGAGTCTGGGACGGACAGGCCACTCCT  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  46179694  GAAGCATGCCCATCACCCACCGGTGATGGGAGCCGAGTCTGGGACGGACAGGCCACTCCT  46179753

Query  181       TACCCTCCCCGCGGAATCGCAACTCACCCGCCACCCCCACTCCAAACGGACCCAGAACTT  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  46179754  TACCCTCCCCGCGGAATCGCAACTCACCCGCCACCCCCACTCCAAACGGACCCAGAACTT  46179813

Query  241       CAGCCCCGCGCCGACGGACTCAGGACAGGCCTCCTCTCCGGTCGGCGCTGGGACGAACTT  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  46179814  CAGCCCCGCGCCGACGGACTCAGGACAGGCCTCCTCTCCGGTCGGCGCTGGGACGAACTT  46179873

Query  301       TTCCTCCCGGGTCCTGGGGCCAGTCAGGGGTGACACAAAGTTCCCAGCCGCCGGGCGCGC  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  46179874  TTCCTCCCGGGTCCTGGGGCCAGTCAGGGGTGACACAAAGTTCCCAGCCGCCGGGCGCGC  46179933

Query  361       GAGTTCTGAGTTGGCTGCtccccctaaccctgctctgtccctgcccagctcccgccgccc  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  46179934  GAGTTCTGAGTTGGCTGCTCCCCCTAACCCTGCTCTGTCCCTGCCCAGCTCCCGCCGCCC  46179993

Query  421       gggatccccgcgcccgcccgctccgtacccgcagccccgagggctccccggctcgagcag  480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  46179994  GGGATCCCCGCGCCCGCCCGCTCCGTACCCGCAGCCCCGAGGGCTCCCCGGCTCGAGCAG  46180053

Query  481       gcggcgccgggccccccggcctgggccgccgcggcccccctccccccggGTGCGGACGT  536
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  46180054  GCGGCGCCGGGCCCCCGGCCTGGGCCGCCGCGGCCCCCTCCCCCGGGTGCGGACGT  46180109
```

APPENDIX III

PEER-REVIEWED PUBLICATION

"A Standardized Human Embryoid Body Platform for the Detection and Analysis of Teratogens" begins on the following page and is copied in its final form as published.

Citation:

Flamier A, **Singh S**, Rasmussen TP (2017) A standardized human embryoid body platform for the detection and analysis of teratogens. PLoS ONE 12(2): e0171101. doi:10.1371/journal.pone.0171101

A Standardized Human Embryoid Body Platform for the Detection and Analysis of Teratogens

Anthony Flamier[1], Supriya Singh[2] and Theodore. P. Rasmussen[1, 2, 3*]

[1]Department of Pharmaceutical Sciences, University of Connecticut, Storrs, CT, USA;
[2]Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA;
[3]University of Connecticut Stem Cell Institute, Storrs/Farmington, CT, USA.

Running title: TERATOLOGY AND EMBRYOID BODIES

Corresponding author:

Theodore Rasmussen
Associate Professor of Pharmaceutical Sciences
U. Conn. Stem Cell Institute
U. Conn. Institute for Systems Genomics

University of Connecticut
Department of Pharmaceutical Sciences
69 North Eagleville Road
Storrs, Connecticut 06269-3092

*Corresponding author
E-mail: theodore.rasmussen@uconn.edu

Funding

This research was supported Connecticut stem cell research grants 09-SCB-UCON-18 and 11-SCD-I02.

Abstract

Teratogens are compounds that can induce birth defects upon exposure of the developing fetus. To date, most teratogen studies utilize pregnant rodents to determine compound teratogenicity *in vivo*. However, this is a low throughput approach that cannot easily meet the need for comprehensive high-volume teratogen assessment, a goal of the US Environmental Protection Agency. In addition, rodent and human development differ substantially, and therefore the use of assays using relevant human cells has utility. For these reasons, interest has recently focused on the use of human embryonic stem cells for teratogen assessment. Here we present a highly standardized and quantitative system for the detection and analysis of teratogens that utilizes well-characterized and purified highly pluripotent stem cells. We have devised strategies to mass-produce thousands of uniformly sized spheroids of human ESCs (hESCs) that can be caused to undergo synchronous differentiation to yield embryoid bodies (EBs) in the presence and absence of suspected teratogens. The system uses all human cells and rigorously controlled and standardized EB culture conditions. Furthermore, the approach has been made quantitative by using high-content imaging approaches. Our system offers distinct advantages over earlier EB systems that rely heavily on the use on mouse ESCs and EB aggregates of stochastic sizes. Together, our results show that thousands of suspected teratogens could be assessed using human EB-based approaches.

Key words

Teratology; Embryoid Bodies; Embryonic Stem Cells.

Introduction

Annually, 3 to 5% live births in the United States are impacted by birth defects leading to over 100,000 occurrences of birth defects per year [1, 2]. Teratogens include environmental compounds to which women are exposed before or during pregnancy through food, drinking water, airborne sources, and through dermatological exposure. In addition, many pharmaceuticals have teratogenic activities. The U.S. Environmental Protection Agency has a long-standing mission to identify potentially toxic compounds through the Toxcast program which aims to assess thousands of chemicals of potential concern [3]. However, methods to screen vast collections of compounds for teratogenicity are currently laborious and plagued by relatively low throughput, though recent computational approaches have identified a subset of ToxCast compounds, many with potentially teratogenic effects [4].

Teratogens are compounds that cause birth defects, and these can be either pharmaceutical agents to which conceptuses are exposed *in utero*, or environmental compounds. Thalidomide is a sedative used widely form 1957 to 1961 during pregnancy, and over 10,000 infants were born with birth defects after thalidomide exposure in utero [5-8]. These unfortunate individuals presented with limb abnormalities and neural tube closure defects [7]. Sodium valproate is a widely prescribed anti-epileptic and mood stabilizing agent which is also a histone deacetylase inhibitor and known teratogen [9-13]. Though women who may potentially become pregnant are advised to avoid valproate, unintentional exposure may occur during the first days of pregnancy at the zygote, morula, blastocyst, and gastrulation stages prior to detection of pregnancy.

Pharmaceuticals and xenobiotic compounds have been traditionally tested for teratogenicity in pregnant animal models and rodent embryo culture [14], and human cell culture models for teratogenicity are not well developed at present. However, recent innovative approaches include the use of embryonic stem cells to monitor metabolomic changes in response to toxicants [15], and assessments of teratogenic effects in zebrafish, where embryonic development can be monitored by virtue of a transparent fish body plan [16].

Traditional animal assays for teratogens are excellent (since the readout is based on direct teratological impacts on developing embryos), but practical difficulties exist with these approaches. The first impediment is that animal teratological assays are very low throughput. In addition, there is concern that species-specific differences between rodent and human prenatal development impact the interpretation of suspected teratogens. Furthermore, in mammals (including both rodents and humans), both male and female germ line development occurs during embryonic development. For this reason, when dosing a pregnant rodent, the mother, the developing embryo, and the developing embryonic gonadal primordia are simultaneously dosed. Thus, it is not immediately apparent whether the effects of a compound act directly upon the mother, the fetus, or the offspring of the exposed fetus, thus creating a conundrum that requires the assessment of at least 2 generations of offspring for birth defects.

For the reasons above, recent interest has focused on the development of cell culture-based systems to detect teratogens. ESCs can be cultured indefinitely and coaxed to differentiate in vitro into a multitude of terminally-differentiated cell types, thus making them an attractive choice for use in teratology assays. Initial attempts with ESCs relied upon direct exposure of undifferentiated mouse ESCs to potential teratogens followed by assays of differentiation. More recently, approaches have been developed that use ESC aggregates to initiate the formation of embryoid bodies (EBs). Standard methods for producing EBs involve partially disaggregating mouse or human ESC colonies to yield agglomerates of dozens to hundreds of cells, and these aggregates are cultured in media that lack signalling molecules that preserve the pluripotent state (i.e. LIF in the case of mouse ESCs and basic FGF (b-FGF) in the case of human ESCs). Under these conditions, EBs grow and spontaneously differentiate into cellular lineages that contain cells of ectodermal, endodermal, and mesodermal identities [17-21]. Together, these and many other studies show that differentiating EBs come to contain a great many cell types derived from all three of the major germ layers, and that the accumulation of these cells can be monitored by cell biological assays. Furthermore, developing EBs exhibit the induction of a wide array of germ layer-specific gene expression patterns. However, monitoring complex alterations of cellular content and gene expression in EBs developing in the presence of compounds is labor intensive, and EB approaches using uncontrolled aggregation yield EBs of random size and shape, making the assessment of EB growth difficult to quantitate and of low throughput. We

aimed to improve the throughput of initial EB teratogen screens by developing standardized systems to produce uniformly sized EBs combined with synchronous differentiation.

Here we present a standardized human EB system for the detection and modeling of teratogen action. Our goals were to develop a highly standardized system, with uniformly-sized human EBs assembled from pristinely pluripotent and quality-controlled hESC cultures. Our aim was to improve the quality and accuracy of a human EB system for the detection of teratogens that could be easily adapted to automated high content imaging solutions, with improved throughput so that in the near future, hundreds to thousands of compounds can be expediently assessed for teratological risk.

Materials And Methods

Cell culture

H9 hESCs were obtained from WiCell and cultured on hESC-qualified Matrigel (Corning #354277). For all assays, we utilized a highly characterized lot of H9 hESCs curated in the Rasmussen Laboratory established from an initial freeze at passage 19. Media was changed every day with mTeSR1 (Stem Cell technologies #05857) and mechanically passaged every 4-6 days. Embryoid bodies were formed and maintained in EB Aggrewell medium (Stem Cell Technologies #05893). Fibroblast cell lines IMR-90 and IMR-91 (Coriell, #I90-83 and #I91S-04) were grown at 37°C with 5% CO2 on 0.1% gelatin-coated cell culture-treated plates under aseptic conditions. Culture media consisted of 4.5 g/L glucose DMEM (Invitrogen, #10938025) supplemented with 10% FBS (Atlanta Biologicals, #S10250H) and 100X Glutamax (Invitrogen, #35050061).

Cytotoxicity Assays

IMR-90 (I90-83) and IMR-91 (I91S-04) cells were plated on day zero at passages 10 and 14, respectively. Fibroblasts were treated in quadruplicate with 0 mM, 0.1 mM, 1 mM, 10 mM, 100 mM, and 1 M concentrations of VPA, caffeine, and penicillin G dissolved in culture media on day 2. On day 3, cells were measured for viability using an MTS assay (Promega, #G3580) in a SpectraMax Paradigm plate reader from Molecular Devices.

Flow Cytometry

H9 hESCs were incubated with Accutase (Invitrogen, #A11105-01) for 5 minutes or until partial detachment and passed through a 40 μm cell strainer. Single cells were then pelleted and resuspended with fresh 4% paraformaldehyde (PFA). After 10 minutes of incubation at room temperature, cells were exposed to methanol for 15 minutes on ice. Following cell fixation and permeabilization, the cells were incubated with primary antibodies for 30 minutes on ice using Stain buffer (BD Biosciences, #554657): Mouse α-human Tra1-60-FITC (BD Biosciences, #560380), Mouse α-human SSEA4 PE (BD Biosciences, #560128), Mouse α-human Oct3/4

AlexaFluor 647 (BD Biosciences, #560329), Mouse α-human Nanog-PerCP-Cy5.5 (BD Biosciences, #562259).

10,000 events were collected using FACS Calibur instrument (BD Biosciences) after compensation using CompBeads plus (BD Biosciences, #560497). All data were analyzed using FlowJo software.

Live cell staining

Live stains of H9 colonies were performed using StainAlive α-SSEA-4 antibody (DyLight 550) (Stemgent, #09-0097). Briefly, fresh mTeSR1 media containing 10 μg of StainAlive α-SSEA-4 antibody was applied to cultures and incubated 30 minutes at 37°C. After two washes with fresh warm mTeSR1 media, hESC cultures were analyzed on an inverted fluorescence microscope.

Magnetic immunoselection of pluripotent cells

Magnetic purification was conducted using a TRA-1-60 selection kit (Miltenyi Biotec, #130-100-832) following manufacturer's instructions. Briefly, H9 cells at the day of harvest were incubated with Accutase (Invitrogen, #A11105-01) for 5 minutes or until partial detachment, passed through a 40 μm cell strainer and resuspended into fresh medium. $2x10^6$ single cells were then incubated for 5 minutes at 4°C with anti-TRA-1-60 MicroBeads with ROCK inhibitor. Cell-bead conjugates were then collected on magnetic columns under magnetic field and TRA-1-60 positive cells were collected using 1 ml of culture medium.

Embryoid body formation

Single cell suspensions of TRA-1-60 positive cells were deposited into AggreWell plates (Stem Cell Technologies) and allowed to aggregate into spheroids. Briefly, $1x10^6$ cells positive for TRA-1-60 were suspended in 5 ml of EB Aggrewell medium (Stem Cell technologies #05893) and added into a well of an AggreWell400Ex plate (Stem Cell Technologies, # 27840) containing approximately 4,700 microwells with a diameter of 400 μm and pretreated with AggreWell rinsing solution (Stem Cell Technologies, # 07010). AggreWell plates were centrifuged at 100 x g for 3 minutes to deposit the cells in the microwells and incubated at 37°C

for 24 hours, during which time spherical aggregates of deposited pluripotent cells (spheroids) were formed.

qEB assay

Spheroid aggregates were collected by firmly pipetting in the AggreWell plates and EB cultures were initiated by depositing spheroids into low-adherence culture plates treated with either DMSO (vehicle only), Valproic acid (VPA) (Sigma-Aldrich, #P4543-10G; 1mM), caffeine (Sigma-Aldrich, #C0750-5G; 0.26mM) or penicillin G (Sigma-Aldrich, #P3032-1MU; 2.8mM).

Gene expression analysis

RNA was extracted using the RNeasy Mini kit (Qiagen, # 74104) according to manufacturer instructions. 1μg of RNA was further converted into complementary DNA using iScript cDNA Synthesis Kit (Bio-Rad, #1708890). PCR reactions were performed using GoTaq green master mix (Promega. #M7112) and loaded on a 1% agarose gel.

Primer sequences are:

Nestin_F: 5'-GAAGGGCAATCACAACAGGT-3'

Nestin_R: 5'-GGGGCCACATCATCTTCCA-3'

B-III tub F: 5'-GGCCAAGGGTCACTACACG-3'

B-III tub R: 5'-GCAGTCGCAGTTTTCACACTC-3'

NEFH F: 5'-GTGAAGAGTGTCGGATTGGCT-3'

NEFH R: 5'-ACACAGAGGGAATTTTGGGGA-3'

Brach_F: 5'-TATGAGCCTCGAATCCACATAGT-3'

Brach_R: 5'-CCTCGTTCTGATAAGCAGTCAC-3'

Myf5 F: 5'-CTGCCAGTTCTCACCTTCTGA-3'

Myf5 R: 5'-AACTCGTCCCCAAATTCACCC-3'

c-actin F: 5'-GTACCCTGGTATTGCTGATCG-3'

c-actin R: 5'-CCTCATCGTACTCTTGCTTGCT-3'

Results

Purification and assessment of highly pluripotent human embryonic stem cells

We aimed to develop a cell culture system that can detect teratological impacts on human embryonic cells so we chose to use hESCs as they can be coaxed to differentiate into a wide variety of endodermal, ectodermal, and mesodermal differentiated cell types. However, working with pluripotent cells can be challenging since hESCs can undergo undesired alterations during cell culture that can reduce their pluripotency [20, 22]. Thus we devised a battery of pluripotency assays based on both pluripotency markers and functional assessments of pluripotency so that hESC cultures of validated quality could be used. We also developed strategies to purify highly pluripotent cells from initial hESC cultures for subsequent use in teratology assays.

We first assessed cultures of hESCs for 4 key markers of pluripotency. These consisted of two cell surface markers (TRA-60 and SSEA4) and two nuclear markers (OCT4 and NANOG) which were chosen because they are master regulatory transcription factors necessary for the maintenance of the pluripotent state in hESCs and induced pluripotent stem cells human (hiPSCs). Though we routinely stain such cells for these markers by immunofluorescence, we wished to quantitatively assess these markers using flow cytometry to more rigorously assess the expression of pluripotency-associated markers in our starter cell cultures (Fig 1A). To achieve this goal, we dissociated cultures to single cells, fixed, and permeablized the cells so that nuclear the nuclear markers (OCT4 and NANOG) are accessible to antibodies, while preserving the surface markers TRA1-60 and SSEA4. We found that human hESC line H9 (WA09) is strongly positive for all 4 markers as compared to control cells stained with non-specific IgG. As expected, undifferentiated H9 cells were positive for TRA1-60 (>95%), SSEA4 (>75%), NANOG (>95%) and OCT4 (>95%) when compared to IgG stained control cells (Fig 1A). To functionally investigate the suitability of these markers, H9 cells were differentiated by removing basic FGF for 9 days, after which only few cells remained positive for TRA1-60 (<10%),

NANOG (<5%) and OCT4 (<10%) (Fig 1A). However, we found that SSEA4 is a less robust marker of pluripotency since after 9 days of differentiation about 50% of cells retained SSEA4 expression. We also performed two channel flow cytometry on undifferentiated H9 cells and found that 96% of H9 cells were doubly positive for pluripotency markers TRA1-60 and OCT4 (Fig 1B). We also quantitated the decreases in pluripotency marker expression caused by removal of basic FGF from the medium for 9 days. All markers declined robustly upon differentiation, except SSEA4, whose expression was the slowest to respond to removal of basic FGF (Fig 1C). These studies revealed that TRA1-60 is a superior surface marker of pluripotency as compared to SSEA4, and that both nuclear transcription factors (NANOG and OCT4) were expressed only in the undifferentiated cells.

To further assess the behaviour of SSEA4 in pluripotent cell cultures, we noticed that larger hESC colonies at passage >50 exhibit two patterns of SSEA4 marker expression (S1A Fig). Using immunofluorescence, we found that cells at colony peripheries were highly positive for SSEA4, while cells more central to the colony have lower levels of SSEA4. We then microdissected colonies to isolate peripheral cells (SSEA4 HIGH) and central cells (SSEA4 LOW) and initiated new colonies from the two microdissected subpopulations. In both cases, colonies arose that still exhibited SSEA4 HIGH cells on their peripheries, suggesting an intrinsic regulation of SSEA4 expression that is a function of the position of cells within hESC colonies (S1B Fig). To further investigate this observation, we performed flow cytometry on higher passage cells and found two peaks (S1C Fig), corresponding to SSEA4 HIGH and SSEA4 LOW cells. To further investigate molecular differences between both populations, we isolated SSEA4 HIGH and SSEA4 Low cells using magnetic beads cell purification to positively select for SSEA4 expression cells (S1D Fig). We isolated total RNA from SSEA4 HIGH and LOW cells and performed semi-quantitative reverse transcriptase PCR (RT-PCR) and found that both SSEA4 HIGH and LOW cells still exhibited similar expression levels of OCT4, NANOG, REX1, and hTERT (S1C Fig). We also made standard (random sized) EBs from SSEA4 LOW and HIGH cells, and found that EBs derived form both cell populations robustly expressed mesodermal or ectodermal genes after 17 days of EB differentiation (S1D Fig). We conclude that TRA1-60 is a superior cell surface marker of pluripotency as compared to SSEA4. In addition to the above, we wished to perform functional tests of pluripotency by differentiating

cells to ectoderm, endoderm, and mesoderm lineages. We used a kit containing supplements and growth factors allowing the directed differentiation into the three germinal layers in 4 days and the staining of SOX17 (endoderm), OTX2 (ectoderm) and BRACHYURY (mesoderm) (RnD Systems; cat. no. SC027). We found that early passage H9 cells can readily form cells of all three germ layers by this functional test (data not shown).

Development a standardized and quantitative human EB cell culture system suitable for high-content imaging

hESCs and EBs derived from them have been used as cell culture models to assess toxicological responses to exogenous compounds, but most of these studies have used mouse ES cells and EBs of random size and shape. Though these past approaches have been useful, we wished to develop a highly standardized and quantitative method to detect and analyze teratogens that could be adapted to high-content imaging strategies using mass-produced human EBs of uniform size and shape. One design consideration for us was to insure that only the very best (most highly pluripotent) cells are used to form EBs. We therefore made use of our findings outlined in the section above, and utilized early passage (passage 19) H9 hESCs for this purpose. Furthermore, we found it prudent to dissociate hESC colonies to single cells using accutase-mediated disaggregation to single cells followed by purification of TRA1-60 positive cells by magnetic bead selection (Fig 2A), so that we could be sure that only highly purified populations of hESCs would be used for the initiation of EB cultures. Though we used only highly pluripotent low passage H9 hESCs, we decided to routinely purify TRA1-60 positive cells by this method to initiate all subsequent EB cultures. This strategy insures that starting cell populations used to initiate EB formation are devoid of any contaminating differentiated cells, and this approach also provides a good measure of quality control of the cells used to initiate EB culture.

We also wished to mass-produce EBs of initial uniform size to make the EB cultures uniform and synchronous in their growth and differentiation. To achieve this goal, we plated disaggregated and TRA1-60 positive (magnetically purified) cells into Aggrewell plates (Stem

Cell Technologies) at an initial seeding density of $2x10^5$ cells per ml in EB culture medium (Stem Cell Technologies). The pyramidal wells of Aggrewell plates are hydrophobic and prevent cellular adhesion to the substratum. We found that after 3 days of aggregation under these conditions, we could produce about 4000 spherical aggregates (spheroids) of hESCs, which were used to initiate EB culture. (Fig 2B). Using these approaches, we were able to produce thousands of uniformly sized human EBs from cultures of hESC line H9 using starting cell populations that are highly uniform and pluripotent, as the cells were positively purified using magnetic beads that bind TRA1-60$^+$ pluripotent cells (Fig 3). Once we developed the capability to mass-produce uniform spheroids, we wished to explore alternative means of EB culture and quantification to render the system more versatile of teratogen testing. Uniformly sized and spherical day 0 EBs (spheroids) were harvested from Aggrewell plates, and further cultured in three different ways: First, we deposited individual day 0 EBs into individual wells of low adherence multi-well plates containing differentiation medium using a 200 μl-tip cut at the extremity. The individually plated EBs were allowed to grow and differentiate for an additional 12 days, and since single EBs were deposited into each well, we could monitor the growth and behavior of individual EBs as a function of time. Secondly, we plated pools consisting of hundreds of uniform day 0 spheroids into low adherence 6 cm dishes and allowed them to differentiate for 12 days. This strategy allowed us to monitor EB growth on populations of EBs so that statistically rigorous assessments of mean EB size (growth) could be obtained. Lastly, we plated EBs into Matrigel-coated wells. This strategy allows EBs to adhere to the cell culture substrate, which allows cells to proliferate and grow outward from the core of attached EBs. These combined approaches allowed us to quantitatively assess the growth of EBs under 3 conditions, using cell culture strategies that could easily be adapted for high-content image analysis.

Performance of the quantitative human EB model with selected teratogens

To assess the sensitivity and accuracy of our EB system (hereafter called the human quantitative embryoid body teratology model system, hqEB system, outlined in Fig. 2) we chose

to use three compounds known to have teratogenic activities. We chose to study caffeine (a weak teratogen), penicillin-G (a moderate teratogen), and Valproic acid (VPA, highly teratogenic). Caffeine has been studied in the context of embryonic development and was reported to impair normal chick neurodevelopment especially affecting eye formation [23, 24]. Treatment of pluripotent stem cells with caffeine also dysregulates the transcription of important early differentiation genes [19]. Penicillin-G was implicated in kidney development defects in the rat and has been widely studied for its developmental neurotoxicity [25, 26]. Neural teratogenicity has been observed after treatment of human embryonic stem cells with VPA [27]. VPA exposure during pregnancy can cause ear malformation [28]. Our goal was to determine if the hqEB system can robustly and quantitatively assess teratogenicity based on the potential effects of these reference teratogens in our newly developed system.

Determination of caffeine, penicillin-G and valproic acid sub-cytotoxic concentrations and teratogen assay dosing

In principle, EB growth kinetics could be affected by either simple cellular cytotoxicity, or due to teratogenic effects that alter cellular proliferation and differentiation. Since we were more interested in teratogenicity as opposed to overt cytotoxicity, we performed dose-response cytotoxicity assays to determine the LC_{50} for each of the three compounds. We first determined by dose response the cytotoxicity of the drugs at five concentrations in two human fibroblast cell lines: I90-83 (also called IMR-90) and I91S-04 (also called IMR-91) (Fig S2A). The results indicated that both fibroblast cell lines had similar LC_{50} values upon exposure to each teratogen. This dosing data allowed us to choose teratogen concentrations for use in the hqEB system that do not elicit overt cytotoxic effects. The concentrations chosen for hqEB analyses in this study were 0.26 mM, 2.8 mM and 1 mM for caffeine, penicillin-G and VPA respectively. These concentrations are 10-fold or more below the LC_{50} for each compound. To further control for potential cytotoxicity of the compounds at these concentrations, we treated hESC H9 cells for 72 hours at these sub-cytotoxic concentrations. We found no significant impact on hESC cell colony growth or morphology after 72 hours of treatment (S2B Fig).

Individual EB monitoring

We first investigated the use of individually cultured, non-adherent EBs exposed to the reference teratogens. First, we prepared hESC spheroids using Aggrewell plates as described above, using magnetically purified TRA1-60^{+} cells to initiate spheroid formation. After 3 days of aggregation (hereafter designated as day 0), spheroids of uniform size and shape were deposited individually into single wells of low adherence 24-well plates, the media was changed from mTeSR1 to EB differentiation medium, and 3 reference teratogens were added at the established sub-cytotoxic concentrations. Four individual EBs were grown in the presence of each teratogen (quadruplicate). We analyzed the growth and shape of individual EBs by microscopy every 3 days for a total of 12 days in differentiation medium. (Fig 4A). Control EBs that were not treated with teratogens exhibited consistent growth over the course of the experiment. In contrast, exposure to VPA greatly curtailed the growth of EBs. After only 3 days of teratogen exposure, we observed that the growth of VPA-treated EBs was arrested, and this failure to grow persisted throughout the course of the experiment (Fig 4A,B). In the case of caffeine (weakly teratogenic), we observed a modest but detectable effect on EB growth, which became apparent and statistically significant at days 9 and 12. The effect of penicillin-G (moderately teratogenic) was stronger than that of caffeine, but less severe than VPA. EBs treated with Penicillin-G failed to increase significantly in size after 6 days of treatment (Fig 4A,B). We used microscopy imaging software to quantitatively assess impacts on the shape of EBs. To do this, we calculated the circularity coefficient for each time point (expressed as circularity = 4π(area/perimeter^2)). This measure showed that shape alterations can also be detected in the hqEB system. We found subtle effects on EB circularity using this approach in the case of caffeine (Fig 4C). Together these results show that teratogenic effects can be readily detected by the hqEB system. We note that these findings suggest that the hqEB system could easily be adapted to utilize automated high-content imaging solutions in future larger scale applications.

Additional analyses of teratogen-exposed EBs

In addition to individual EB monitoring, we wanted to see if similar results could be obtained by analyzing pools of EBs exposed to teratogens. To do this, dishes containing at least 100

unattached EBs of the same size were treated with the three compounds for 12 days and analyzed by microscopy every three days (Fig 5A). The diameters and growth kinetics of EBs were consistent with the previous assay, and VPA teratogenicity was again detected after only three days of teratogen exposure, while EBs treated with Penicillin-G stopped growth by 9 days after aggregation (Fig 5B). Using this approach, caffeine treatment exhibited little or no effect on EB growth (Fig 5B). EBs were then analyzed by qRT-PCR for induction of selected germ layer gene expression (Fig 5C). We found that mesoderm and ectoderm genes were induced and robustly expressed by day 12. We found that caffeine treatment had no major effects on expression of these genes, while penicillin-G treatment had some effect on expression of these genes. Lastly, we wished to see if we could track compound teratogenicity in culture conditions in which spheroids are allowed to attach to the culture substratum. We therefore aggregated H9 hESCs for 3 days and plated the spheroids onto Matrigel-coated plates concurrently with change to differentiation medium. As expected, we noted outgrowths of differentiated cells the spread out from the centers originally formed at the sites of EB attachment. The outgrowths were irregular in shape, and after 72 h of treatment with the three compounds, no phenotypic differences were observed (Fig 5D). Our results indicate that monitoring individual EBs grown under non-adherent conditions constituted an approach that robustly and quantitatively detected teratogenic effects. In contrast, we found the adherent EB strategy less useful, because outgrowth area and shape are more variable.

Discussion

We developed and refined a human EB system that utilizes highly characterized cells of demonstrated pluripotency in such a way that the effects of teratogens can be evaluated and quantitated by numerical descriptors with statistical accuracy. Since teratogens impact the course of directed cellular differentiation during embryogenesis, we strived for utilizing only the very highest quality of cells to initiate EB cultures. To this end, we assessed hESCs rigorously for their pluripotency-associated marker expression. We found that TRA1-60 is a superior surface pluripotency marker as compared to SSEA4. Furthermore, we wished to initiate our EB cultures with spherical aggregates of hESCs, which we found could be mass-produced in aggrewell plates. To do this, we faced a challenge in that suspensions of single cells are required in order to seed aggrewell plates. This is complicated by the fact that cell-cell contacts (mediated by E-Cadherins in hESCs) are required to maintain the pluripotent state of hESCS. We therefore developed a strategy whereby feeder cell-free starter cultures of hESCs are disaggregated to single cells by Accutase treatment. The singularized cells were maintained in the presence of ROCK inhibitor, which can stabilize the maintenance of pluripotency of single hESCs during the disaggregation of hESC colonies to single cells. We also immunologically purified TRA1-60^{+} cells using magnetic beads and used these to seed aggrewell plates to prepare large numbers of uniform spheroids that could in turn be used to initiate EB cultures. We found that we can prepare large numbers of EBs by this method, which are of initial uniform size, and whose growth and proliferation occur synchronously. We found that by simply monitoring a cohort of EBs treated with teratogens (at doses that are sub-cytotoxic), that we could observe defects in EB growth characteristics that could be quantified precisely. We were able to reproducibly and quantitatively detect effects on EB development using a panel of three reference teratogens. This report builds on previous successes with human EBs, but our system makes use of highly standardized EBs of known pluripotency which are of standardized and uniform size and pluripotency prior to differentiation in the presence of potentially teratogenic compounds.

The hqEB system has features that make it attractive for future assessments of teratogen action. First, our system utilizes relevant human cells of embryonic cellular identity (hESCs), which model the inner cell mass of blastocyst-stage preimplantation embryos. Furthermore, EBs

differentiate in vivo to form primitive cells of ectodermal, endodermal, and mesodermal lineages. Also, hESCs and EBs derived from them contain intact complete genomes, and offer obvious advantages as compared to immortalized cell lines. We also initiated our EBs with known numbers using TRA1-60^{+} immunopurified hESC starter cells and found that EB growth (and the impacts of teratogen exposure) can be monitored continuously and assessed quantitatively. The hqEB system is of much higher throughput than costly animal experiments. Though these are useful features, the EB system also has limitations. Firstly, metabolites of parent compounds may be responsible for teratogenicity *in vivo*, and the EB system presented here utilizes relevant embryonic cells (the targets of teratogens) but these cell types cannot metabolize parent compounds. Thus the hqEB system is at present not suited to the assessment of unknown metabolites of parent compounds. In addition, EB systems in general contain early embryonic cells, and therefore later developmental effects cannot easily be modeled with such cells. Therefore, future uses of the hqEB system and its future iterations should be of good use, but it is unlikely that its use will fully supplant the use of animals studies, which could better detect subtle and late-term impacts on embryo health (though with the caveat that rodent embryonic development is clearly not the same as human embryonic development).

In the future, it should be relatively straight forward to adapt the hqEB system to automated high content imaging, since individual EBs can be cultured individually in multi-well plates and assessed periodically by automated cell culture microscopy approaches and associated image acquisition and analysis software. We also note that our uniform EBs could also be cultured in multiplexed microfluidics devices in which continuous perfusion and compound dosing can be performed. It is also noteworthy that complex mixtures (including drinking water samples) could readily be assessed by these approaches. For these reasons, we suspect that improved human EB methods such as those described here will gain added importance in future large scale assays of teratogens. Furthermore, the system would be useful during safety assessment stages of development of new pharmaceuticals. The approaches described in this study should allow the assessment of hundreds of potential teratogens since the assay is based on morphological assays of EBs that are initially uniform, develop synchronously, and that can be assessed at intervals over the course of differentiation. Furthermore, our approaches are readily adaptable to microfluidics and high-content imaging. Once prospective teratogens are identified by these expedient means, it should then be possible to perform more rigorous studies of teratogenic

mechanisms by investigating the impacts of compounds on the cellular composition of impacted EBs (including their relative content and identities of cells of ectodermal, endodermal, and mesodermal origin) as well as detailed analyses of altered gene and protein expression, which collectively, should shed light on specific developmental pathways impacted by specific teratogens.

Acknowledgements

This research was supported in part by grants from the Connecticut Regenerative Medicine Research Fund grants nos. 11SCDI02 and 06SCA034.

Declaration of Interest

The authors confirm that there is no conflict of interest that affects the impartiality of the research reported herein.

References

1. Mai CT, Isenburg J, Langlois PH, Alverson CJ, Gilboa SM, Rickard R, et al. Population-based birth defects data in the United States, 2008 to 2012: Presentation of state-specific data and descriptive brief on variability of prevalence. Birth Defects Res A Clin Mol Teratol. 2015;103(11):972-93. doi: 10.1002/bdra.23461. PubMed PMID: 26611917; PubMed Central PMCID: PMCPMC4769882.
2. Update on Overall Prevalence of Major Birth Defects --- Atlanta, Georgia, 1978--2005. Morbidity and Mortality Weekly Report (MMWR). 2008;57(1):1-5.
3. Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ. The ToxCast program for prioritizing toxicity testing of environmental chemicals. Toxicological sciences : an official journal of the Society of Toxicology. 2007;95(1):5-12. doi: 10.1093/toxsci/kfl103. PubMed PMID: 16963515.
4. Sipes NS, Martin MT, Reif DM, Kleinstreuer NC, Judson RS, Singh AV, et al. Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. Toxicological sciences : an official journal of the Society of Toxicology. 2011;124(1):109-27. Epub 2011/08/30. doi: 10.1093/toxsci/kfr220. PubMed PMID: 21873373.
5. Franks ME, Macpherson GR, Figg WD. Thalidomide. Lancet. 2004;363(9423):1802-11. Epub 2004/06/03. doi: 10.1016/S0140-6736(04)16308-3. PubMed PMID: 15172781.
6. Gaffield W, Incardona JP, Kapur RP, Roelink H. A looking glass perspective: thalidomide and cyclopamine. Cell Mol Biol (Noisy-le-grand). 1999;45(5):579-88. Epub 1999/10/08. PubMed PMID: 10512190.
7. Holmes LB. Teratogen-induced limb defects. American journal of medical genetics. 2002;112(3):297-303. Epub 2002/10/03. doi: 10.1002/ajmg.10781. PubMed PMID: 12357474.
8. Matthews SJ, McCoy C. Thalidomide: a review of approved and investigational uses. Clinical therapeutics. 2003;25(2):342-95. Epub 2003/05/17. PubMed PMID: 12749503.
9. Lammer EJ, Sever LE, Oakley GP, Jr. Teratogen update: valproic acid. Teratology. 1987;35(3):465-73. Epub 1987/06/01. doi: 10.1002/tera.1420350319. PubMed PMID: 3114906.
10. Nau H. Valproic acid-induced neural tube defects. Ciba Foundation symposium. 1994;181:144-52; discussion 52-60. Epub 1994/01/01. PubMed PMID: 8005022.
11. Nau H, Hauck RS, Ehlers K. Valproic acid-induced neural tube defects in mouse and human: aspects of chirality, alternative drug development, pharmacokinetics and possible mechanisms. Pharmacology & toxicology. 1991;69(5):310-21. Epub 1991/11/01. PubMed PMID: 1803343.
12. Ornoy A. Valproic acid in pregnancy: how much are we endangering the embryo and fetus? Reprod Toxicol. 2009;28(1):1-10. Epub 2009/06/06. doi: 10.1016/j.reprotox.2009.02.014. PubMed PMID: 19490988.

13. Smith J, Whitehall J. Sodium valproate and the fetus: a case study and review of the literature. Neonatal network : NN. 2009;28(6):363-7. Epub 2009/11/07. PubMed PMID: 19892633.

14. Steele CE. Whole embryo culture and teratogenesis. Hum Reprod. 1991;6(1):144-7. Epub 1991/01/01. PubMed PMID: 1874948.

15. Kleinstreuer NC, Smith AM, West PR, Conard KR, Fontaine BR, Weir-Hauptman AM, et al. Identifying developmental toxicity pathways for a subset of ToxCast chemicals using human embryonic stem cells and metabolomics. Toxicology and applied pharmacology. 2011;257(1):111-21. Epub 2011/09/20. doi: 10.1016/j.taap.2011.08.025. PubMed PMID: 21925528.

16. Padilla S, Corum D, Padnos B, Hunter DL, Beam A, Houck KA, et al. Zebrafish developmental screening of the ToxCast Phase I chemical library. Reprod Toxicol. 2012;33(2):174-87. Epub 2011/12/21. doi: 10.1016/j.reprotox.2011.10.018. PubMed PMID: 22182468.

17. Krueger WH, Tanasijevic B, Barber V, Flamier A, Gu X, Manautou J, et al. Cholesterol-secreting and statin-responsive hepatocytes from human ES and iPS cells to model hepatic involvement in cardiovascular health. PLoS One. 2013;8(7):e67296. doi: 10.1371/journal.pone.0067296. PubMed PMID: 23874411; PubMed Central PMCID: PMCPMC3708950.

18. Krueger WH, Tanasijevic B, Norris C, Tian XC, Rasmussen TP. Oct4 promoter activity in stem cells obtained through somatic reprogramming. Cell Reprogram. 2013;15(2):151-8. doi: 10.1089/cell.2012.0059. PubMed PMID: 23550731; PubMed Central PMCID: PMCPMC3616451.

19. Mayshar Y, Yanuka O, Benvenisty N. Teratogen screening using transcriptome profiling of differentiating human embryonic stem cells. J Cell Mol Med. 2011;15(6):1393-401. doi: 10.1111/j.1582-4934.2010.01105.x. PubMed PMID: 20561110; PubMed Central PMCID: PMCPMC4373338.

20. Tanasijevic B, Dai B, Ezashi T, Livingston K, Roberts RM, Rasmussen TP. Progressive accumulation of epigenetic heterogeneity during human ES cell culture. Epigenetics. 2009;4(5):330-8. PubMed PMID: 19571681.

21. Tanasijevic B, Rasmussen TP. X chromosome inactivation and differentiation occur readily in ES cells doubly-deficient for macroH2A1 and macroH2A2. PLoS One. 2011;6(6):e21512. doi: 10.1371/journal.pone.0021512. PubMed PMID: 21738686; PubMed Central PMCID: PMCPMC3127949.

22. Chan EM, Ratanasirintrawoot S, Park IH, Manos PD, Loh YH, Huo H, et al. Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat Biotechnol. 2009;27(11):1033-7. doi: 10.1038/nbt.1580. PubMed PMID: 19826408.

23. Ma ZL, Wang G, Cheng X, Chuai M, Kurihara H, Lee KK, et al. Excess caffeine exposure impairs eye development during chick embryogenesis. J Cell Mol Med. 2014;18(6):1134-43. doi: 10.1111/jcmm.12260. PubMed PMID: 24636305; PubMed Central PMCID: PMCPMC4508153.

24. Ma ZL, Qin Y, Wang G, Li XD, He RR, Chuai M, et al. Exploring the caffeine-induced teratogenicity on neurodevelopment using early chick embryo. PLoS One. 2012;7(3):e34278. doi: 10.1371/journal.pone.0034278. PubMed PMID: 22470550; PubMed Central PMCID: PMCPMC3314624.

25. Nathanson S, Moreau E, Merlet-Benichou C, Gilbert T. In utero and in vitro exposure to beta-lactams impair kidney development in the rat. J Am Soc Nephrol. 2000;11(5):874-84. PubMed PMID: 10770965.

26. Kuegler PB, Zimmer B, Waldmann T, Baudis B, Ilmjarv S, Hescheler J, et al. Markers of murine embryonic and neural stem cells, neurons and astrocytes: reference points for developmental neurotoxicity testing. ALTEX. 2010;27(1):17-42. PubMed PMID: 20390237.

27. Colleoni S, Galli C, Gaspar JA, Meganathan K, Jagtap S, Hescheler J, et al. A comparative transcriptomic study on the effects of valproic acid on two different hESCs lines in a neural teratogenicity test system. Toxicol Lett. 2014;231(1):38-44. doi: 10.1016/j.toxlet.2014.08.023. PubMed PMID: 25192806.

28. Van Houtte E, Casselman J, Janssens S, De Kegel A, Maes L, Dhooge I. Middle and inner ear malformations in two siblings exposed to valproic acid during pregnancy: a case report. Int J Pediatr Otorhinolaryngol. 2014;78(11):2007-10. doi: 10.1016/j.ijporl.2014.08.030. PubMed PMID: 25216807.

Figures

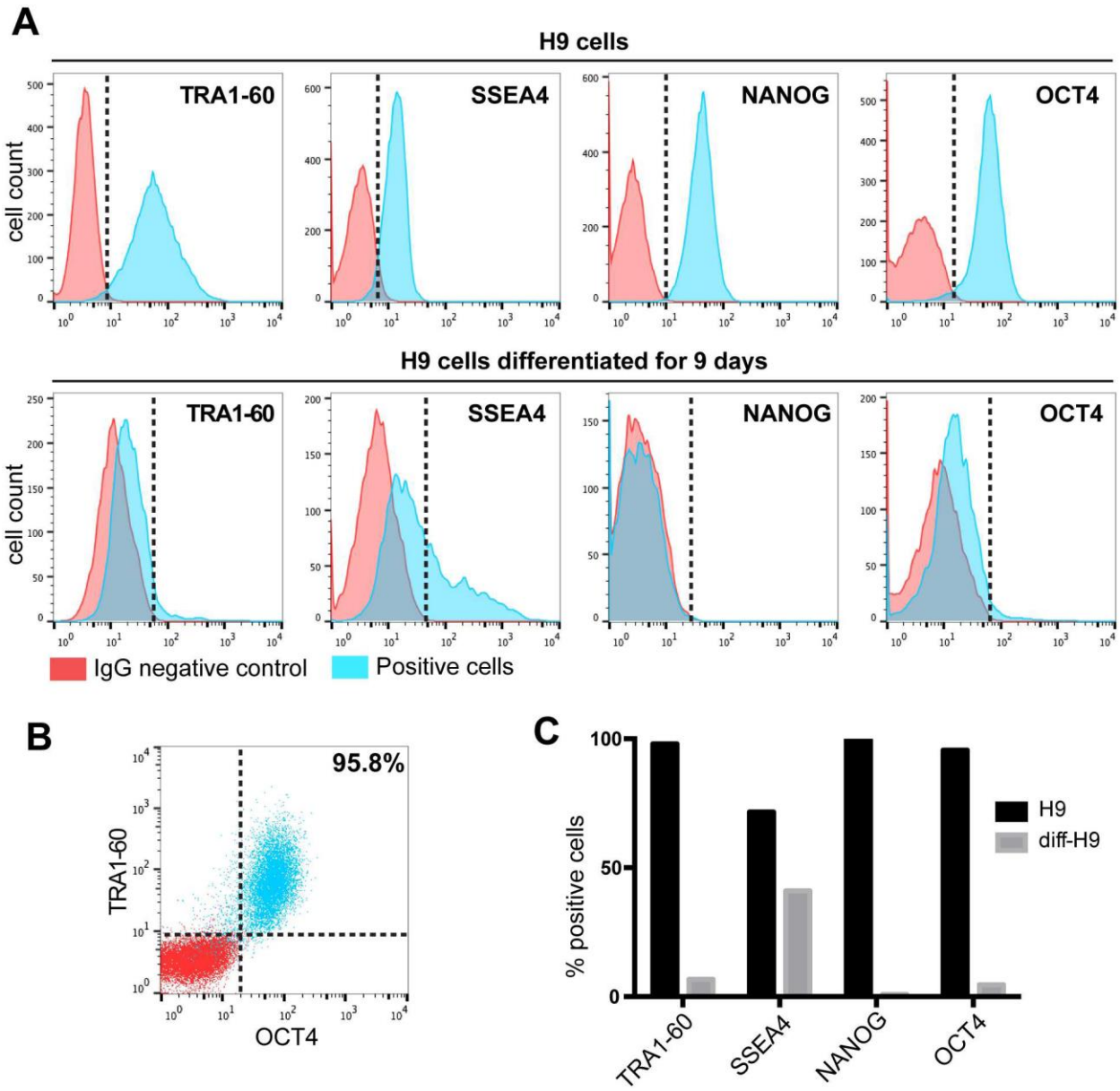


Fig 1. Simultaneous detection of four pluripotency markers by flow cytometry. (A) Flow cytometric analysis of permeablized H9 hESCs. Cells were exposed to antibodies against TRA1-60, SSEA4, NANOG and OCT4. IgG antibody was used as negative control. Top: Cytometry profiles of undifferentiated H9 cells. Bottom: Cytometry histograms of H9 cells 9 days after removal of bFGF. (B) Flow cytometric analysis of undifferentiated H9 cells for the co-stained for TRA1-60 and OCT4. (C) Quantification of positive cells for TRA1-60,

SSEA4, NANOG and OCT4 in comparison to IgG control for both undifferentiated and differentiated H9 cells.

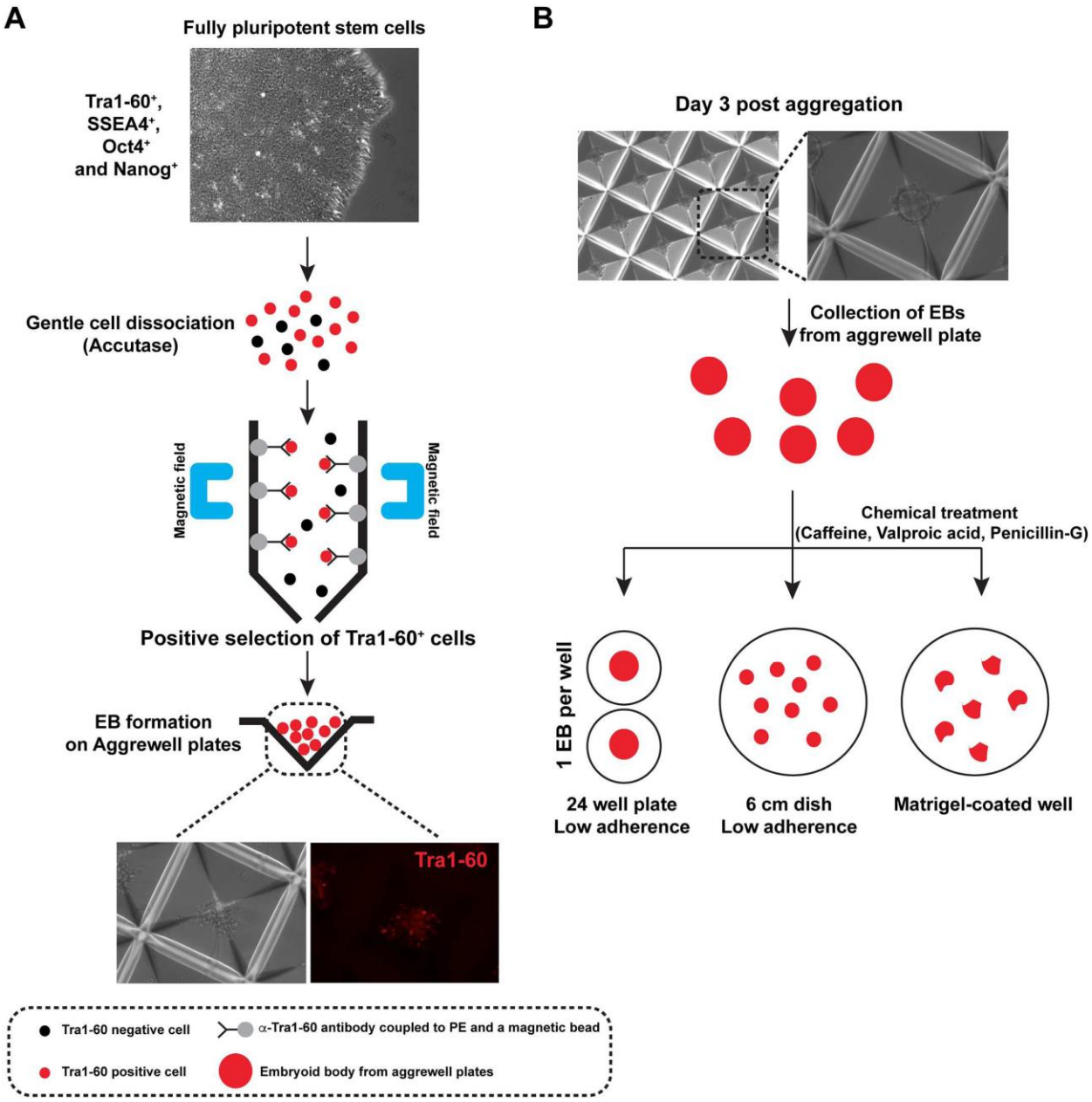


Fig 2. Outline of human quantitative embryoid body (hqEB) system strategies. (A) Preparation of spheroids from TRA1-60+ magnetically purified hESCs. Low passage H9

cells were dissociated to single cells by Accutase treatment, then TRA1-60+ cells were magnetically purified. These were then plated at known density into Aggrewell plates to from spheroids of uniform size. (B) Alternative plating and differentiation approaches. Three methods were assessed: (1) Spheroids were deposited individually into cell culture plate wells, and imaged every 3 days. (2). Pools of EBs are plated into 6 well dishes under non-adherent conditions, imaged, quantitated for size. (3) Spheroids are attached onto Matrigel coated plates and allowed to proliferate and spread under adherent conditions.

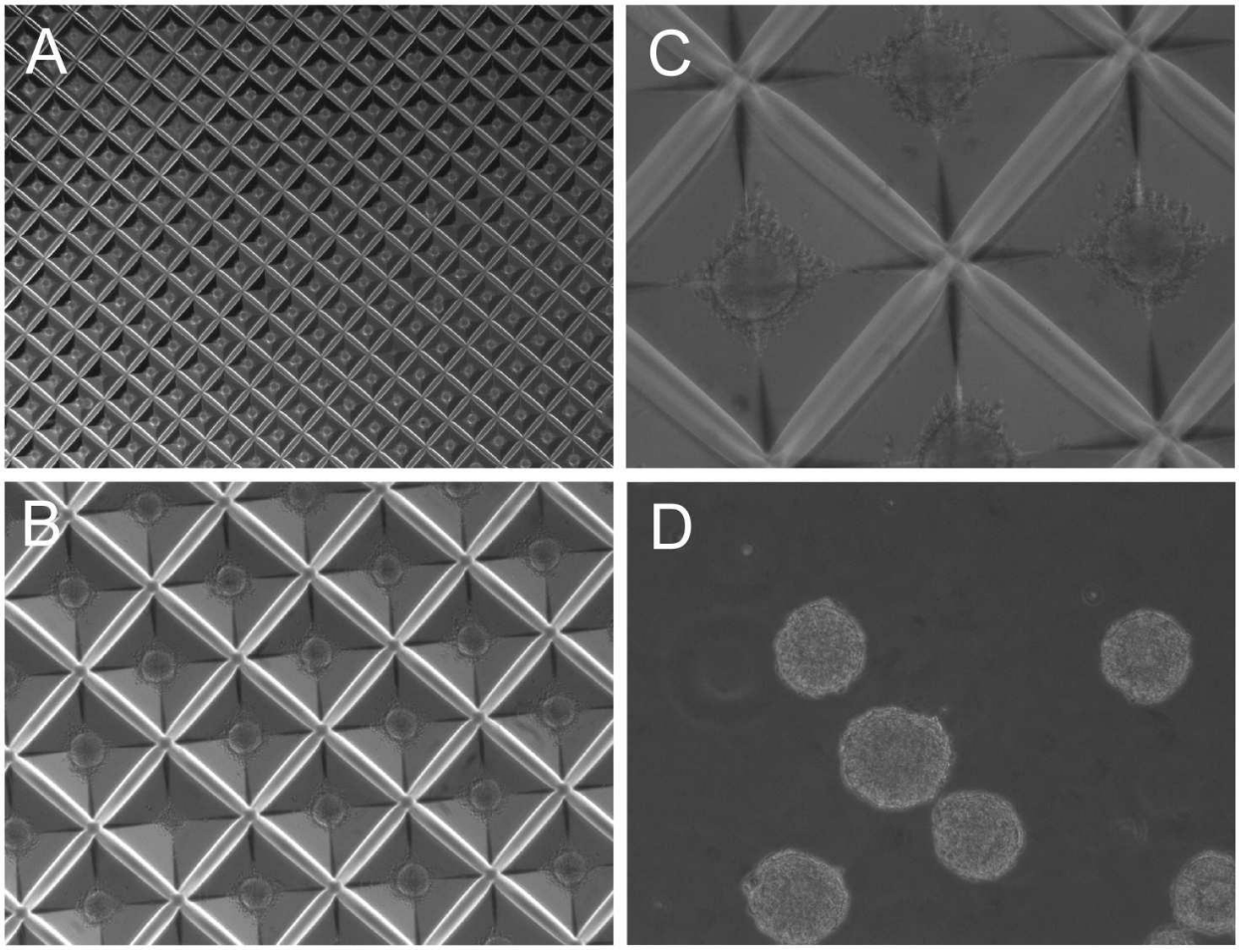


Fig 3. Representative images of hESC spheroids prepared from TRA1-60+ purified cells.

Spheroids containing uniform numbers of cells were prepared in Aggrewell plates using TRA1-60+ cells. (A) Aggrewell plate containing spheroids aggregated for 3 days (1.25X magnification). (B) Aggrewell plate containing spheroids aggregated for 3 days (4X magnification). (C) Aggrewell plate containing spheroids aggregated for 3 days (10 X magnification). (D) Spheroids imaged immediately after aggregation.

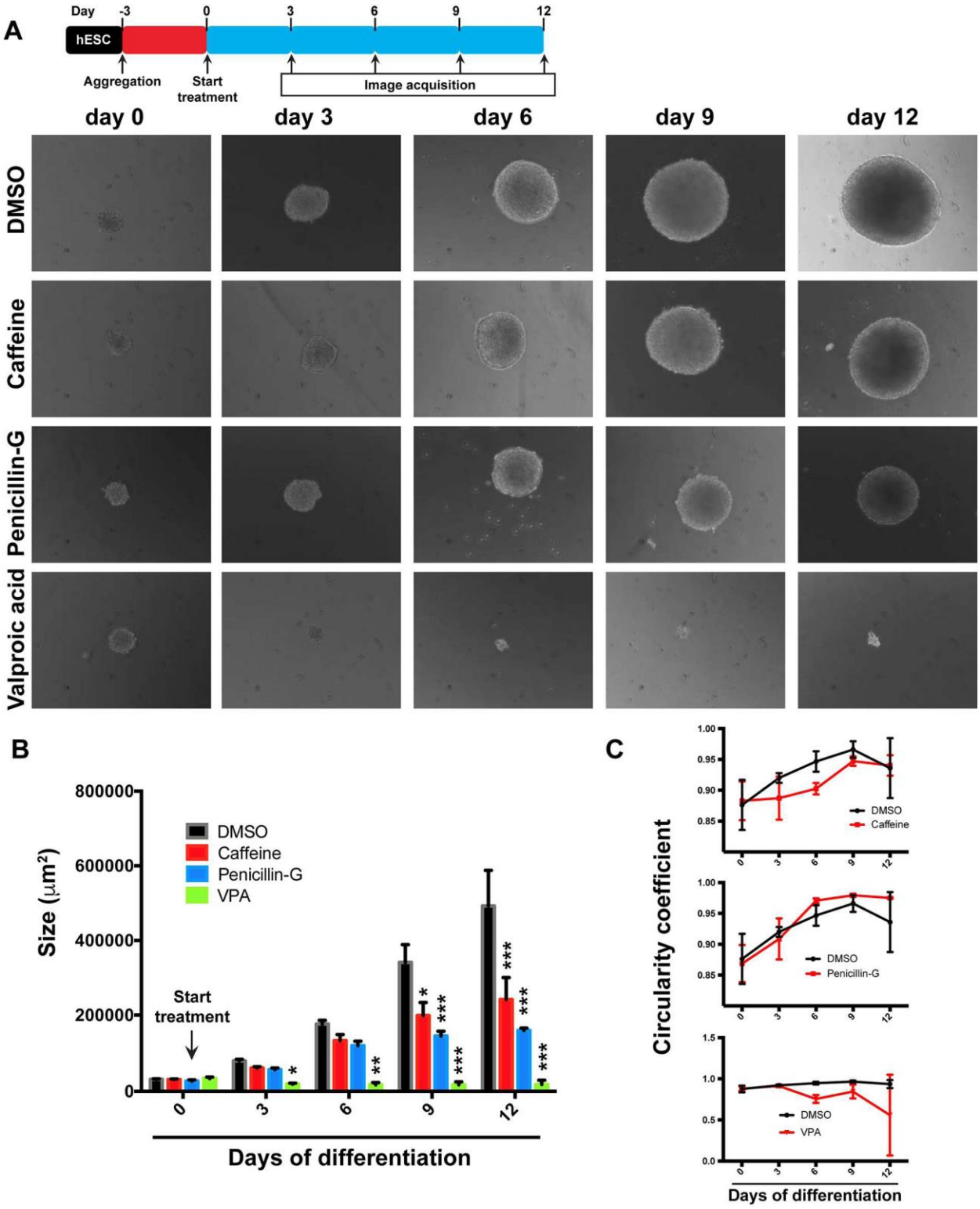


Fig 4. Consequences of teratogen exposure on individually cultured EBs. (A) H9 hESCs were dissociated with Accutase, formed into spheroids, and plated into individual wells

under non-adherent conditions (day 0). Media was changed to EB differentiation medium and individual EBs were imaged microscopically every 3 days for a total of 12 days of differentiation. Individual EBs were treated with vehicle only (DMSO), Caffeine, Penicillin-G, and Valproic acid and images were acquired after 0, 3, 6, 9, and 12 days of differentiation. (B) EBs were quantitated by measuring the apparent size (μm2) of individual embryoid bodies after treatment with DMSO, Caffeine, Penicillin-G or Valproic acid. Results shown indicate the means and standard deviations by quantitatively imaging a cohort of 3 independently cultured EBs each tested with each of the 3 compounds. (*P<0.05; **P<0.01; ***P<0.001 as compared to DMSO). (C) Shape quantification expressed by the circularity coefficient of individual EBs treated with DMSO, Caffeine, Penicillin-G or Valproic acid. Results shown indicate the means and standard deviations obtained by quantitatively imaging a cohort of 3 independently cultured EBs each tested with each of the 3 compounds.

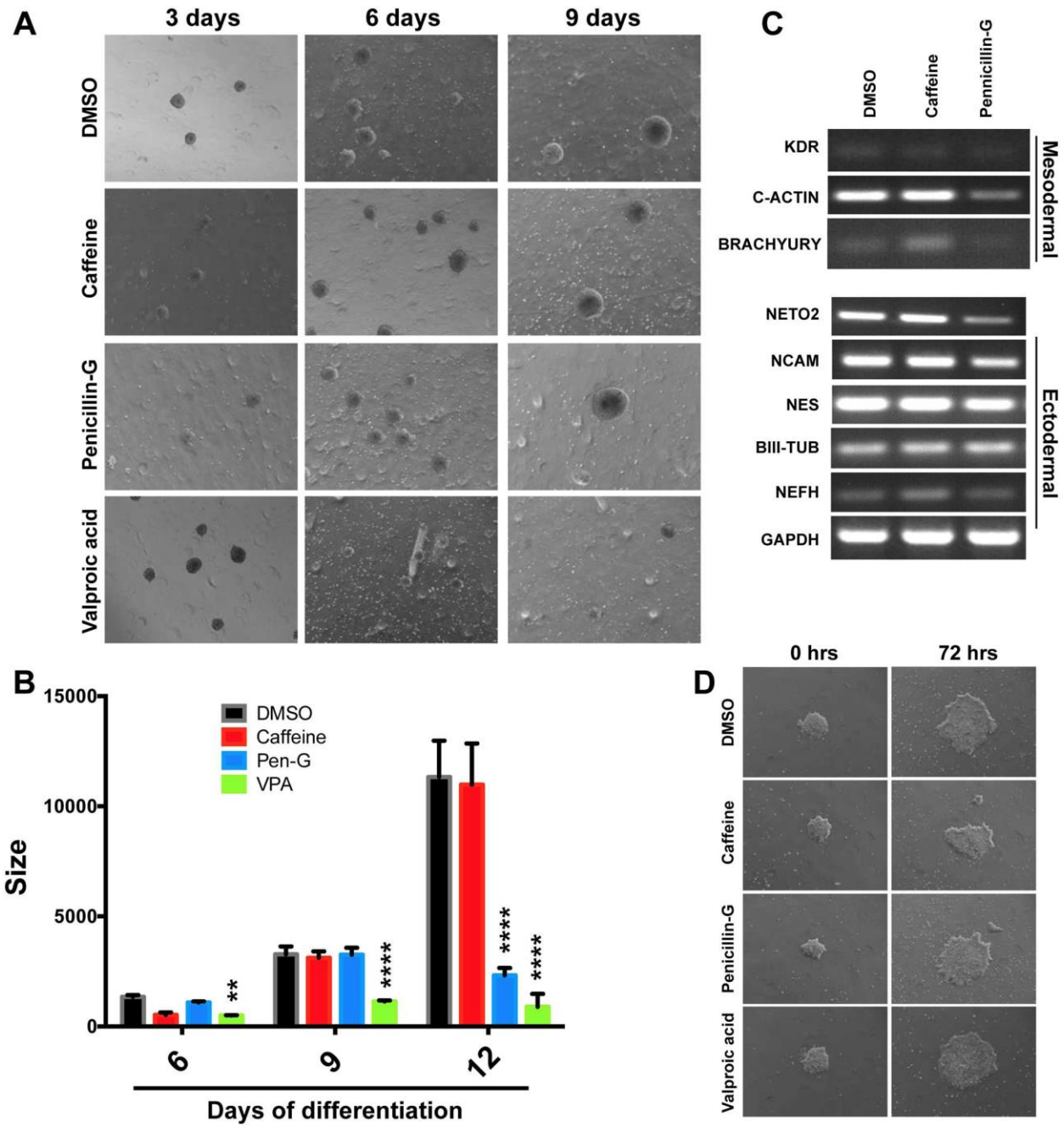
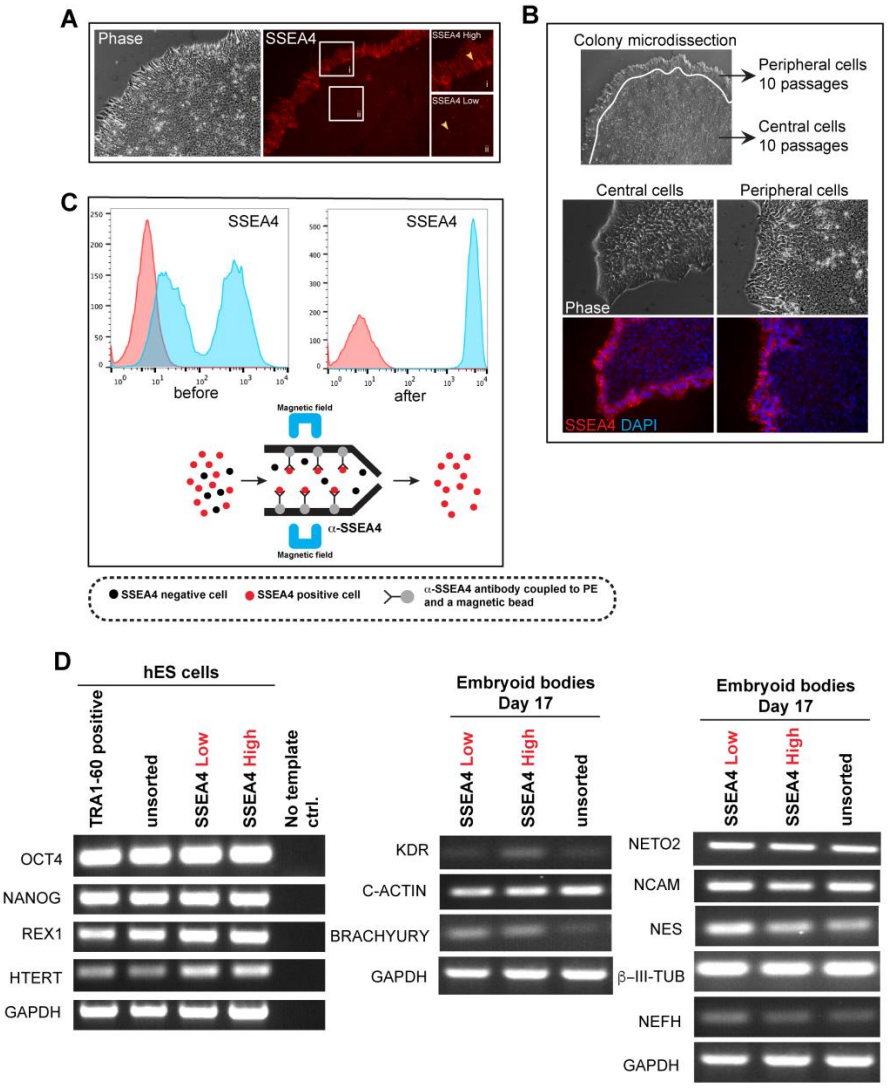


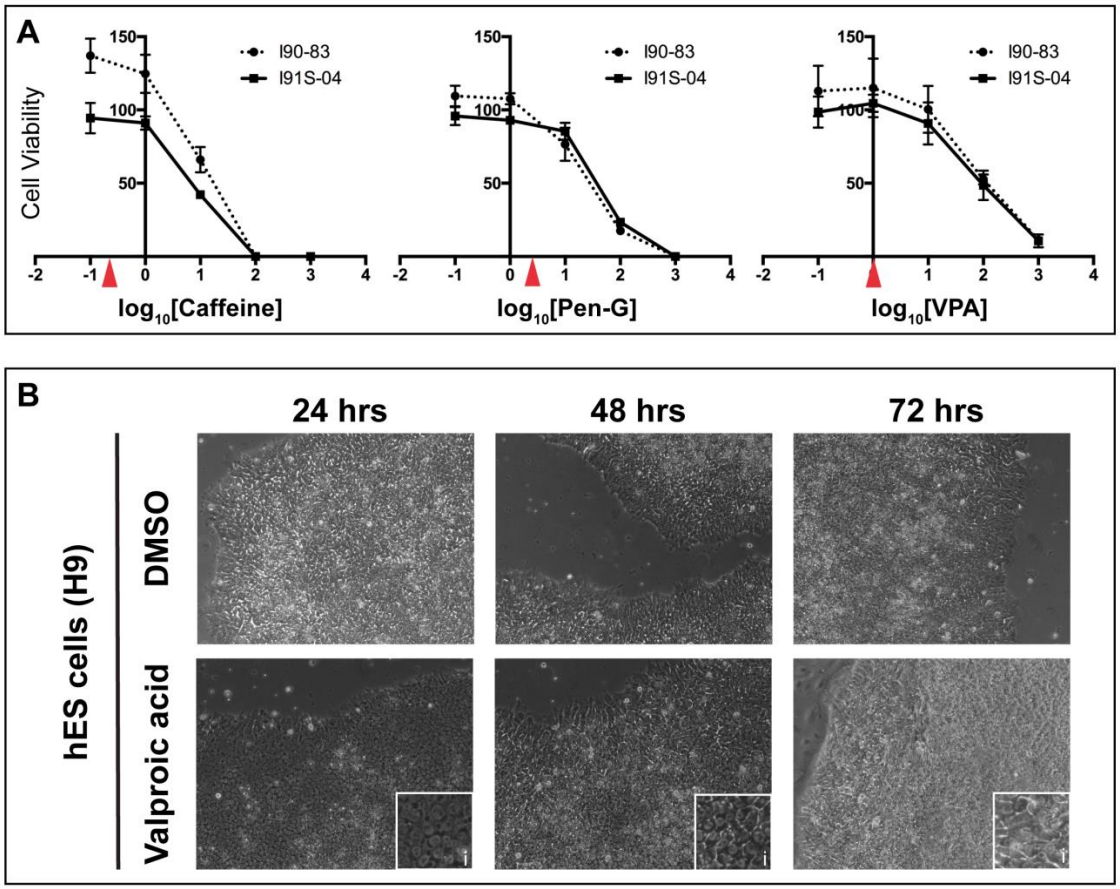
Fig 5. Monitoring EB growth in pooled populations of EBs grown under non-adherent conditions. (A) Representative images of EBs grown in the presence of DMSO (vehicle only), Caffeine, Penicillin-G and Valproic acid. (B) Size quantification of EB populations treated with DMSO, Caffeine, Penicillin-G or Valproic acid. Results are the means +/- standard deviation. (n = 17 to 69 independent EBs after 3, 6, and 9 days of culture in the

presence of teratogens (*P<0.05; **P<0.01; ***P<0.001 as compared to DMSO). (C) Analysis of mRNA expression of differentiation markers from EBs treated with DMSO, Caffeine or Penicillin-G over the course of 12 days of EB differentiation. RT-PCR performed to detect mesodermal (KDR, C-ACTIN and BRACHYURY) and ectodermal mRNAs (NETO2, NCAM, NES, BIII-TUB and NEFH). GAPDH is used as loading control. (D) EB spreading assay performed on attached EBs cultured on Matrigel (0 hours) and treated with DMSO, Caffeine, Penicillin-G or Valproic acid for 72 hours in differentiation medium.



S1 Fig. Analysis of pluripotency marker expression in hESC line H9. (A) High passage H9 cell colony live stained with α-SSEA4 antibody coupled to PE. i and ii (insets) show colony regions expressing high (colony peripheries) and low (colony interior) levels of SSEA4 respectively. (B) Microdissection of H9 cell colonies to isolate peripheral cells (SSEA4 HIGH) and central cells (SSEA4 LOW). Bottom: representative colonies initiated with SSEA4 HIGH and LOW dissected cells after 10 passages. (C) Flow cytometry analysis of high passage H9 cells for the stem cell marker SSEA4. Note the presence of two populations of cells expressing different levels of SSEA4 as compared to IgG negative

control. H9 colonies were dissociated to single cells and incubated with an anti-SSEA4 antibody coupled to magnetic beads. Immunopurified cells express high levels of SSEA4. (D) Pluripotency marker expression analysis on hESC line H9 sorted for TRA1-60 or SSEA4 (low and high expression). RT-PCR was used to detect OCT4, NANOG, REX1 and hTERT mRNAs and GAPDH was used as loading control. To determine if the expression of SSEA4 status impacts differentiation, H9 EBs were made from hESCs expressing high or low levels of SSEA4 and then differentiated for 17 days. RT-PCR performed on mRNA isolated from the EBs and assessed for expression of KDR, C-ACTIN, BRACHYURY (mesoderm) and NETO2, NCAM, NES, BIII-TUB and NEFH (ectoderm). GAPDH is used as an internal loading control.



S2 Fig. Determination of sub-cytotoxic compound doses. (A) Dose-response curves of teratogens two human fibroblast cell lines (I90-83 and I91S-04) exposed to caffeine, penicillin-G, and Valproic acid. Red arrow indicates the concentration chosen for the hqEB system. (n = 3 independent experiments). (B) Representative images of hESC line H9 treated with DMSO or Valproic acid for 24, 48 and 72 hours. Insets show higher magnification images of cellular morphology within colonies.